POLITYKA GOSPODARCZA

Wybrane zagadnienia

Dariusz K. Rosati

SGH

OFICYNA WYDAWNICZA SGH
SZKOŁA GŁÓWNA HANDLOWA W WARSZAWIE

POLITYKA GOSPODARCZA

Wybrane zagadnienia

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

POLITYKA GOSPODARCZA

Wybrane zagadnienia

Dariusz K. Rosati

SGH

OFICYNA WYDAWNICZA SGH
SZKOŁA GŁÓWNA HANDLOWA W WARSZAWIE
WARSZAWA 2017

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Recenzent
Andrzej Wojtyna

Redakcja językowa
Patrycja Czarnecka

© Copyright by Dariusz K. Rosati & Szkoła Główna Handlowa w Warszawie, Warszawa 2016
Wszelkie prawa zastrzeżone. Kopiowanie, przedrukowywanie i rozpowszechnianie całości lub fragmentów niniejszej publikacji bez zgody wydawcy zabronione.

Wydanie I

ISBN 978-83-8030-148-1

Oficyna Wydawnicza SGH – Szkoła Główna Handlowa w Warszawie
02-554 Warszawa, al. Niepodległości 162
www.wydawnictwo.sgh.waw.pl
e-mail: wydawnictwo@sgh.waw.pl

Projekt i wykonanie okładki
Monika Trypuz
Zdjęcie na okładce
Włodzimierz Gajda

Skład i łamanie
DM Quadro

Druk i oprawa
QUICK-DRUK s.c.
e-mail: quick@druk.pdi.pl

Zamówienie 69/IV/17

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Spis treści

Przedmowa 15

Rozdział 1. Podstawy polityki gospodarczej 19
1.1. Wprowadzenie 19
1.2. Niesprawności rynku 21
1.2.1. Zasada Pareto i pierwsze podstawowe twierdzenie ekonomii dobrobytu . 21
1.2.2. Konkurencja nie jest doskonała 27
1.2.3. Niekompletność rynków 29
1.2.4. Efekty zewnętrzne 30
1.2.5. Dobra publiczne 41
1.2.6. Głosowanie większościowe i paradoks głosowania* 48
1.2.7. Dobra społecznie pożądane 51
1.2.8. Niesprawność rynku na szczeblu makroekonomicznym 52
1.2.9. Polityka gospodarcza jako reakcja władzy publicznej na szoki zewnętrzne 54
1.2.10. Redystrybucja i drugie podstawowe twierdzenie ekonomii dobrobytu . . 57
1.3. Normatywna teoria polityki gospodarczej 60
1.3.1. Funkcje polityki gospodarczej 61
1.3.2. Cele 62
1.3.3. Instrumenty 68
1.3.4. Związki między celami a instrumentami. Modele gospodarcze 69
1.3.5. Reguły polityki 74
1.3.6. Krytyka Lucasa 76
1.3.7. Instytucje 82
1.3.8. Polityka gospodarcza jako sztuka wyboru w przestrzeni i w czasie (sekwencja wyborów) 84
1.3.9. Planowanie 87
1.4. Wzrost roli państwa w gospodarce w XX wieku 88
1.5. Pozytywna teoria polityki gospodarczej: niesprawności państwa 91
1.5.1. Czy istnieje interes publiczny? 92
1.5.2. Problem reprezentacji 95
1.5.3. Problem ograniczonej wiedzy i niepełnej informacji 98
1.5.4. Rola niepewności 100
1.5.5. Rola oczekiwań 103
1.5.6. Zawodne państwo? 105
1.6. Rodzaje polityki gospodarczej. Próba klasyfikacji 107
Literatura cytowana 109

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozdział 2. Polityka gospodarcza w modelach makroekonomicznych 113
2.1. Wprowadzenie 113
2.2. Podstawowe zależności makroekonomiczne w modelu IS-LM 115
2.2.1. Rynek produktów 116
2.2.2. Rynek pieniądza 122
2.2.3. Równowaga ogólna i obszary nierównowagi 126
2.3. Model IS-LM z uwzględnieniem rynku pracy 128
2.4. Przywracanie równowagi i rola pieniądza 130
2.4.1. Przywracanie równowagi przy cenach giętkich i cenach sztywnych („lepkich") 130
2.4.2. Klasyczna dychotomia i neutralność pieniądza 133
2.5. Polityka makroekonomiczna w modelu IS-LM 134
2.6. Uwzględnienie wymiany z zagranicą i bilansu płatniczego 139
2.6.1. Model IS-LM dla gospodarki otwartej 139
2.6.2. Model Mundella-Fleminga 143
2.7. Model AS-AD 144
2.7.1. Wpływ zmian cen i krzywa łącznego (zagregowanego) popytu *AD* 144
2.7.2. Krzywa łącznej (zagregowanej) podaży *AS* 147
2.8. Polityka pieniężna i polityka fiskalna w modelu Mundella-Fleminga w warunkach różnych systemów kursu walutowego 149
2.8.1. W systemie stałego kursu walutowego 149
2.8.2. W systemie płynnego kursu walutowego 152
2.9. Reguła efektywnego przyporządkowania 154
2.10. Przywracanie równowagi ogólnej przy pomocy polityki fiskalnej i polityki pieniężnej w warunkach stałego kursu walutowego* 157
2.11. Alternatywne modele makroekonomiczne 164
2.11.1. Wprowadzenie 164
2.11.2. Synteza neoklasyczna 164
2.11.3. Monetaryzm i Nowa Synteza Neoklasyczna 166
2.11.4. Modele realnego cyklu koniunkturalnego 169
2.11.5. Nowe modele keynesowskie 170
Literatura cytowana 172

Rozdział 3. Polityka fiskalno-budżetowa 175
3.1. Zakres, cele i funkcje polityki fiskalno-budżetowej 175
3.1.1. Wprowadzenie 175
3.1.2. Udział sektora publicznego w gospodarce 177
3.1.3. Cele i funkcje polityki fiskalnej 186
3.2. Budżet 189
3.2.1. Podstawowa klasyfikacja pozycji budżetowych 189
3.2.2. Równowaga i nierównowaga budżetowa 193
3.2.3. Federalizm fiskalny 196
3.2.4. Unia fiskalna 199
3.2.5. Finanse samorządów terytorialnych 200

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.3. Skłonność do deficytu (*deficit bias*) 206
3.4. Finansowanie deficytu 211
3.4.1. Finansowanie deficytu przy pomocy zwiększania bazy monetarnej (monetyzacja deficytu) 212
3.4.2. Finansowanie deficytu z długu 214
3.4.3. Zaciąganie pożyczek za granicą 217
3.4.4. Przychody ze sprzedaży aktywów 218
3.5. Rachunkowe metody obliczania salda budżetu 219
3.5.1. Metoda kasowa i metoda memoriałowa 219
3.5.2. Zakres operacji finansowych objętych rachunkiem 221
3.5.3. Salda budżetowe poszczególnych części sektora finansów publicznych . . 222
3.5.4. Saldo pierwotne, saldo kapitałowe, saldo zagraniczne 222
3.5.5. Saldo strukturalne budżetu 224
3.5.6. Deficyt faktyczny i deficyt strukturalny w Polsce w latach 1995–2014 . . . 228
3.5.7. Potrzeby pożyczkowe sektora publicznego 232
3.6. Ekonomiczne konsekwencje deficytu budżetowego 233
3.7. Dług publiczny – podstawowe zależności 239
3.7.1. Wprowadzenie 239
3.7.2. Płynność a wypłacalność 249
3.7.3. Związek między deficytem a długiem 251
3.7.4. Mechanizm rosnącego ryzyka suwerennej niewypłacalności 256
3.7.5. Zakres swobody fiskalnej 258
3.8. Relacje między długiem publicznym a wzrostem gospodarczym 260
3.8.1. Obserwacje empiryczne 260
3.8.2. Wpływ zadłużenia publicznego na tempo wzrostu gospodarczego i poziom stóp procentowych w krajach Unii Europejskiej i strefy euro . . . 262
3.8.3. Maksymalna politycznie możliwa nadwyżka pierwotna i pułapka długu . 264
3.8.4. Czy dług trzeba spłacać? 267
3.9. Wpływ wydatków publicznych i deficytu na popyt krajowy. Spór wokół polityki stabilizacyjnej 269
3.9.1. Mnożnik wydatków w modelu IS-LM 269
3.9.2. Mnożnik wydatków w modelu neoklasycznym* 274
3.9.3. Mnożnik w modelach neokeynesowskich 278
3.9.4. Polityka fiskalna w gospodarce otwartej (na podstawie modelu Mundella-Fleminga) 279
3.9.5. Zwiększać wydatki czy zmniejszać podatki? 281
3.9.6. Równoważność rikardiańska – jak finansować deficyt: z długu czy z podatków? 282
3.9.7. Efekty podażowe i spór wokół polityki stabilizacyjnej 288
3.9.8. Mnożnik wydatków rządowych – wyniki badań empirycznych 291
3.10. Reguły fiskalne 294
3.10.1. Polityka dyskrecjonalna a polityka oparta na regułach 295
3.10.2. Rodzaje reguł fiskalnych 296

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.10.3. Reguły fiskalne w Unii Europejskiej 303
3.10.4. Reguły fiskalne w Polsce. Hamulec długu i progi ostrożnościowe 305
3.10.5. Stabilizująca reguła wydatkowa 308
3.10.6. Reguły fiskalne jednostek samorządu terytorialnego w Polsce 313
3.10.7. Reguły czy instytucje? Rady polityki fiskalnej 315
3.11. Polityka konsolidacji finansów publicznych 318
3.11.1. Struktura dostosowania fiskalnego: podnosić podatki czy obniżać wydatki? 318
3.11.2. Sekwencja działań – co czynić najpierw: oszczędzać czy pobudzać gospodarkę? 321
3.11.3. Tempo konsolidacji fiskalnej – co jest korzystniejsze: więcej konsolidacji teraz i mniej później czy odwrotnie? 327
3.11.4. Tempo konsolidacji fiskalnej – próby weryfikacji statystycznej w przypadku strefy euro 329
Literatura cytowana 336

Rozdział 4. Polityka podatkowa 345
4.1. Podstawowe pojęcia i definicje 345
4.2. Zasady opodatkowania 347
4.2.1. Potrzeba równoważenia skutków podatku 347
4.2.2. Zasada równości poziomej i zasada równości pionowej 348
4.3. Rodzaje podatków 349
4.3.1. Podatki bezpośrednie 350
4.3.2. Podatki pośrednie 350
4.3.3. Efekt zniekształcenia i podatki korekcyjne 354
4.4. Stawki podatkowe 355
4.4.1. Podatki ryczałtowe, proporcjonalne i progresywne 355
4.4.2. Nominalne i efektywne stawki podatkowe 359
4.5. Pożądane cechy systemu podatkowego 360
4.5.1. Efektywność fiskalna 362
4.5.2. Koszt funkcjonowania systemu podatkowego 366
4.5.3. Efektywność ekonomiczna: wpływ podatków na produkcję i konsumpcję 368
4.5.4. Efektywność ekonomiczna: wpływ podatków na zatrudnienie i dochód 380
4.5.5. Efektywność ekonomiczna: podatek optymalny według Ramseya 384
4.6. Podatki a podział dochodu 391
4.6.1. Efektywność a równość 391
4.6.2. Zróżnicowanie dochodów: krzywa Lorenza i współczynnik Giniego 394
4.6.3. Ulgi i zwolnienia podatkowe 398
Literatura cytowana 401

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozdział 5. Polityka pieniężna ... 403
5.1. Podstawowe pojęcia ... 403
5.1.1. Wprowadzenie ... 403
5.1.2. Funkcje, rodzaje i miary pieniądza ... 405
5.1.3. Popyt na pieniądz ... 415
5.1.4. Pieniądz w bilansie banku centralnego, w bilansach banków komercyjnych i w bilansie skonsolidowanym systemu bankowego ... 419
5.1.5. Dwuwalutowość ... 424
5.1.6. Rezerwy bankowe ... 427
5.2. Mnożnik pieniężny ... 430
5.2.1. Definicja mnożnika ... 430
5.2.2. Wielkość mnożnika bazy monetarnej ... 433
5.3. Cele polityki pieniężnej ... 437
5.3.1. Tradycyjne cele polityki pieniężnej ... 437
5.3.2. Stabilność cen jako cel polityki pieniężnej ... 439
5.3.3. Pozostałe cele współczesnej polityki pieniężnej: stabilizacja makroekonomiczna, stabilność kursu walutowego i stabilność finansowa ... 441
5.3.4. Cele polityki pieniężnej w wybranych krajach ... 444
5.3.5. Mierniki inflacji ... 448
5.3.6. Dlaczego (wysoka) inflacja jest szkodliwa? ... 450
5.3.7. Epizody hiperinflacji na świecie ... 455
5.3.8. Dlaczego deflacja jest szkodliwa? ... 459
5.3.9. Cel nadrzędny (ostateczny) i cele pośrednie ... 461
5.4. Instrumenty polityki pieniężnej ... 466
5.4.1. Mechanizm zmian podaży pieniądza przez bank centralny ... 466
5.4.2. Operacje otwartego rynku ... 469
5.4.3. Stopa procentowa: podstawowe pojęcia ... 475
5.4.4. Stopa procentowa jako instrument polityki pieniężnej ... 479
5.4.5. Pozostałe instrumenty ... 484
5.4.6. Wskaźnik restrykcyjności polityki pieniężnej ... 489
5.5. Transmisja impulsów polityki pieniężnej do gospodarki ... 492
5.5.1. Mechanizm transmisji monetarnej ... 492
5.5.2. Rynek pieniężny ... 493
5.5.3. Instrumenty dłużne ... 497
5.5.4. Krzywa rentowności ... 500
5.5.5. Kanały transmisji ... 503
5.5.6. Oczekiwania inflacyjne ... 508
5.6. Reguły polityki pieniężnej ... 513
5.6.1. Dyskrecjonalność polityki pieniężnej i zjawisko niespójności czasowej ... 513
5.6.2. Rozwiązanie problemu niespójności czasowej w obrębie jednego kraju. Niezależny bank centralny ... 524
5.6.3. Podstawowe reguły polityki pieniężnej ... 530
5.6.4. Reguły polityki pieniężnej stosowane na świecie ... 537

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.7. Strategia bezpośredniego celu inflacyjnego 538
5.7.1. Istota i elementy składowe strategii BCI 538
5.7.2. Wybór miernika inflacji 541
5.7.3. Wybór celu inflacyjnego 543
5.7.4. System prognoz inflacji 545
5.7.5. Funkcja reakcji banku centralnego. Reguła Taylora 546
5.7.6. Inflacja jako skutek szoków zewnętrznych i efekty „drugiej rundy" 551
5.7.7. Rola niepewności 553
5.7.8. Strategia komunikowania się z rynkami 555
5.8. Polityka pieniężna w alternatywnych modelach makroekonomicznych 557
5.8.1. Model syntezy neoklasycznej 557
5.8.2. Monetaryzm i teoria realnego cyklu koniunkturalnego 557
5.8.3. Modele nowej syntezy neoklasycznej 558
5.9. Polityka pieniężna w nowych modelach keynesowskich 560
5.9.1. Wprowadzenie 560
5.9.2. Model Clarida, Galì i Gertner: polityka dyskrecjonalna 561
5.9.3. Model CGG: polityka pieniężna oparta na wiążących regułach 569
5.10. Konflikt celów w polityce pieniężnej i „trójkąt niemożności" 573
5.10.1. Parytet stóp procentowych 573
5.10.2. Konflikt celów i „trójkąt niemożności" (*impossible trinity*) 578
5.11. Zależność między polityką pieniężną a polityką fiskalną 582
5.11.1. Czy polityka pieniężna i polityka fiskalna są od siebie zależne? 582
5.11.2. „Nieprzyjemna arytmetyka monetarna" 585
5.12. Polityka pieniężna w Polsce 590
5.12.1. Lata 1990–1998: reguła celu kursowego i strategia „eklektyczna" 590
5.12.2. Lata 1999–2015: reguła BCI i stabilizacja inflacji 594
5.13. Polityka pieniężna a kryzys finansowy 599
5.13.1. Wpływ globalizacji 599
5.13.2. Kryzys finansowy w latach 2008–2009 601
5.13.3. Polityka pieniężna a bańki spekulacyjne na rynku aktywów 603
5.13.4. Czy realizacja strategii celu inflacyjnego może zapobiegać kryzysom finansowym? 607
5.14. Nowe tendencje w polityce pieniężnej 610
5.14.1. Problem „zerowej dolnej granicy" i potrzeba zastosowania niekonwencjonalnych instrumentów polityki pieniężnej 610
5.14.2. Polityka luzowania ilościowego 617
5.14.3. Pieniądze z „helikoptera" (*helicopter money*) 626
5.14.4. Polityka „wytycznych na przyszłość" (*forward guidance*) 631
5.14.5. Zmiany w strategii BCI 634
5.14.6. Funkcja „pożyczkodawcy ostatniej szansy" dla rządów i wielorakie równowagi na rynku długu 636
5.15. Przyszłość pieniądza 643

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.15.1. Jaka będzie przyszłość pieniądza gotówkowego? 643
5.15.2. Pieniądz wirtualny 645
Literatura cytowana 646

Rozdział 6. Polityka kursu walutowego 657
6.1. Kurs walutowy: podstawowe pojęcia 657
6.1.1. Kurs walutowy jako cena waluty obcej 657
6.1.2. Kurs nominalny, realny i efektywny 659
6.1.3. Kurs parytetu siły nabywczej 663
6.1.4. Krótka historia kursów walutowych 667
6.1.5. Podstawowe decyzje z zakresu polityki kursowej 673
6.2. Wybór systemu kursowego 674
6.2.1. Wybór między kursem stałym a kursem płynnym 674
6.2.2. Rozwiązania pośrednie 680
6.3. Wpływ systemu kursowego na skuteczność polityki pieniężnej i fiskalnej 684
6.3.1. System kursowy a skuteczność polityki pieniężnej 684
6.3.2. System kursowy a skuteczność polityki fiskalnej 689
6.4. Wymienialność waluty krajowej 695
6.4.1. Przyczyny występowania ograniczeń w zakresie wymienialności walut krajowych na waluty obce 695
6.4.2. Zakres wymienialności 698
6.4.3. Czy kapitał powinien swobodnie przepływać między krajami? 700
6.5. Rynek walutowy 705
6.5.1. Podaż i popyt na rynku walutowym 705
6.5.2. Bilans płatniczy 708
6.5.3. Równowaga na rynku walutowym 714
6.5.4. Krótkookresowa zmienność kursu walutowego – zjawisko „przestrzelenia" (*exchange rate overshooting*)* 718
6.6. Czynniki określające kurs walutowy 722
6.6.1. Kurs walutowy i strumienie handlu: podejście tradycyjne (strumieniowe) i warunek Marshalla-Lernera-Robinson 722
6.6.2. Kurs walutowy i rynek aktywów: model monetarny i model równowagi portfela 730
6.7. Kurs walutowy w krótkookresowym modelu makroekonomicznym* 736
6.7.1. Kurs walutowy jako czynnik określający popyt w gospodarce otwartej . . 736
6.7.2. Równowaga na rynku produktów i konstrukcja krzywej *DD* 737
6.7.3. Równowaga na rynku pieniądza krajowego i na rynku pieniądza zagranicznego. Konstrukcja krzywej *AA* 739
6.7.4. Równowaga ogólna i przywracanie równowagi 744
6.8. Kurs walutowy w średnim i długim okresie 747
6.8.1. Kurs równowagi długookresowej (*equilibrium exchange rate*) 747
6.8.2. Kurs parytetu siły nabywczej: rozwinięcie 748
6.8.3. Efekt Balassy-Samuelsona 753

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

6.8.4. Efekt Balassy-Samuelsona a zmiany kursu walutowego 758
6.8.5. Wpływ zmian podaży i popytu na kurs walutowy w długim okresie 762
6.8.6. Wpływ trwałych zmian w zakresie polityki makroekonomicznej na kurs walutowy w długim okresie 763
6.9. Dewaluacja jako instrument polityki gospodarczej 766
6.9.1. Argumenty na rzecz dewaluacji 766
6.9.2. Argumenty przeciwko dewaluacji 769
6.10. Interwencje walutowe 773
6.10.1. Cele i zasady interwencji walutowych 773
6.10.2. Rodzaje interwencji walutowych 774
6.10.3. Skuteczność interwencji walutowych 778
6.10.4. Interwencja sterylizowana 782
6.11. Polityka kursowa w Polsce 784
6.11.1. Polityka kursowa w Polsce w latach 1990–1999 784
6.11.2. Polityka kursowa w Polsce w latach 2000–2015 787
Literatura cytowana 790

Rozdział 7. Polityka makroostrożnościowa 793
7.1. Wprowadzenie 793
7.2. Potrzeba stosowania nadzoru nad sektorem bankowym 795
7.2.1. Specyfika banków 795
7.2.2. Nadzór mikroostrożnościowy i jego ograniczenia 798
7.3. Ryzyko systemowe w sektorze bankowym 801
7.3.1. Słabości wspólne oraz instytucje „ważne systemowo” 801
7.3.2. Rodzaje ryzyka systemowego 805
7.3.3. Pomiar ryzyka systemowego 808
7.3.4. Modele ryzyka systemowego* 810
7.4. Pojęcie, cele i zakres polityki makroostrożnościowej 816
7.4.1. Regulacja „mikro” i regulacja „makro” 816
7.4.2. Polityka makroostrożnościowa a polityka makroekonomiczna 818
7.5. Instrumenty polityki makroostrożnościowej 820
7.5.1. Instrumenty regulujące rozmiary i tempo wzrostu kredytu 820
7.5.2. Instrumenty regulujące płynność aktywów 823
7.5.3. Instrumenty regulujące wielkość kapitału banku 825
7.5.4. Instrumenty regulujące transgraniczne przepływy kapitału 828
7.5.5. Instrumenty oddziałujące na stronę aktywną i na stronę pasywną bilansu banku 831
7.6. Zastosowanie polityki makroostrożnościowej 833
7.6.1. Jeden czy kilka instrumentów jednocześnie? 833
7.6.2. Reguły czy dyskrecjonalność? 833
7.6.3. Instrumenty stałe czy zmieniane w czasie? 834
7.6.4. Polityka makroostrożnościowa a polityka makroekonomiczna (fiskalna, pieniężna, kursowa) 835

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

7.6.5. Ramy instytucjonalne polityki makroostrożnościowej 838
7.6.6. Podstawy polityki makroostrożnościowej w Unii Europejskiej 840
7.6.7. Ramy instytucjonalne polityki makroostrożnosciowej w Polsce 841
7.7. Transgraniczna transmisja szoków i potrzeba międzynarodowej koordynacji polityki makroostrożnościowej 842
7.8. Zakończenie 844
Literatura cytowana 845

Rozdział 8. Koordynacja polityki gospodarczej w skali międzynarodowej 849
8.1. Globalizacja 849
8.2. Potrzeba międzynarodowej koordynacji polityki 852
8.2.1. Międzynarodowe dobra publiczne 852
8.2.2. Efekty „rozlewania" (*spillovers*) 854
8.3. Granice koordynacji 861
8.3.1. Zawodność koordynacji 861
8.3.2. Koordynacja polityki w krajach rozwiniętych 863
8.4. Teoria unii gospodarczej 865
8.5. Teoria unii walutowej 870
8.5.1. Podstawy teorii optymalnych obszarów walutowych 870
8.5.2. Dalszy rozwój teorii optymalnych obszarów walutowych 875
8.5.3. Wstrząsy asymetryczne 879
8.6. Potrzeba koordynacji polityki gospodarczej w unii walutowej 885
8.6.1. Transmisja impulsów polityki fiskalnej w unii walutowej: kanał handlu i kanał stóp procentowych (*fiscal policy spillovers*) 885
8.6.2. Potrzeba stosowania dyscypliny fiskalnej w obrębie unii walutowej: uzasadnienie teoretyczne 891
8.7. Znaczenie konwergencji nominalnej. Kryteria konwergencji określone w Traktacie z Maastricht 894
8.7.1. Zapisy traktatowe 894
8.7.2. Ocena kryteriów konwergencji nominalnej 899
8.8. Koordynacja polityki gospodarczej w Unii Europejskiej 902
8.8.1. Ramy instytucjonalne koordynacji polityki gospodarczej w Unii Europejskiej 902
8.8.2. Pakt Stabilności i Wzrostu 905
8.8.3. Procedura nadmiernego deficytu 906
8.8.4. Ocena systemu koordynacji polityki fiskalnej w Unii Europejskiej w latach 1999–2009 908
8.9. Kryzys w strefie euro 911
8.9.1. Kryzys w strefie euro w latach 2009–2012 911
8.9.2. Czy brak konwergencji mógł być przyczyną kryzysu w strefie euro? 918
8.9.3. Co ujawnił kryzys? Niedostosowanie instytucjonalne w strefie euro 925
8.9.4. Brak mechanizmu podziału ryzyka 931
8.9.5. „Trójkąt" współzależności 935

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

8.10. Nowy system zarządzania ekonomicznego w Unii Europejskiej 937
8.10.1. Semestr Europejski i Procedura Nierównowagi Makroekonomicznej . . 937
8.10.2. Zreformowany Pakt Stabilności i Wzrostu 941
8.10.3. Traktat o Stabilizacji, Koordynacji i Zarządzaniu 945
8.10.4. Unia Bankowa 947
8.10.5. Europejski Mechanizm Stabilizacyjny 949
8.11. Przyszłość strefy euro 953
8.11.1. Czy strefa euro jest optymalnym obszarem walutowym? 953
8.11.2. Kierunki dalszych reform 954
8.11.3. Czy Polska powinna dołączyć do strefy euro? 959
Literatura cytowana 960

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Przedmowa

Niniejsza monografia została opracowana na podstawie cyklu wykładów z polityki gospodarczej prowadzonych w latach 2010–2016 dla studentów Szkoły Głównej Handlowej w Warszawie. Decyzję o napisaniu książki podjąłem, gdy uświadomiłem sobie, że polska literatura ekonomiczna liczy bardzo niewiele całościowych prac obejmujących zagadnienia polityki gospodarczej, w tym uwzględniających zwłaszcza w większym stopniu doświadczenia wyniesione z kryzysu finansowego i gospodarczego lat 2008–2012.

Politykę gospodarczą traktuję jako ważną dziedzinę polityki publicznej (polityki państwa), polegającą na takim kształtowaniu działalności gospodarczej przez władzę publiczną, które zmierza do zapewnienia możliwie szybkiego i zrównoważonego rozwoju gospodarczego kraju. Ta zwięzła definicja pozwala określić cel, przedmiot oraz podmiot polityki gospodarczej. Celem jest szybki i zrównoważony rozwój gospodarczy, a więc rozwój zapewniający wysokie tempo wzrostu gospodarczego oraz równowagę gospodarczą, nazywaną też stabilnością makroekonomiczną przy uwzględnieniu także innych celów rozwojowych. Przedmiot polityki gospodarczej stanowi gospodarka krajowa, obejmująca przedsiębiorstwa prywatne i publiczne, gospodarstwa domowe oraz instytucje rynkowe i system prawa gospodarczego, a także relacje gospodarki krajowej z otoczeniem międzynarodowym. Podmiotem polityki gospodarczej jest zaś władza publiczna, czyli z reguły rząd wyposażony w kompetencje do podejmowania decyzji oraz parlament stanowiący prawo.

Główny obszar zainteresowań w niniejszej publikacji stanowi polityka makroekonomiczna. Tradycyjnie zalicza się do niej politykę fiskalno-budżetową, politykę pieniężną oraz politykę kursową, przy czym ta ostatnia jest w istocie częścią szeroko rozumianej polityki pieniężnej. W książce dodatkowo omawiana jest także polityka nadzoru makroostrożnościowego nad sektorem bankowym, po ostatnim kryzysie finansowym i gospodarczym będąca powszechnie akceptowaną częścią polityki gospodarczej na szczeblu makroekonomicznym. Kwestie mikroekonomiczne, takie jak funkcjonowanie poszczególnych rynków (związanych z rynkiem

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pracy i polityką zatrudnienia czy też z rynkiem kapitałowym), a także stosowanie instrumentów mikroekonomicznych i sektorowych (takich jak polityka cen, polityka przemysłowa bądź polityka handlowa) nie są tu zasadniczo omawiane. Wyjątkiem jest polityka podatkowa, której poświęcono osobny rozdział ze względu na jej ścisły związek z polityką fiskalno-budżetową. Ponieważ rozważania koncentrują się przede wszystkim na stabilizacji makroekonomicznej, na uboczu pozostawiono również – mimo ich niewątpliwej wagi – zagadnienia polityki wzrostu gospodarczego, polityki zrównoważonego rozwoju oraz zagadnienia polityki społecznej (w szczególności ta ostatnia z nich powinna być traktowana jako odrębna dziedzina polityki publicznej). Dlatego też książka nosi podtytuł „wybrane zagadnienia".

Głównym celem publikacji jest dostarczenie czytelnikom podstawowego kompendium współczesnej wiedzy na temat teorii i praktyki polityki gospodarczej. Mam nadzieję, że książka okaże się przydatna dla wszystkich tych, którzy zainteresowani są zagadnieniami polityki gospodarczej. Powstała bowiem z myślą zarówno o studentach, jak i o politykach oraz praktykach, zaangażowanych w projektowanie i prowadzenie polityki gospodarczej w kraju, a także o wszystkich osobach, które chciałyby pogłębić swoją wiedzę dotyczącą funkcjonowania gospodarki rynkowej oraz sposobów i możliwości osiągania ważnych celów gospodarczych. Dodatkowym zamiarem jest również przybliżenie czytelnikowi najświeższych osiągnięć światowej literatury ekonomicznej z zakresu tej dziedziny. Niedawny kryzys finansowy i gospodarczy rzucił nowe światło na wiele zasadniczych założeń tradycyjnej polityki gospodarczej, zmuszając ekonomistów do ich ponownego przemyślenia i weryfikacji. Dlatego też staram się przedstawić w książce – oprócz „uznanych" prawd – także najważniejsze wnioski płynące z ostatniego kryzysu oraz odpowiedzi, jakich wobec tych wyzwań pozwala udzielić współczesna nauka ekonomii.

W dwóch pierwszych rozdziałach opisane zostały ramy i podstawy polityki gospodarczej. Rozdział pierwszy prezentuje podstawowe założenia polityki gospodarczej przy jednoczesnym wskazaniu na to, że istnienie nieusuwalnych niesprawności mechanizmu rynkowego stanowi obiektywne uzasadnienie dla interwencji władzy publicznej w obrębie gospodarki rynkowej, w tym zwłaszcza interwencji w ramach polityki gospodarczej. Do zasadniczych niedoskonałości mechanizmu rynkowego zaliczam występowanie efektów zewnętrznych (pozarynkowych), potrzebę dostarczania dóbr publicznych, redystrybucję dochodu mogącą odbiegać od społecznych preferencji oraz istnienie cyklu koniunkturalnego powodującego cykliczne spadki produkcji i zatrudnienia w gospodarce rynkowej. W rozdziale drugim ukazane zostały podstawowe zależności makroekonomiczne w gospodarce,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

a za punkt wyjścia przyjęto klasyczne modele IS-LM i AS-AD, które następnie rozwinięto w ramach nowszych modeli makroekonomicznych, reprezentujących założenia głównych nurtów współczesnej myśli ekonomicznej. Do najważniejszych osiągnięć, jakie wypracowano w obrębie tych modeli, mających największy wpływ na kształtowanie współczesnej polityki gospodarczej, zaliczam uwzględnienie występowania racjonalnych oczekiwań, niedoskonałej konkurencji oraz wielookresowej optymalizacji zachowań konsumentów.

Kolejne pięć rozdziałów koncentruje się już ściśle na polityce gospodarczej. Omawiając politykę fiskalno-budżetową w rozdziale trzecim, poza opisem jej podstawowych zadań i narzędzi, najwięcej miejsca poświęcono związkom między deficytem, długiem i wzrostem gospodarczym, a także wpływowi wydatków publicznych na łączny popyt oraz polityce konsolidacji finansów publicznych. Niezbędnym uzupełnieniem tych zagadnień okazała się polityka podatkowa, która została opisana w rozdziale czwartym. Z kolei omawianie polityki pieniężnej w rozdziale piątym rozpoczynam od przypomnienia celów, instrumentów i kanałów transmisji pieniężnej. Następnie wyjaśnione zostały reguły polityki pieniężnej, w tym przede wszystkim reguła bezpośredniego celu inflacyjnego; przedstawiono także konflikt celów występujący w polityce pieniężnej w warunkach gospodarki otwartej. Ostatnie fragmenty rozdziału poświęcone są zaś nowym tendencjom, jakie pojawiły się w polityce pieniężnej w następstwie kryzysu finansowego. W rozdziale szóstym zawarte zostały główne zagadnienia polityki kursu walutowego. Punktem wyjścia jest tu przypomnienie podstawowych pojęć i najważniejszych decyzji, jakie podejmuje się w ramach polityki kursowej – dotyczących wyboru systemu kursowego, zakresu wymienialności oraz pożądanego poziomu kursu. Następnie przedstawiono funkcjonowanie rynku walutowego oraz czynniki kształtujące kurs walutowy w krótkim i długim okresie, a także politykę interwencji. Przedmiotem rozdziału siódmego jest natomiast polityka makroostrożnościowa. Na tle mechanizmu powstawania ryzyka systemowego omówiono tu narzędzia polityki i regulacji nadzorczych na szczeblu makroekonomicznym oraz wskazano główne zależności zachodzące między nimi. Ostatni rozdział, poświęcony międzynarodowej koordynacji polityki gospodarczej, rozpoczyna się od przedstawienia argumentów przemawiających zarówno na rzecz, jak i przeciw koordynacji, a także zawiera omówienie elementów teorii unii gospodarczej oraz teorii unii walutowej. Następnie przechodzę do wyjaśnienia zależności i powiązań między polityką poszczególnych państw, przejawiającą się w postaci efektów zewnętrznych (*spillovers*), by na tym tle omówić koordynację polityki gospodarczej w ramach integracji europejskiej, a zwłaszcza w obrębie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

unii walutowej. W dalszych częściach rozdziału przedstawione zostały przyczyny i przebieg kryzysu zadłużenia w strefie euro, wskazano kierunki wprowadzanych reform na tle głównych dylematów koordynacji polityki narodowej oraz zarysowano alternatywne scenariusze dalszego funkcjonowania unii walutowej w Europie.

W toku wykładu, oprócz omawiania zależności ekonomicznych, nawiązuję do doświadczeń gospodarczych różnych krajów, w tym przede wszystkim Polski. Przykłady praktyczne, a także opis i wyprowadzenie niektórych twierdzeń podane są najczęściej w ramkach, które służą jako materiał uzupełniający. Fragmenty tekstu opatrzone gwiazdką (*) dotyczą zagadnień bardziej zaawansowanych i mogą być pominięte podczas lektury bez uszczerbku dla zrozumienia głównej myśli wywodu.

Pragnę podziękować prof. dr hab. Andrzejowi Wojtynie, który podjął się trudu zrecenzowania pracy, a jego krytyczne uwagi pozwoliły mi usunąć wiele błędów znajdujących się w początkowej wersji monografii. Dziękuję również kolegom z Instytutu Handlu Zagranicznego i Studiów Europejskich SGH, z którymi wielokrotnie dyskutowałem na temat szeregu zagadnień podjętych w książce. Podziękowania przekazuję także Oficynie Wydawniczej SGH – za cenną pomoc i współpracę przy publikacji książki.

Dariusz K. Rosati

Warszawa, czerwiec 2016

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozdział 1

PODSTAWY POLITYKI GOSPODARCZEJ

1.1. Wprowadzenie

Samo pojęcie „polityka" wywodzi się od greckiego słowa *polis*, oznaczającego miasto-państwo. Arystoteles w *Etyce nikomachejskiej* mianem polityki (gr. *πολιτική*) określił „sztukę rządzenia państwem", mającą na celu osiąganie dobra wspólnego. Współcześnie politykę traktuje się nie tylko jako „sztukę" sprawowania władzy, lecz także jako „sposób" rządzenia oparty na racjonalnych przesłankach i konkretnej wiedzy.

W niniejszej książce polityka gospodarcza rozumiana jest jako świadome oddziaływanie władzy publicznej na gospodarkę z zamiarem osiągnięcia określonych celów, przy wykorzystaniu określonych narzędzi ekonomicznych i prawnych. Ponieważ podstawowym mechanizmem regulującym procesy gospodarcze jest mechanizm rynkowy, polityka gospodarcza musi wiązać się z ingerencją w jego funkcjonowanie. Wyjaśniając cele i zadania polityki gospodarczej, należy więc w pierwszej kolejności odpowiedzieć na pytanie o to, dlaczego oraz w jakich okolicznościach ingerencja władzy publicznej w mechanizm rynkowy jest uzasadniona i pożądana.

Mechanizm rynkowy, dostosowując poziom i strukturę podaży do popytu, kształtuje alokację zasobów produkcyjnych z podziałem na różne możliwości zastosowania. Funkcjonuje on jednak w pełni sprawnie tylko w warunkach doskonałej konkurencji. Dostosowanie podaży do popytu staje się następstwem decyzji podejmowanych przez producentów, dotyczących rozdysponowania zasobów między różne rodzaje produkcji na podstawie sygnałów cenowych i kosztowych płynących z konkurencyjnego rynku. Skłaniają one indywidualnych producentów do alokowania posiadanych zasobów w te rodzaje produkcji, które zapewniają

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

największy zysk. Z kolei konsumenci, kierując się sygnałami cenowymi pochodzącymi z rynku, wydają posiadane środki na zakup dóbr i usług, które najpełniej zaspokajają ich potrzeby. W ten sposób mechanizm rynkowy pozwala gospodarce osiągnąć optymalną równowagę – tj. w taki sposób wykorzystywać ograniczone zasoby, aby w jak najlepszym stopniu zaspokoić potrzeby społeczne.

Lecz w pełni doskonała konkurencja istnieje jedynie w teorii. W rzeczywistym świecie konkurencja rynkowa jest z reguły niedoskonała i podlega wielu ograniczeniom: zasoby produkcyjne nie są w pełni przenośne i podzielne, ceny nie stają się w pełni giętkie, a informacja o popycie i podaży pozostaje niepełna. Działalność gospodarcza prowadzi ponadto do powstawania efektów pozarynkowych, zewnętrznych (*externalities*), które nie są brane pod uwagę przez uczestników rynku w procesie podejmowania decyzji o zakupie bądź sprzedaży, choć mają istotne znaczenie dla całego społeczeństwa. Mechanizm rynkowy nie uwzględnia wreszcie celów nieekonomicznych, takich jak sprawiedliwy podział wytworzonego produktu, społeczna solidarność, bezpieczeństwo pracowników i konsumentów, edukacja, kultura czy ochrona środowiska naturalnego.

Te niedostatki rzeczywistej konkurencji i mechanizmu rynkowego stały się przesłanką do poszukiwania alternatywnych – w stosunku do rynków – mechanizmów alokacji zasobów i zaspokajania potrzeb. Najbardziej znanym z tych rozwiązań był system gospodarki centralnie planowanej, wykorzystujący instrumenty nakazowo-rozdzielcze. Zastosowanie tego systemu w praktyce w wielu krajach świata po II wojnie światowej zarówno w jego ortodoksyjnej postaci, jak i w postaci systemu mieszanego, uzupełnionego o elementy mechanizmu rynkowego, nie przyniosło jednak oczekiwanych rezultatów. Gospodarki centralnie planowane okazały się mało efektywne, niezdolne do innowacji i elastycznych dostosowań, wysoce zbiurokratyzowane. Na przełomie lat 80. i 90. większość krajów, które próbowały budować gospodarkę centralnie planowaną, odeszła od tego rozwiązania i powróciła do systemu gospodarki rynkowej.

Politycy i ekonomiści zdają sobie obecnie sprawę z immanentnych słabości oraz zawodności mechanizmu rynkowego. Jeśli ma on dobrze służyć zaspokajaniu potrzeb społeczeństwa, wymaga korekt i uzupełnień. Zadanie to spada na barki władzy publicznej, która – występując w roli reprezentanta interesów całego społeczeństwa, a więc interesu publicznego – stara się w sposób świadomy oddziaływać na procesy gospodarcze. Polityka gospodarcza jest zatem formą interwencji państwa w gospodarkę, której celem jest takie korygowanie mechanizmu rynkowego, które służyć ma poprawie efektywności gospodarowania i osiąganiu rezultatów

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

społecznie pożądanych. Tymczasem, choć sama potrzeba prowadzenia polityki gospodarczej nie jest na ogół kwestionowana, ekonomiści i politycy spierają się o to, jaki powinien być zakres tej interwencji oraz jak powinny przedstawiać się jej cele i instrumenty, a także warunki realizacji.

1.2. Niesprawności rynku

Potrzebę stosowania polityki gospodarczej uzasadnia się głównie niesprawnością mechanizmu rynkowego, czyli niedoskonałością mechanizmu konkurencji. Zgodnie z tym poglądem, w dziedzinach, w których rynek nie funkcjonuje bądź działa zawodnie, władza publiczna powinna podejmować działania uzupełniające albo korygujące mechanizm rynkowy. Uzasadnienie potrzeby przeprowadzenia interwencji publicznej w ramach mechanizmu rynkowego obejmuje zasadniczo dwa etapy. Po pierwsze, zanim podejmie się interwencję, należy wcześniej upewnić się, że mechanizm konkurencji rynkowej w danej dziedzinie pozostaje rzeczywiście niesprawny i nie uwzględnia potrzeb społecznych oraz że w efekcie w danej sytuacji rynek nie zapewnia dostępu do optymalnego rozwiązania. Po drugie, można uznać, że interwencja jest uzasadniona, gdy w jej rezultacie osiąga się efekt lepszy ze społecznego punktu widzenia, niż gdyby pozostawić sprawy do rozstrzygnięcia rynkowi. Poniżej zostanie przedstawiona pierwsza z tych kwestii, podczas gdy do drugiej przyjdzie powrócić w dalszych rozdziałach.

1.2.1. Zasada Pareto i pierwsze podstawowe twierdzenie ekonomii dobrobytu

Włoski ekonomista Vilfredo Pareto sformułował w 1897 roku zasadę porównywania różnych stanów gospodarki pod względem osiągniętego poziomu dobrobytu (*welfare*). Miarą dobrobytu jest funkcja użyteczności społecznej (funkcja społecznego dobrobytu – *social welfare function*), na którą składają się indywidualne użyteczności poszczególnych jednostek, tworzących dane społeczeństwo. Funkcja użyteczności społecznej wynika z określonego uporządkowania, tj. uszeregowania stanów społecznych (stanów rzeczywistości), które mogą zachodzić w społeczeństwie. To uszeregowanie (*social ordering*) polega na wskazaniu, które stany społeczne są jednakowo preferowane (czyli obojętne względem siebie), lepsze od innych lub od nich gorsze (zob. Acocella 2002, s. 24–25). Zasada Pareto (zwana

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

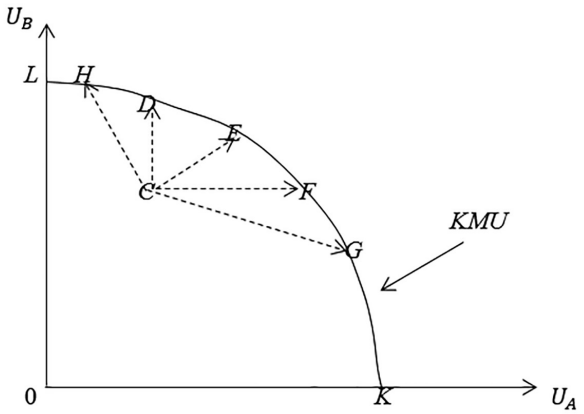
także kryterium Pareto) pozwala dokonywać uszeregowania tych różnych stanów z określonego punktu widzenia. Głosi, że społeczeństwo poprawia swój dobrobyt (mierzony wartością funkcji użyteczności społecznej), przechodząc od stanu A do stanu B, jeśli przynajmniej jedna jednostka znajdzie się w lepszej sytuacji w stanie B niż w stanie A, a sytuacja pozostałych jednostek się nie pogorszy. Taką sytuację nazywa się też korzyścią w sensie Pareto (zob. Stiglitz 2004, s. 69–88). Reguła ta jest w istocie miarą efektywności produkcji, ponieważ uzależnia poziom dobrobytu społecznego od jak najlepszego wykorzystania zasobów produkcyjnych.

Zasada Pareto pozwala na sformułowanie warunku optymalności stanu gospodarki – jest to wówczas optymalność w sensie Pareto (*Pareto optimality*). Jeśli gospodarka osiągnie stan, w którym poprawa sytuacji pewnych jednostek nie jest możliwa bez jednoczesnego pogorszenia się sytuacji innych jednostek, stan ten jest optymalny w rozumieniu Pareto (optimum Pareto). Mówi się też, że osiągnięta równowaga jest optymalna w sensie Pareto.

Rysunek 1.1 stanowi ilustrację zasady Pareto przy założeniu, że społeczeństwo składa się tylko z dwóch osób, A i B, konsumujących dobro x. Na osiach odłożono użyteczności tych dwóch osób, U_A i U_B, wynikające z konsumpcji dobra x. Przy danym poziomie podaży dobra x maksymalna suma możliwych użyteczności obu osób określona jest krzywą maksymalnych użyteczności *KMU*. Krzywa ta jest wklęsła do początku układu współrzędnych, co oznacza, że – zgodnie ze standardowym założeniem ekonomii neoklasycznej – krańcowa użyteczność konsumpcji dobra x przez poszczególne osoby maleje wraz ze wzrostem poziomu konsumpcji tego dobra. Załóżmy, że osoba A konsumuje całość podaży dobra x, a osoba B nie konsumuje go wcale (punkt K na rysunku). Wówczas przesunięcie jednostki x od A do B pociąga za sobą jedynie niewielki spadek użyteczności osoby A (U_A), lecz zarazem powoduje bardzo wysoki przyrost użyteczności osoby B (U_B). W takiej sytuacji nachylenie krzywej *KMU* będzie bardzo strome (i takie rzeczywiście jest w sąsiedztwie osi poziomej U_A). Również odwrotnie: gdy osoba B konsumuje całość podaży x (punkt L na rysunku), to przesunięcie jednostki x od B do A pociąga za sobą jedynie niewielki spadek użyteczności osoby B (U_B), ale zarazem powoduje bardzo wysoki przyrost użyteczności osoby A (U_A). W takiej sytuacji nachylenie krzywej *KMU* będzie bardzo płaskie (i takie rzeczywiście jest w sąsiedztwie osi pionowej U_B).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 1.1. Zasada Pareto – ilustracja



Źródło: opracowanie własne.

Zasada malejącej użyteczności konsumpcji pozwala porównywać ze sobą poziomy łącznej użyteczności różnych wariantów podziału dobra *x* pomiędzy konsumpcje osoby A i osoby B. Załóżmy, że punktem wyjścia jest taki podział konsumpcji dobra *x* między obie osoby, który opisany jest punktem *C*. Wówczas punkt *E* pozostaje bezwzględnie lepszy, bo użyteczności obu osób są oczywiście większe w punkcie *E* niż w punkcie *C*. Korzystne jest też przejście z *C* do *D*, lub z *C* do *F*, ponieważ poprawa użyteczności jednej osoby nie powoduje obniżenia użyteczności drugiej osoby. Punkty *D*, *E* i *F* są więc bezwzględnie lepsze w sensie Pareto niż punkt *C*. Przejście do każdego z tych punktów zapewnia osiągnięcie korzyści w rozumieniu Pareto. Nie można jednak już tego powiedzieć o przejściu z punktu *C* do punktu *G* lub *H* ze względu na to, że poprawa użyteczności jednej osoby powoduje w tym przypadku pogorszenie użyteczności drugiej osoby. Ponieważ krzywa *KMU* jest na całej długości wklęsła do początku układu, wszystkie położone na niej punkty pozostają optymalne w sensie Pareto, bo nie można poprawić użyteczności jednej osoby bez jednoczesnego zmniejszenia użyteczności drugiej osoby.

Czy istnieje zatem jakiś mechanizm, który zapewniałby, że gospodarka znajdzie się na krzywej *KMU*, czyli osiągnie stan optymalny, w którym poprawa sytuacji pewnych osób nie jest już możliwa bez jednoczesnego pogorszenia sytuacji innych osób? Można wykazać, że gdy spełnione są warunki doskonałej konkurencji (*perfect competition*) i rynki pozostają „kompletne" (*complete markets*), mechanizm rynkowy

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

powoduje, że gospodarka zawsze zmierza do osiągnięcia optimum Pareto. Prawidłowość tę nazywa się pierwszym podstawowym twierdzeniem ekonomii dobrobytu (*first fundamental theorem of welfare economics*). Gdy spełnione są bowiem warunki doskonałej konkurencji, mechanizm rynkowy wymusza alokację zasobów produkcyjnych w takich sektorach produkcji, w których produktywność zasobów jest najwyższa, a produkcja najbardziej opłacalna. Proces ten będzie przebiegał aż do momentu, kiedy nastąpi zrównanie krańcowej produktywności zasobów we wszystkich możliwych zastosowaniach, czyli wyrównanie się tzw. krańcowych stóp technicznej transformacji w przypadku każdej pary produktów i dla wszystkich zasobów. Zapewniona jest wówczas maksymalna *efektywność produkcji*, co oznacza, że nie da się wytworzyć większej ilości jednego dobra bez jednoczesnego zmniejszenia poziomu produkcji innego dobra. Oznacza to, że gospodarka znajduje się na krzywej *KMU* i wykorzystuje zasoby w sposób najbardziej efektywny.

Podobnie funkcjonuje mechanizm rynkowy w warunkach doskonałej konkurencji, wymuszając taką strukturę konsumpcji, która zapewnia wyrównanie się krańcowych stóp substytucji w przypadku każdej pary dóbr i dla wszystkich konsumentów. Zapewniona jest wówczas maksymalna *efektywność wymiany*, co oznacza, że nikomu nie opłaca się dokonywać dalszych transakcji zamiennych, ponieważ nie da się poprawić sytuacji jednej osoby bez jednoczesnego pogorszenia sytuacji drugiej osoby. Wreszcie musi nastąpić jednoczesne zrównanie krańcowych stóp transformacji i krańcowych stóp substytucji w przypadku każdej pary dóbr na poziomie równym relacji ich cen rynkowych. Zapewniona jest wówczas *efektywność struktury produkcji*.

Spełnienie trzech wymienionych warunków – wyrównanie się krańcowych stóp technicznej transformacji zasobów w przypadku każdej pary produktów, wyrównanie się krańcowych stóp substytucji dla każdej pary dóbr, oraz wyrównanie się krańcowych stóp transformacji i krańcowych stóp substytucji wobec każdej pary dóbr na poziomie cen rynkowych – oznacza osiągnięcie pełnej efektywności w sensie Pareto. Każde odchylenie od tak określonego stanu równowagi powoduje pogorszenie funkcji użyteczności społecznej, a to oznacza, że dany stan równowagi jest optymalny w sensie Pareto[1]. Formalny zapis warunków optymalności Pareto dla przypadku wielu dóbr i wielu zasobów jest przedstawiony w ramce 1.1.

[1] Więcej na temat warunków efektywności w rozumieniu Pareto zob. m.in. Stiglitz (2004), s. 76–87; Acocella (2002), s. 37–50 oraz Rosati (1990), zwłaszcza rozdział 2.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 1.1. Optimum Pareto

Optimum Pareto oznacza taki stan równowagi systemu gospodarczego, w którym nie jest możliwe zwiększenie poziomu dobrobytu jakiegokolwiek uczestnika systemu bez równoczesnego obniżenia poziomu dobrobytu innego uczestnika. Można wykazać, że osiągnięty stan równowagi jest optymalny w sensie Pareto, gdy jednocześnie spełnione są cztery warunki[2]:

a) krańcowe produkcyjności poszczególnych zasobów produkcyjnych wyrównują się w przypadku wszystkich rodzajów produkcji – jest to optimum techniczne;
b) krańcowe stopy substytucji (KSS) konsumpcji w przypadku każdej pary towarów wyrównują się dla wszystkich konsumentów – jest to optimum wymiany;
c) krańcowe stopy transformacji (KST) produkcji poszczególnych towarów wyrównują się w przypadku wszystkich rodzajów produkcji – jest to optimum produkcji;
d) krańcowe stopy substytucji konsumpcji i krańcowe stopy transformacji produkcji wyrównują się w przypadku tych samych towarów – jest to optimum ogólne.

Pierwszy warunek, tj. wyrównanie się krańcowych produkcyjności poszczególnych zasobów, oznacza, że gospodarka znajduje się na tak zwanej krzywej kontraktu (*contract curve*), której odpowiada krzywa maksymalnych możliwości produkcyjnych, czyli krzywa transformacji. Jest to optimum techniczne, gwarantujące pełne wykorzystanie zasobów. Zakładając, że w gospodarce wytwarza się n produktów przy wykorzystaniu m zasobów oraz oznaczając produkty, ich ceny oraz zasoby odpowiednio jako $(x_1, x_2, \ldots, x_n)$, $(p_1, p_2, \ldots, p_n)$ i $(R_1, R_2, \ldots, R_m)$, optimum techniczne można zapisać w postaci następującego układu równań:

$$\left.\begin{aligned}
&p_1\left(\frac{\delta x_1}{\delta R_1}\right)=p_2\left(\frac{\delta x_2}{\delta R_1}\right)=\ldots=p_n\left(\frac{\delta x_n}{\delta R_1}\right)\\
&p_1\left(\frac{\delta x_1}{\delta R_2}\right)=p_2\left(\frac{\delta x_2}{\delta R_2}\right)=\ldots=p_n\left(\frac{\delta x_n}{\delta R_2}\right)\\
&\ldots\\
&p_1\left(\frac{\delta x_1}{\delta R_m}\right)=p_2\left(\frac{\delta x_2}{\delta R_m}\right)=\ldots=p_n\left(\frac{\delta x_n}{\delta R_m}\right).
\end{aligned}\right\} \tag{1.1}$$

[2] Rosati (1990), s. 51–53.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Każde z równań dowodzi, że w warunkach optymalnej równowagi produkcyjności krańcowe każdego z zasobów produkcyjnych w różnych zastosowaniach (w różnych rodzajach produkcji) wyrównują się. Nie można zatem poprawić sytuacji, przenosząc zasoby z jednego rodzaju produkcji do innego. Rozwiązaniem powyższego układu względem cen p_i jest pewien wektor cen $[p_1^*, p_2^*, \ldots, p_n^*]$, jednakowych dla wszystkich konsumentów, spełniający układ równań (1.1). Oznacza to optimum wymiany. Z kolei trzeci warunek, czyli optimum produkcji, otrzymuje się, dzieląc kolejne wyrażenia powyższych równań odpowiednio przez siebie, by w rezultacie dla kolejnych $x_1, x_2, \ldots x_n$ otrzymać:

$$\frac{\dfrac{\delta x_1}{\delta R_i}}{\dfrac{\delta x_1}{\delta R_j}} = \frac{\dfrac{\delta x_2}{\delta R_i}}{\dfrac{\delta x_2}{\delta R_j}} = \ldots = \frac{\dfrac{\delta x_n}{\delta R_i}}{\dfrac{\delta x_n}{\delta R_j}} \quad \text{dla } i, j = 1,2,\ldots, m. \tag{1.2}$$

Wreszcie optimum ogólne oznacza wyrównanie się krańcowej stopy substytucji konsumpcji *KSS* i krańcowej stopy transformacji *KST*. Wyjściowy układ równań (1.1) zostanie teraz przekształcony za sprawą podzielenia przez siebie parami kolejnych wyrazów w poszczególnych równaniach. W rezultacie otrzymuje się warunki równowagi dla każdej pary cen *i*, *j*:

$$\frac{p_i}{p_j} = \frac{\dfrac{\delta x_j}{\delta R_1}}{\dfrac{\delta x_i}{\delta R_1}} = \frac{\dfrac{\delta x_j}{\delta R_2}}{\dfrac{\delta x_i}{\delta R_2}} = \ldots = \frac{\dfrac{\delta x_j}{\delta R_m}}{\dfrac{\delta x_i}{\delta R_m}} = \frac{\delta x_j}{\delta x_i} \quad \text{dla } i, j = 1, 2,\ldots, n. \tag{1.3}$$

Tak zdefiniowane pojęcie optimum Pareto może stanowić dogodne kryterium racjonalności i skuteczności polityki gospodarczej. W sytuacji suboptymalnej istnieje bowiem możliwość zwiększenia wartości funkcji dobrobytu i uzyskania dodatkowych korzyści społecznych dzięki świadomemu wyeliminowaniu rozbieżności istniejących między poszczególnymi wartościami krańcowymi. W takim ujęciu istota polityki gospodarczej polega na eliminowaniu odchyleń od optymalnych wartości krańcowych.

Optymalność w sensie Pareto stanowi elegancką konstrukcję teoretyczną, pozwalającą kierunkowo oceniać racjonalność poszczególnych decyzji podejmowanych w ramach polityki gospodarczej. Jej praktyczna przydatność podlega

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jednak wielu ograniczeniom. Krytyka tej koncepcji odnosi się głównie do dwóch kwestii. Po pierwsze, zasada Pareto, zapewniając efektywność ekonomiczną z perspektywy indywidualnego dobrobytu poszczególnych osób, nie bierze pod uwagę kryteriów ogólnospołecznych, w tym zwłaszcza wpływu sprawiedliwego podziału na poziom dobrobytu społecznego. Po drugie, zasada Pareto zapewnia optymalność ekonomiczną pod warunkiem spełnienia szeregu wymogów dotyczących istnienia doskonałej konkurencji, kompletnych rynków, pełnej informacji i braku tzw. efektów zewnętrznych (pozarynkowych). Ponieważ w praktyce rzadko mamy do czynienia z realizacją tych założeń, stosowanie zasady Pareto może dawać wyniki, które odbiegają od społecznego optimum (Greenwald i Stiglitz 1986). Występujące w rzeczywistości odstępstwa od konkurencji doskonałej nakazują więc traktować teorię Pareto z odpowiednim dystansem.

1.2.2. Konkurencja nie jest doskonała

Jak wspomniano wyżej, stany konkurencji doskonałej zdarzają się w rzeczywistości nadzwyczaj rzadko. Jednym z najważniejszych warunków powstania tego zjawiska jest występowanie odpowiednio dużej liczby uczestników rynku, tak aby żaden z nich nie mógł samodzielnie wywierać wpływu na ceny produktów. Producenci, dążąc do maksymalizacji zysku, a zarazem nie mając wpływu na poziom cen, zwiększają wówczas produkcję do momentu zrównania krańcowego kosztu wytworzenia (dodatkowej) jednostki produktu z krańcowym przychodem ze sprzedaży tej jednostki produktu. Ów krańcowy przychód równy jest cenie rynkowej w zakresie całej produkcji. Ta natomiast osiąga wówczas optimum społeczne. W rzeczywistości często ma się jednak do czynienia z sytuacją, w której producenci dysponują możliwością wywierania wpływu na poziom cen, na przykład dzięki posiadaniu dużego udziału w rynku, co zapewnia im tzw. pozycję dominującą. Są to sytuacje monopolistyczne lub oligopolistyczne. Taki producent będzie generalnie oferował do sprzedaży mniejszą ilość produktu, ale po wyższej cenie jednostkowej, osiągając w ten sposób dodatkowy zysk, wynikający z renty monopolowej. W rezultacie poziom produkcji ukształtuje się poniżej społecznego optimum. Podobna w skutkach może być zmowa kilku producentów, mająca na celu utrzymanie zawyżonego poziomu cen na oferowane przez nich wyroby.

Sytuacje monopolistyczne powstają na ogół wtedy, gdy potencjalni konkurenci muszą ponosić duże koszty, aby wejść do danej branży i konkurować z monopolistą. Mówi się wówczas o istnieniu wysokich barier wejścia do (lub wyjścia z) danej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

branży (*barriers to entry, barriers to exit*). Przykładem tego typu działów są sektory chronione przez politykę państwa, takie jak na przykład krajowa produkcja żywności chroniona ze względów bezpieczeństwa wysokimi cłami przed konkurencją z zagranicy, wynalazki krajowych producentów objęte ochroną patentową czy też pewne rodzaje działalności wymagające uzyskiwania koncesji lub zezwoleń, takie jak eksploatacja zasobów naturalnych. Inny przykład stanowią dziedziny, w ramach których podjęcie produkcji wymaga poniesienia bardzo dużych nakładów finansowych na badania i rozwój technologii produkcyjnych (np. produkcja samolotów) lub poniesienia dużych nakładów inwestycyjnych na uruchomienie produkcji (np. energetyka jądrowa). W przypadku wszystkich tych sytuacji stan gospodarki odbiega od społecznego optimum.

Sytuacja monopolistyczna powstaje również wtedy, gdy w danym sektorze występują rosnące przychody ze skali produkcji. Wobec tego krańcowy koszt produkcji jest niższy od kosztu przeciętnego i przedsiębiorstwo ustala cenę sprzedaży na poziomie kosztu przeciętnego, a nie krańcowego. Gdyby monopolista ustalił cenę na poziomie kosztu krańcowego, zgodnie z kryterium optymalności Pareto, ponosiłby wówczas straty w sprzedaży, ponieważ koszty wytwarzania byłyby wyższe od ceny. Monopolista wybierze więc wyższe ceny i niższy poziom produkcji. Sytuacja rosnących przychodów ze skali (spadających kosztów przeciętnych) jest typowa wśród monopoli naturalnych[3]. W tych przypadkach osiągnięcie optimum Pareto jest niemożliwe.

Inny ważny czynnik konkurencji doskonałej stanowi doskonała przenośność i podzielność zasobów. Gdy zasoby nie są w pełni przenośne (sektorowo i geograficznie) ani w pełni podzielne, nie dochodzi do wyrównania krańcowych stóp transformacji, wobec czego struktura produkcji nie jest optymalna. Podobna w skutkach jest niejednorodność wytwarzanych dóbr, z czym można mieć do czynienia w przypadku różnic jakościowych oraz przy uwzględnianiu przez konsumentów czynników pozaekonomicznych, takich jak reputacja marki handlowej czy reklama.

Bardzo istotnym warunkiem doskonałej konkurencji jest wreszcie pełna informacja o cenach, podaży i popycie. W praktyce występuje ona tylko w bardzo

[3] Monopol naturalny powstaje w sytuacji, w której jedno przedsiębiorstwo może wytwarzać dane dobro taniej i efektywniej niż kilka konkurencyjnych firm. Najczęściej wiąże się to z korzyściami skali – kiedy wraz ze wzrostem produkcji spada jednostkowy koszt wytworzenia dobra. Przykładem monopoli naturalnych są sektory sieciowe – wodociągi, gazociągi, przesyłowe linie energetyczne. Zapewnienie konkurencji wymagałoby budowy równoległych sieci dostaw, co byłoby bardzo kosztowne i ekonomicznie nieuzasadnione wówczas, gdy istniejąca sieć jest w stanie zaspokoić potrzeby wszystkich odbiorców.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

szczególnych przypadkach, na przykład na rynkach dóbr jednorodnych, najmniej przetworzonych, takich jak surowce czy paliwa. Brak pełnej informacji o alternatywnych możliwościach zakupu i sprzedaży czy o jakości nabywanych produktów wprowadza element ryzyka do transakcji rynkowych, co w rezultacie ogranicza ich skalę. Niedostatek informacji (bądź asymetria informacji w relacjach nabywca – sprzedawca) powoduje, że poziom produkcji w danej gospodarce kształtuje się poniżej poziomu optymalnego.

1.2.3. Niekompletność rynków

W pełni doskonały mechanizm rynku konkurencyjnego zapewnia możliwość dokonywania transakcji przez wszystkich uczestników rynku, w przypadku wszystkich rodzajów produktów i aktywów oraz na wszystkie możliwe okresy (transakcje bieżące i terminowe). Gdy jeden lub więcej z tych warunków nie jest spełniony, mamy wówczas do czynienia z niekompletnością rynku. Zjawisko to często występuje w rzeczywistej gospodarce. Dobrym przykładem jest sektor usług edukacyjnych. Ich pełna komercjalizacja – czyli sprzedaż przez dostawców na wolnym rynku – spowodowałaby ukształtowanie się równowagi podaży i popytu przy takim poziomie ceny, który eliminowałby z tego rynku większość rodzin o najniższych dochodach. Dzieje się tak dlatego, że koszt dostarczania usług edukacyjnych jest na tyle wysoki, że rodziny o niższych dochodach zmuszone byłyby ograniczyć zakup tych usług lub zrezygnować z nich całkowicie, ponieważ całość ich dochodów pochłonęłyby wydatki na podstawowe utrzymanie (żywność, mieszkanie, odzież, leki).

Z kolei brak rynkowych instrumentów długoterminowego zastawu praktycznie uniemożliwia zaciąganie przez osoby w wieku szkolnym kredytów na edukację. Ten rodzaj niesprawności rynku nie może być usunięty za sprawą samego mechanizmu rynkowego. Przyczyną takiego stanu rzeczy jest wysoki poziom ryzyka związany z długim okresem zwrotu nakładów przeznaczonych na naukę oraz brakiem możliwości zabezpieczenia kredytu. Tego typu ograniczenia istnieją także na rynku usług opieki zdrowotnej. Podobnie jak w przypadku edukacji, wysoki koszt usług zdrowotnych uniemożliwiłby w systemie wolnorynkowym dostęp do tych usług osobom najuboższym[4]. Bez interwencji państwa mechanizm rynkowy w tym zakresie

[4] Przykładem niekompletności rynku jest rynek ubezpieczeń zdrowotnych w Stanach Zjednoczonych. W 2004 roku 45,8 mln Amerykanów w wieku produkcyjnym nie posiadało żadnego ubezpieczenia zdrowotnego. W ogromnej większości były to osoby o niskich dochodach, najczęściej w młodym wieku, które rezygnowały z wykupu ubezpieczenia ze względu na brak środków (Gruber 2007, s. 5). Jedną z głównych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ograniczałby podaż wskazanych usług poniżej poziomu społecznie pożądanego, co obniżałoby inwestycje w kapitał ludzki oraz intelektualny, negatywnie wpływałoby na wydajność pracy i hamowałoby tempo wzrostu gospodarczego. Niekompletność rynków przekłada się na występowanie w gospodarce pozytywnych lub negatywnych efektów zewnętrznych, tj. pozarynkowych.

1.2.4. Efekty zewnętrzne

Efekt zewnętrzny, czyli pozarynkowy (*external effect, externality*), powstaje, gdy działanie jednego uczestnika rynku generuje określone koszty (lub korzyści) dla innego uczestnika rynku i kiedy pierwszy z nich nie ponosi z tego tytułu kosztów (ani nie odnosi korzyści)[5]. Oznacza to, że dodatkowe koszty (lub korzyści) nie są odnotowywane w indywidualnym rachunku ekonomicznym danego uczestnika rynku.

Efekty zewnętrzne dotyczą zarówno produkcji, jak i konsumpcji. Przykładem negatywnego efektu zewnętrznego powstającego w zakresie produkcji jest zanieczyszczanie środowiska naturalnego (powietrza, wody, gleby) przez zakłady przemysłowe, takie jak huty metali, kopalnie odkrywkowe czy zakłady chemiczne. Zanieczyszczenie środowiska niesie szkodliwe skutki dla innych uczestników rynku – zarówno producentów, jak i konsumentów, którzy ponoszą dodatkowe koszty, przeznaczane m.in. na instalowanie filtrów oczyszczających wodę, straty związane ze spadkiem biologicznych zasobów naturalnych wód (ryby), degradację lasów na skutek tzw. kwaśnych deszczów (*acid rain*) czy dodatkową opiekę zdrowotną dla osób narażonych na oddychanie zanieczyszczonym powietrzem. Wskazani użytkownicy faktycznie ponoszą dane koszty, nie są one jednak wliczane do rachunku zysków i strat w hutach i kopalniach, które zanieczyszczają środowisko.

Przykładem negatywnego efektu zewnętrznego w zakresie konsumpcji jest spożywanie niektórych używek, zwłaszcza alkoholu i wyrobów tytoniowych. Jak wykazują liczne badania, palenie tytoniu i nadużywanie alkoholu zwiększa ryzyko zachorowań na groźne choroby i prowadzi do ogólnego pogorszenia stanu zdrowotności

reform podjętych przez administrację prezydenta Baracka Obamy było objęcie części tych osób ubezpieczeniem zdrowotnym dofinansowanym przez państwo.

[5] Efekty zewnętrzne dzieli się często na efekty technologiczne (*technological externalities*), które wpływają bezpośrednio na wyniki finansowe innych uczestników rynku, oraz efekty cenowe (*pecuniary externalities*), które wywierają wpływ na ceny rynkowe, i za pośrednictwem cen na pozostałych uczestników rynku (Greenwald i Stiglitz 1986).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

społeczeństwa. Powoduje to spadek produktywności siły roboczej i kapitału ludzkiego, wzrost kosztów opieki zdrowotnej, a także ogólne pogorszenie jakości życia pozostałych członków społeczeństwa. Podobnie jak w sytuacji zanieczyszczenia środowiska, koszty te nie są uwzględniane bezpośrednio w kosztach wytwarzania i rynkowej cenie używek. W obu przypadkach zasadna staje się zatem interwencja władzy publicznej w celu „zinternalizowania" kosztów zewnętrznych, tak aby nie zniekształcały one rzeczywistego obrazu rachunku ekonomicznego zarówno wśród producentów tych dóbr, jak i ich konsumentów.

Innym przykładem negatywnego efektu zewnętrznego (koszt) jest emisja dwutlenku węgla przez pojazdy samochodowe napędzane silnikami spalinowymi. Powoduje to zagęszczenie dwutlenku węgla w górnych warstwach atmosfery, co prowadzi do ocieplenia klimatu. Koszty związane z ociepleniem klimatu, takie jak skutki podnoszenia się poziomu wód mórz i oceanów, stepowienia lądów oraz nasilanie się intensywności tropikalnych burz i huraganów, ponoszą wszyscy mieszkańcy naszego globu, choć koszty te nie są wliczone ani w cenę samochodów, ani w cenę zużywanego przez nie paliwa.

Inne jeszcze typowe przykłady negatywnych efektów zewnętrznych dotyczą wpływu zaniedbanego sąsiedztwa na wartość nieruchomości, skutków zdrowotnych nadmiernej otyłości, a także nieostrożnego kredytowania dłużników przez dany bank, co wobec systemowych powiązań zwiększa ryzyko niewypłacalności także dla pozostałych banków.

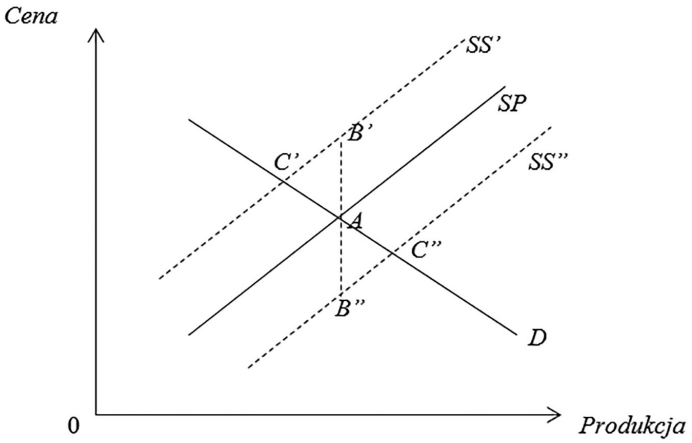
Przykładem z kolei dodatniego efektu zewnętrznego są korzyści związane z edukacją. Nakłady przeznaczane na kształcenie nie tylko umożliwiają podniesienie poziomu wydajności i kreatywności samych pracowników, co przekłada się na wzrost zarówno ich zarobków, jak i indywidualnego dobrobytu, lecz sprawiają także, że ludzie wykształceni przyczyniają się do wzrostu wydajności pracy swoich współpracowników poprzez zapewnienie lepszej organizacji pracy i skuteczniejszego nadzoru. Ponadto ludzie wykształceni łatwiej i lepiej dostosowują się do zmieniających się warunków rynkowych, są bardziej skłonni prowadzić zdrowy tryb życia, rzadziej też wchodzą w konflikt z prawem. Pozwala to zmniejszyć niezbędne nakłady na opiekę socjalną, ochronę zdrowia i ochronę porządku publicznego. Uwolnione w ten sposób środki można przeznaczyć na inne wydatki, o większym potencjale rozwojowym (np. inwestycje w infrastrukturę), bądź też alternatywnie obniżyć podatki. Wszystkie te dodatkowe efekty edukacji niosą określone korzyści społeczne, które nie są jednak uwzględnione w bezpośrednich nakładach na edukację i kształcenie, ponoszonych przez indywidualne osoby. Należy więc pamiętać,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

że łączne korzyści społeczne uzyskiwane dzięki edukacji znacznie wykraczają poza korzyści indywidualne, co uzasadnia interwencję władzy publicznej wszędzie tam, gdzie pojedyncze osoby nie są skłonne – z różnych powodów – aby przeznaczać nakłady finansowe na wykształcenie własne lub swoich dzieci.

Kolejnym przykładem dodatniego efektu zewnętrznego jest zlecona przez daną firmę na potrzeby własnej działalności gospodarczej budowa drogi, która pozostaje jednak ogólnie dostępna i wpływa na poprawę warunków transportu w danym regionie, zaspokajając też tym samym potrzeby lokalnej społeczności. Inne jeszcze przykłady dodatnich efektów zewnętrznych dotyczą zarówno rynku ubezpieczeń zdrowotnych, które zapobiegają pogorszeniu stanu zdrowotnego społeczeństwa i chronią przed epidemiami chorób zakaźnych, jak i komercyjnej działalności badawczo-rozwojowej.

Rysunek 1.2. Efekt zewnętrzny (pozarynkowy) w produkcji



Źródło: opracowanie własne.

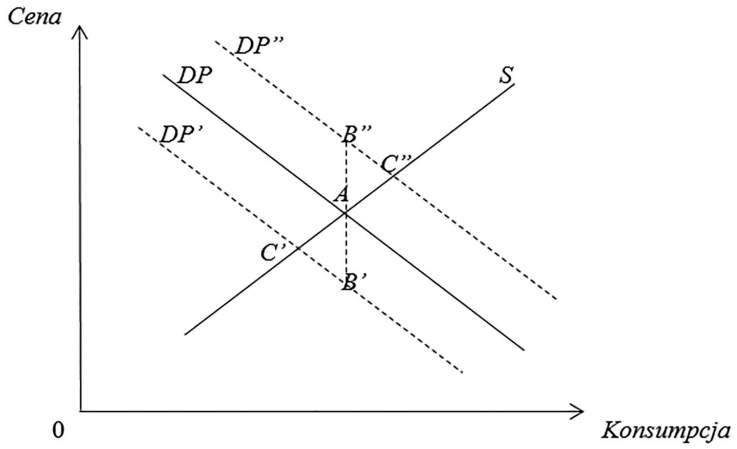
Działanie efektu zewnętrznego – zarówno dodatniego, jak i negatywnego – w zakresie produkcji można przedstawić przy pomocy prostego schematu równowagi cząstkowej (rysunek 1.2). Punktem równowagi *rynkowej* jest *A* – punkt przecięcia się krzywej popytu *D* i prywatnej (indywidualnej) krzywej podaży *SP*. Obecność negatywnego efektu zewnętrznego w produkcji sprawia, że krańcowy koszt społeczny w danym punkcie jest wyższy od krańcowego kosztu prywatnego dla każdego poziomu produkcji. Oznacza to, że społeczna krzywa podaży *SS'* leży powyżej prywatnej krzywej podaży *SP*, a dystans pomiędzy krzywą *SP* i *SS'* (odcinek *AB'*) jest

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

równy wielkości negatywnego efektu zewnętrznego. Punktem równowagi *społecznej* jest punkt przecięcia się krzywej *D* ze społeczną krzywą podaży *SS'* (punkt *C'*). Społecznie optymalny poziom produkcji w tym punkcie jest więc niższy niż w punkcie *A*, a poziom cen staje się wyższy, co odzwierciedla dodatkowe koszty społeczne związane z produkcją. Pole trójkąta *AB'C'* wyraża stratę społeczną z tytułu istnienia efektu zewnętrznego przy poziomie produkcji w punkcie *A*[6].

Analogicznie w przypadku dodatniego efektu zewnętrznego społeczna krzywa podaży *SS"* znajduje się na prawo od krzywej prywatnej, a punktem równowagi społecznej jest *C'*. Dodatni efekt zewnętrzny w produkcji powoduje, że krańcowy koszt społeczny staje się niższy od krańcowego kosztu prywatnego w przypadku każdego poziomu produkcji. W tym punkcie społecznie optymalna wielkość produkcji jest większa niż w punkcie *A*, a poziom cen pozostaje niższy, co odzwierciedla społeczne korzyści związane z produkcją. Pole trójkąta *AB"C"* wyraża wielkość utraconych społecznych korzyści z tytułu istnienia dodatniego efektu zewnętrznego przy produkcji na poziomie *A*.

Rysunek 1.3. Efekt zewnętrzny (pozarynkowy) w konsumpcji



Źródło: opracowanie własne.

Z kolei oddziaływanie efektu zewnętrznego (pozarynkowego) – zarówno dodatniego, jak i ujemnego – w konsumpcji zostało przedstawione na rysunku 1.3.

[6] Wzrost produkcji poza punkt *C'* przynosi rosnące straty społeczne, ponieważ cena rynkowa nie pokrywa pełnego kosztu społecznego produkcji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Punktem równowagi rynkowej jest punkt *A* – punkt przecięcia krzywej podaży *S* i krzywej prywatnego (indywidualnego) popytu *DP*. W przypadku ujemnego efektu zewnętrznego w konsumpcji społeczna krzywa popytu *DS'* leży poniżej krzywej popytu prywatnego *DP* i punktem równowagi społecznej staje się *C'*. Społecznie optymalny poziom konsumpcji jest mniejszy niż w punkcie *A*, a cena pozostaje niższa. Pole trójkąta *AB'C'* wyraża stratę społeczną z tytułu ujemnego efektu zewnętrznego w zakresie konsumpcji, gdy ta utrzyma się w punkcie *A*. Analogicznie w przypadku dodatniego efektu zewnętrznego w konsumpcji punktem równowagi społecznej jest *C"*, a społecznie optymalny poziom konsumpcji jest wyższy niż w punkcie *A*. Pole trójkąta *AB"C"* wyraża utraconą korzyść społeczną z tytułu dodatniego efektu zewnętrznego, gdy konsumpcja zatrzyma się w punkcie *A*.

Kluczową cechą tych pozarynkowych efektów – pozytywnych lub negatywnych – jest to, że nie są one uwzględniane w decyzjach produkcyjnych i inwestycyjnych firm, których działalność staje się źródłem tych efektów. W bilansie wyników danej firmy nie pojawia się ani koszt zanieczyszczenia środowiska w związku z produkcją, ani korzyści płynące dla lokalnej społeczności z tytułu wybudowania powszechnie dostępnej drogi. Oznacza to jednak, że w przypadku generowania zanieczyszczeń ogólnospołeczny koszt wytworzenia danej produkcji jest w rzeczywistości wyższy niż koszt prywatny (rynkowy) i co oznacza negatywny efekt zewnętrzny, podczas gdy wobec wybudowania drogi korzyść społeczna przewyższa korzyść prywatną, co przekłada się na dodatni efekt zewnętrzny. W rezultacie w pierwszym przypadku rzeczywisty poziom produkcji jest za wysoki w stosunku do poziomu społecznie optymalnego, w drugim zaś poziom produkcji pozostaje za niski w stosunku do poziomu społecznie optymalnego. W takiej sytuacji wolny rynek jako mechanizm alokacji zasobów zawodzi, co może stanowić uzasadnienie dla interwencji władzy publicznej w celu skorygowania tej niesprawności.

Co można zrobić, aby skorygować efekt zewnętrzny i przybliżyć rozwiązanie rynkowe do społecznego optimum? Jednym z możliwych sposobów jest stworzenie odpowiednich warunków prawno-instytucjonalnych, które skłonią sektor prywatny do internalizacji efektu zewnętrznego. Amerykański ekonomista Ronald Coase wykazał, że o ile spełnione są pewne określone warunki (dobrze zdefiniowane prawa własności i brak kosztów negocjacji), o tyle sektor prywatny będzie w stanie samodzielnie uwzględnić działanie efektu zewnętrznego – czyli zinternalizować efekt zewnętrzny – poprzez system płatności kompensacyjnych, bez konieczności interwencji władzy publicznej. To właśnie jest istotą słynnego twierdzenia Coase'a (zob. ramka 1.2).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 1.2. Twierdzenie Coase'a

W pracy opublikowanej w 1960 roku Ronald Coase wykazał, że – przy spełnieniu pewnych warunków prawno-instytucjonalnych – skutki niesprawności rynkowej można skutecznie wyeliminować poprzez współdziałanie uczestników rynku (Coase 1960). Twierdzenie Coase'a składa się z dwóch części. Pierwsza z nich mówi o tym, że gdy prawa własności są dobrze zdefiniowane i negocjacje nie pociągają za sobą kosztów, uczestnik rynku generujący efekt zewnętrzny i uczestnik absorbujący efekt zewnętrzny mogą poprzez negocjacje osiągnąć społecznie efektywne optimum. Druga część twierdzenia mówi, że społecznie efektywne optimum nie zależy od tego, który uczestnik rynku posiada prawo własności o tyle, o ile tylko któryś z nich to prawo posiada.

Istotę twierdzenia Coase'a można wyjaśnić przy pomocy prostego przykładu. Załóżmy, że pewien zakład chemiczny odprowadza ścieki produkcyjne do najbliższej rzeki, która jednocześnie jest jedynym źródłem dochodu dla miejscowych rybaków. Zanieczyszczenie wody w rzece obniża poziom połowów, co powoduje utratę części dochodów przez łowiących w niej ryby. Jest to negatywny efekt zewnętrzny, generowany przez produkcję zakładu chemicznego. Sytuację tę można jednak poprawić pod warunkiem, że prawo własności do rzeki jest ściśle określone. Jeśli to rybacy posiadają prawo własności do rzeki, mogą zażądać od zakładu chemicznego rekompensaty za utracony dochód lub zainstalowania urządzeń oczyszczających ścieki. W ostateczności mają możliwość zwrócenia się do sądu z wnioskiem o zamknięcie zakładu. Zakładowi bardziej opłaca się w takiej sytuacji zapłacić rybakom rekompensatę w rozsądnej wysokości (lub zainstalować oczyszczalnię ścieków), niż zamknąć produkcję. Gdy rybacy są w stanie prawidłowo wycenić straty z tytułu zanieczyszczenia rzeki, mogą domagać się rekompensaty w wysokości ponoszonych strat. Wówczas kwota rekompensaty pokrywa w pełni koszty efektu zewnętrznego.

Kiedy z kolei to zakład chemiczny ma prawo własności do rzeki, rybakom opłaca się negocjować z zakładem ograniczenie produkcji i związanych z nią zanieczyszczeń do poziomu zapewniającego im pewien rozsądny poziom dochodu z połowu ryb w zamian za uiszczenie określonej rekompensaty. Jeśli ustalona kwota rekompensaty pokryje straty zakładu chemicznego z tytułu ograniczenia produkcji, wyrówna również w pełni koszt negatywnego efektu zewnętrznego. W obu przypadkach prywatny koszt krańcowy zostanie powiększony o kwotę rekompensaty,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

co ograniczy szkodliwą produkcję i pozwoli osiągnąć społecznie efektywne optimum. Mechanizm, za pośrednictwem którego efekt zewnętrzny stanowiący dodatkowy koszt społeczny ujawnia się w postaci wzrostu kosztu prywatnego, nazywa się internalizacją efektu zewnętrznego.

Twierdzenie Coase'a pokazuje, w jaki sposób można osiągnąć ekonomicznie efektywną alokację zasobów produkcyjnych w warunkach występowania efektów zewnętrznych. Zastosowanie tego twierdzenia w praktyce napotyka jednak na szereg przeszkód, takich jak kwestia ustalenia uczestnika rynku odpowiedzialnego za generowanie efektu zewnętrznego (tzw. problem przypisania – *assignment problem*), niepełna informacja o skutkach efektu zewnętrznego, zjawisko „pasażera na gapę", trudności w oszacowaniu kosztów negocjacyjnych czy możliwość wstrzymania negocjacji i blokowanie porozumienia przez uczestnika najmniej skłonnego do ustępstw (tzw. *holdout problem*). Dlatego też rządy poszczególnych państw z reguły odnosiły się sceptycznie do możliwości wykorzystania twierdzenia Coase'a w polityce gospodarczej.

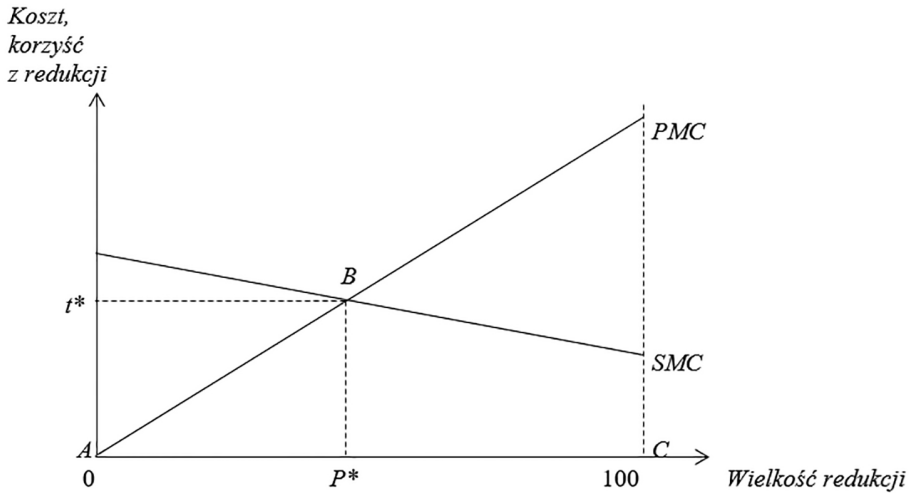
Jednak, jak pokazuje doświadczenie, współdziałanie prywatnych uczestników rynku w eliminacji następstw efektów zewnętrznych napotyka na wiele trudności. Alternatywę stanowi zatem świadoma interwencja władzy publicznej. Korekta efektu zewnętrznego może w tym przypadku polegać na zastosowaniu przez władzę publiczną odpowiednich narzędzi, takich jak podatki, subsydia lub regulacje prawne w postaci nakazów, zakazów, limitów i norm. W przypadku produkcji generującej ujemny efekt zewnętrzny nałożenie podatku korygującego (*corrective taxation, Pigouvian taxation*) w wysokości ujemnego efektu zewnętrznego przesuwa krzywą podaży w górę i ogranicza produkcję do poziomu uzasadnionego społecznie. W przypadku produkcji generującej dodatni efekt zewnętrzny nałożenie subsydium korygującego w wysokości dodatniego efektu zewnętrznego przesuwa krzywą podaży w dół i zwiększa produkcję do poziomu uzasadnionego społecznie. W podobny sposób można korygować efekty zewnętrzne w konsumpcji. Gdy efekt zewnętrzny staje się znaczny lub jego wielkość jest trudna do oszacowania, władza gospodarcza może sięgnąć po instrumenty administracyjne, takie jak nakazy i zakazy.

Czy można określić warunki, w jakich instrumenty administracyjne są bardziej skuteczne w eliminowaniu efektów zewnętrznych niż podatki korekcyjne? Aby odpowiedzieć na to pytanie, należy porównać skutki zastosowania instrumentów administracyjnych, czyli nakazów i zakazów, oraz skutki zastosowania podatków

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w redukowaniu zanieczyszczeń środowiska naturalnego, generowanych przez działalność gospodarczą. Rysunek 1.4 przedstawia sytuację prywatnego zakładu produkcyjnego, który poprzez swoją działalność zanieczyszcza środowisko. Na osi poziomej odłożono wielkość redukcji zanieczyszczeń. Im dalej w prawo od początku układu, tym większa staje się redukcja, aż do całkowitego jej wyeliminowania – w punkcie *A* nie ma żadnej redukcji, podczas gdy w punkcie *C* zanieczyszczenia są całkowicie wyeliminowane. Na osi pionowej zaznaczono koszt prywatny redukcji i korzyść społeczną wynikającą z redukcji. Krzywa *PMC* ukazuje prywatny koszt krańcowy redukowania zanieczyszczeń o kolejne jednostki (koszt ponoszony przez zakład produkcyjny). Krzywa ta ma tendencję wzrostową ze względu na rosnące krańcowe koszty redukcji. Krzywa *SMC* reprezentuje krańcowy koszt społeczny, jakiego można uniknąć, redukując zanieczyszczenia – jest on zarazem równy krańcowej korzyści społecznej z tytułu redukcji. Krzywa ta ma tendencją opadającą ze względu na spadek krańcowych przychodów z redukcji.

Rysunek 1.4. Redukcja zanieczyszczeń w warunkach wolnego rynku (dane w %)



Źródło: opracowanie własne.

Wobec braku interwencji władzy publicznej mechanizm rynkowy doprowadzi oczywiście do rozwiązania w punkcie *A* – ponieważ zakład produkcyjny nie ponosi żadnych kosztów związanych z zanieczyszczeniami, nie otrzymuje też żadnych bodźców, aby je redukować. Taka sytuacja nie jest jednak optymalna ze społecznego punktu widzenia. Jak widać, w danym punkcie korzyść z redukcji

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zanieczyszczeń jest znacznie większa niż ewentualny koszt redukcji – krzywa *SMC* leży powyżej krzywej *PMC*. Opłaca się zatem redukować zanieczyszczenia aż do punktu *B*, w którym koszt redukcji wyrównuje się z korzyścią pochodzącą z redukcji. Punkt *B* wyznacza więc społeczne optimum redukcji zanieczyszczeń na poziomie P^* (zob. Gruber 2007).

W jaki sposób władza publiczna może osiągnąć punkt społecznego optimum? Ma możliwość nałożenia podatku w wysokości t^* na każdą jednostkę generowanych zanieczyszczeń – wówczas zakładowi produkcyjnemu opłaca się redukować zanieczyszczenia tak długo, jak długo koszt redukcji jest niższy od płaconego podatku, czyli do punktu *B*. Powyżej tego poziomu zakładowi produkcyjnemu nie opłaca się już redukować zanieczyszczeń, ponieważ koszty redukcji są wyższe niż uiszczany podatek. Skuteczne zastosowanie podatku korekcyjnego wymaga jednak od władzy publicznej posiadania wiedzy na temat lokalizacji punktu *B*, czyli położenia obu krzywych – *PMC* i *SMC*. Alternatywnie władza publiczna może nałożyć administracyjny nakaz ograniczający generowanie zanieczyszczeń do poziomu P^*, czyli do poziomu optymalnego. Ta metoda także wymaga posiadania wiedzy o położeniu obu krzywych. W takiej sytuacji wybór pomiędzy podatkiem a działaniem administracyjnym dokonywany jest na podstawie innych przesłanek, na przykład po uwzględnieniu skutków budżetowych.

Dodatkową komplikację stanowi włączenie do analizy większej liczby zakładów produkcyjnych o różnych parametrach technicznych produkcji. W takiej sytuacji ma się do czynienia z wieloma różnymi krzywymi *PMC* – powstaje wówczas pytanie o najkorzystniejszy rozkład redukcji zanieczyszczeń pomiędzy poszczególne zakłady. Możliwe w takiej sytuacji do zastosowania warianty polityki ograniczania zanieczyszczeń opisane są w ramce 1.3.

Ramka 1.3. Redukcja zanieczyszczeń w przypadku dwóch zakładów produkcyjnych o różnych kosztach redukcji emisji CO_2 – możliwe warianty polityki*

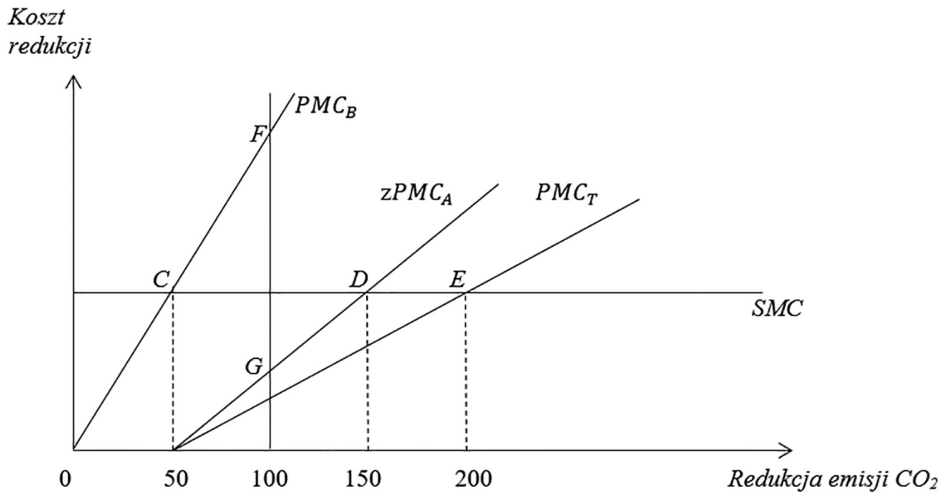
Rozważmy sytuację, w której mamy do czynienia z dwoma zakładami energetycznymi, A i B, które w trakcie wytwarzania energii elektrycznej generują szkodliwą emisję dwutlenku węgla do atmosfery, podczas gdy rząd chce doprowadzić do redukcji emisji CO_2. Następnie należy przyjąć, że redukcja emisji w zakładzie A jest, dzięki zainstalowanej tam nowoczesnej technologii, znacznie tańsza niż redukcja emisji CO_2 w zakładzie B, stosującym technologie tradycyjne. Na

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

rysunku 1.5 przedstawiono przebieg indywidualnych krzywych *PMC* w odniesieniu do obu zakładów oraz krzywej *SMC* (w celu uproszczenia przyjęto, że jest ona linią poziomą). Krzywa PMC_T pokazuje całkowity koszt krańcowy redukcji emisji CO_2 w odniesieniu do obu zakładów łącznie. Przebieg krzywych *PMC* wskazuje, że zakład A może redukować emisję do poziomu 50 jednostek bez ponoszenia żadnych dodatkowych kosztów, a ponadto ma niższe koszty redukcji dwutlenku węgla niż zakład B pod względem każdego poziomu emisji.

Na początku należy zauważyć, że wolny rynek nie skłania do spontanicznej redukcji emisji CO_2, co oznacza, że zakład B w ogóle nie podejmie działań w tym celu, a zakład A może zredukować emisję maksymalnie do poziomu 50 jednostek, bo nie ponosi z tego tytułu żadnych kosztów. Jeśli rządowi zależy na większej skali redukcji, musi wykorzystać określone narzędzia polityki gospodarczej. Jakie skutki przyniesie w tej sytuacji zastosowanie różnych odmian polityki redukcji emisji?

Rysunek 1.5. Redukcja emisji CO_2 z uwzględnieniem wielu firm



Źródło: Gruber (2007), s. 140–143.

Weźmy najpierw pod uwagę regulację ilościową. Zakłada się wówczas, że rząd chce zredukować łączny poziom emisji CO_2 o 200 jednostek i w związku z tym każdy producent otrzymuje nakaz ograniczenia emisji o 100 jednostek. Wówczas zakład B znajdzie się w punkcie *F*, a zakład A w punkcie *G*. Jak widać, nie jest to jest rozwiązanie społecznie efektywne, bo redukcja w zakładzie A kosztuje znacznie mniej niż w zakładzie B, czyli ogólny koszt redukcji pozostaje nadmiernie wysoki

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(punkty *F* i *G*). Bardziej efektywna byłaby redukcja emisji o 150 jednostek w zakładzie A i 50 jednostek w zakładzie B. Rząd mógłby zdecydować się na taki rozkład obciążeń pomiędzy zakładami tylko wtedy, gdy dysponowałby wiedzą na temat faktycznego położenia krzywych PMC_A i PMC_B. Jest to mało prawdopodobne, bo rozkład kosztów produkcji w poszczególnych zakładach stanowi najczęściej informację poufną. Z drugiej strony trzeba jednak podkreślić, że regulacja ilościowa gwarantuje pożądaną redukcję emisji bez względu na koszty.

Weźmy obecnie pod uwagę podatek korekcyjny. Nałożenie podatku w wysokości równej krańcowej korzyści społecznej ($t_c = SMC$) na każdą jednostkę zanieczyszczeń powoduje, że każdy producent będzie redukował emisję CO_2 do momentu zrównania się kosztu krańcowego *PMC* z wysokością podatku. Tak więc zakład B zredukuje spontanicznie emisję o 50 jednostek i zatrzyma się w punkcie *C*, a zakład A zredukuje emisję o 150 jednostek i zatrzyma się w punkcie *D*. Jest to rozwiązanie optymalne z punktu widzenia kosztu ogólnospołecznego (punkt *E*). Podatek korekcyjny pozwala zatem na osiągnięcie najbardziej efektywnego rozwiązania, czyli zrealizowanie założonej redukcji (200 jednostek) możliwie najmniejszym kosztem.

Weźmy wreszcie pod uwagę regulację ilościową, połączoną z handlem zezwoleniami na emisję. Można założyć – tak jak w pierwszym przypadku, że rząd chce zredukować łączny poziom emisji CO_2 o 200 jednostek i że każdy producent otrzymuje nakaz ograniczenia emisji o 100 jednostek. Dodatkowo załóżmy, że każdy producent dostaje zgodę na emisję 100 jednostek CO_2. Początkowo zakład *B* znajdzie się w punkcie *F*, a zakład A w punkcie *G*. W sytuacji, gdy możliwy jest handel zezwoleniami, zakładowi B opłaca się jednak kupować zezwolenia od zakładu A aż do chwili, gdy koszt krańcowy redukcji *PMC* zrówna się z ceną zezwolenia. W rezultacie zakład B zwiększy emisję CO_2 i obniży redukcję ze 100 do 50 jednostek (na pozostałe 50 jednostek ma zezwolenia), a zakład A zmniejszy emisję CO_2 i zwiększy redukcję ze 100 do 150 jednostek. Zaletą tego rozwiązania jest możliwość wykorzystania mechanizmu rynkowego – rząd nie musi znać położenia krzywych *PMC*: to same zakłady produkcyjne decydują o tym, do jakiego momentu opłaca im się handlować zezwoleniami.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

1.2.5. Dobra publiczne

Dobra publiczne to takie dobra, które generują dodatnie efekty zewnętrzne, lecz z różnych powodów nie byłyby wytworzone i dostarczone przez sektor prywatny – albo w ogóle, albo w niedostatecznej ilości. Powodem takiego stanu rzeczy może być nieopłacalność rynkowa, wynikająca z niemożności ograniczenia dostępu do tych dóbr tylko dla tych konsumentów, którzy za nie zapłacą. Tego typu dobra określa się mianem dóbr niewykluczających (*non-excludable*). Należą do nich np. ochrona środowiska, obrona narodowa i straż pożarna a także obiekty takie jak latarnia morska i oświetlenie miejskie. Brak możliwości pobierania opłat od tych, którzy korzystają z danego dobra publicznego, nazywa się w ekonomii problemem „pasażera na gapę" lub „gapowicza" (*free rider problem*). W danej sytuacji sektor prywatny rezygnuje z produkcji tego rodzaju dóbr. Powodem nieopłacalności rynkowej może też być zbyt duże ryzyko związane z nadmiernie długim okresem zwrotu z inwestycji w produkcję danego dobra i niepewnością co do poziomu przyszłych dochodów. Do takich dóbr zalicza się m.in. edukację podstawową, drogi i autostrady, ochronę zdrowia czy podstawowe badania naukowe.

Dobrem publicznym jest również dobro, którego konsumpcja przez danego użytkownika nie ogranicza konsumpcji tego dobra przez innych użytkowników. Takie dobra nazywa się dobrami nierywalizacyjnymi. Należą do nich m.in. prawo i porządek publiczny, wymiar sprawiedliwości, obrona narodowa, straż pożarna, latarnie miejskie, sygnał telewizyjny. Niektóre dobra publiczne (np. latarnia morska czy uliczna, obrona narodowa) spełniają oba warunki – są to tzw. czyste dobra publiczne (*pure public goods*). Inne zaś zapewniają realizację tylko jednego z tych dwóch warunków – na przykład takie dobra jak bezpłatna droga, edukacja czy ochrona zdrowia mogą być użytkowane wyłącznie przez określoną liczbę konsumentów, a odbiór sygnału telewizyjnego może pozostawać płatny dzięki kodowaniu. Dobra takie określamy mianem dóbr publiczno-prywatnych (*public-private goods, impure public goods*). Podział wszystkich dóbr z punktu widzenia dwóch wyróżniających cech – wykluczalności i rywalizacyjności konsumpcji – zawiera tabela 1.1.

Pozytywne efekty zewnętrzne generowane przez dobra publiczne sprawiają, że są one społecznie użyteczne i pożądane, ale mechanizm rynkowy albo nie skłania w ogóle do ich wytwarzania, albo powoduje, że są one produkowane w niedostatecznych ilościach (*underprovisioning*). Podobnie jak w przypadku niekompletności

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

rynków, mechanizm rynkowy okazuje się w takiej sytuacji zawodny. Konieczna okazuje się wówczas interwencja państwa – bądź to w celu skłonienia sektora prywatnego do wytwarzania tych dóbr, bądź w celu ich bezpośredniego dostarczania społeczeństwu.

Tabela 1.1. Dobra publiczne i dobra prywatne

		Czy konsumpcja dobra przez jednego uczestnika rynku ogranicza konsumpcję innych uczestników? (dobra rywalizacyjne i nierywalizacyjne – *rival vs. non-rival goods*)	
		TAK	NIE
Czy można ograniczyć konsumpcję tylko do tych uczestników rynku, którzy za nią zapłacą? (dobra wykluczające i niewykluczające – *excludable vs. nonexcludable goods*)	TAK	czyste dobra prywatne (np. odzież, żywność, meble samochód)	dobra publiczno-prywatne (np. kodowany sygnał TV, dostęp do Internetu)
	NIE	dobra publiczno-prywatne (np. ulica miejska)	czyste dobra publiczne (np. latarnia morska, obrona narodowa, przepisy prawa, stabilność finansowa)

Źródło: opracowanie własne.

Oczywiście, nie wyklucza to sytuacji, w których indywidualne podmioty są gotowe dostarczać pewne dobra publiczne w sposób nieodpłatny z pobudek altruistycznych lub ze względu na poczucie solidarności z innymi jednostkami. Dość często ma się też do czynienia z sytuacją, gdy prywatne firmy finansują w sposób bezpośredni dostarczanie pewnych dóbr publicznych, takich jak edukacja (fundowanie stypendiów), ochrona zdrowia (finansowanie wyposażenia szpitali) czy kultura (sponsorowanie teatrów i filharmonii), lub w sposób pośredni, wspierając organizacje pożytku publicznego. Tego rodzaju działalność charytatywna, choć godna pochwały, ma jednak w zestawieniu ze skalą potrzeb społecznych znaczenie marginalne. Wynika to stąd, że głównym celem takiej aktywności jest najczęściej ocieplanie wizerunku danej firmy (*warm glow model*), a nie korygowanie zawodności rynku.

W jaki sposób państwo może zapewnić dostarczanie społeczeństwu dóbr publicznych na pożądanym poziomie? Po pierwsze, rząd ma możliwość sam dostarczać usługi publiczne poprzez specjalnie powołane do tego celu przedsiębiorstwa państwowe. Ekonomiści są jednak zgodni co do tego, że są one z reguły mniej efektywne niż przedsiębiorstwa prywatne (Mueller 2003). Główne problemy, z jakimi najczęściej ma się tu do czynienia, to nadmierne koszty funkcjonowania, marnotrawstwo zasobów, korupcja, wypieranie sektora prywatnego z danej działalności

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(*crowding out*) oraz nadmierny rozrost biurokracji. Ostatnie z tych zjawisk jest przejawem istnienia zniekształconych bodźców, które kształtują zachowania urzędników państwowych. Zgodnie z teorią agencji (*theory of agency*), gdy władza publiczna powierza określone zadania urzędnikom, zachodzi konflikt między celami władzy publicznej a partykularnymi interesami poszczególnych urzędników, co obniża efektywność działania organów publicznych. Tego rodzaju konflikt celów nosi miano problemu „przełożonego i agenta" (*principal-agent problem*). Z kolei zjawisko szkodliwego rozrostu biurokracji w procesie realizowania celów publicznych jest powszechnie znane i stanowi jeden z najważniejszych argumentów wytaczanych przeciwko angażowaniu się władzy publicznej w realizację działań gospodarczych (zob. m.in. *self-maximizing bureaucracy*, czyli tzw. model Niskanena (1971), lub teorię Lewiatana – *Leviathan theory*, Brennan i Buchanan 1977, 1980). W rezultacie rząd często zawodzi jako bezinteresowny, solidny i efektywny dostawca dóbr publicznych i nie zawsze działa w interesie społecznym.

Po drugie, rząd może powierzyć dostarczanie dóbr publicznych przedsiębiorstwu prywatnemu na podstawie konkurencyjnego przetargu (*competitive bidding*). Indywidualny wykonawca dostarcza wówczas dobra publiczne w ramach i na warunkach określonych w ramach przetargu, co znacznie eliminuje możliwość wystąpienia konfliktu celów, czyli problemu przełożonego i agenta, a także osłabia bodźce sprzyjające rozrostowi biurokracji. Te okoliczności powinny być brane pod uwagę przy dokonywaniu wyboru między bezpośrednim dostarczaniem dóbr publicznych przez przedsiębiorstwa państwowe a otrzymywaniem ich za pośrednictwem przedsiębiorstw prywatnych.

W jaki sposób rząd powinien określać pożądany poziom produkcji dóbr publicznych? Już niemal sto lat temu szwedzki ekonomista Erik Lindahl próbował jako pierwszy odpowiedzieć na to pytanie, wykorzystując koncepcję „krańcowej skłonności do płacenia" (*marginal willingness to pay*). Lindahl przyjął, że każdy obywatel ma jakąś (malejącą) krańcową skłonność do płacenia za dobro publiczne. Innymi słowy, każdy byłby gotów zapłacić pewną cenę za dane dobro, gdyby nie było ono dostarczane bezpłatnie. To wyznacza malejącą funkcję poziomu konsumpcji dóbr publicznych użytkowanych przez daną osobę. Suma indywidualnych preferencji wszystkich obywateli pozwala określić (malejącą) zagregowaną ogólnospołeczną krzywą skłonności do płacenia. Przecięcie tej krzywej z krzywą krańcowego kosztu produkcji dobra publicznego określa optymalny poziom produkcji dóbr publicznych, przy którym społeczna skłonność do płacenia za dane dobro wyrównuje się z krańcowym kosztem produkcji tego dobra (Lindahl 1919).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Choć od strony teoretycznej model Lindahla daje poprawną odpowiedź na pytanie o pożądany poziom dostarczania dóbr publicznych, zastosowanie go w polityce gospodarczej jest praktycznie niewykonalne. Przede wszystkim pojawia się problem z ujawnianiem preferencji indywidualnych (*preference revelation problem*). Uzyskanie wiarygodnych informacji na temat „krańcowej skłonności do płacenia" jest trudne, ponieważ ludzie niechętnie ujawniają swoje preferencje cenowe, i – pytani o to, jaką cenę byliby skłonni zapłacić – mogą oszukiwać, zaniżając te preferencje, aby spróbować „jazdy na gapę" (niech inni zapłacą za dobra publiczne więcej). Ponadto ludzie często nie są w stanie sprecyzować swoich oczekiwań, zwłaszcza w sposób ilościowy, tak aby porównać je z poziomem satysfakcji z nabycia określonych dóbr publicznych (*preference knowledge problem*). Pojawia się wreszcie problem z agregacją preferencji indywidualnych (*preference aggregation problem*) na poziomie makroekonomicznym. Rząd może nie być w stanie oceniać i dodawać preferencji poszczególnych obywateli (np. w jaki sposób osądzić, czy ludzie bardziej cenią sobie szkoły, żłobki, drogi, czy nowe czołgi dla armii). W efekcie trudno jest uzyskać zagregowaną ogólnospołeczną krzywą skłonności do płacenia. Kłopoty z ujawnianiem preferencji, ich porównywaniem i agregowaniem uniemożliwiają wykorzystanie modelu Lindahla w praktycznej polityce.

Wybitny amerykański ekonomista Paul Samuelson sformalizował twierdzenie Lindahla, wykorzystując pojęcie społecznej użyteczności jako sumy użyteczności indywidualnych dla określenia efektywnej podaży dóbr publicznych. Samuelson zwrócił uwagę na to, że przy danych zasobach dostarczanie dóbr publicznych musi jednocześnie oznaczać ograniczenie podaży dóbr prywatnych. Tak zwany warunek optymalności Samuelsona głosi, że optymalna podaż dóbr publicznych jest określona przez punkt wyrównania się sumy krańcowych korzyści z konsumpcji tych dóbr z kosztem krańcowym ich wytwarzania w relacji do dóbr prywatnych (zob. ramka 1.4).

Ramka 1.4. Samuelsona warunek optymalności dostarczania dóbr publicznych

Warunek optymalności Samuelsona określa optymalny poziom produkcji dóbr publicznych (Samuelson 1954). Punktem wyjścia jest stwierdzenie, że w gospodarce, w której zasoby wytwórcze są ograniczone, dostarczenie dodatkowej jednostki dobra publicznego wymaga przesunięcia zasobów od produkcji dóbr prywatnych do produkcji dóbr publicznych, co oznacza spadek produkcji dóbr prywatnych. Proces

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ten jest uzasadniony ekonomicznie aż do momentu, gdy suma korzyści, które indywidualni konsumenci czerpią razem z dodatkowej jednostki dobra publicznego, wyrównuje się z krańcowym kosztem wytworzenia jednostki tego dobra w postaci rezygnacji z dodatkowej jednostki dobra prywatnego.

Przy założeniu, że gospodarka składa się z *n* konsumentów, warunek Samuelsona można zapisać następująco:

$$\sum_{i=1}^{n} MRS_i = MRT, \tag{1.4}$$

gdzie MRS_i jest krańcową stopą substytucji pomiędzy dobrem publicznym a dowolnym dobrem prywatnym dla konsumenta *i*, a *MRT* jest krańcową stopą transformacji (produkcji) pomiędzy tym samym dobrem publicznym a tym samym dowolnym dobrem prywatnym. Jeśli wartości obu dóbr mogą być wyrażone w jednostkach pieniądza, warunek (1.4) przyjmuje postać następującą:

$$\sum_{i=1}^{n} MB_i = MC, \tag{1.5}$$

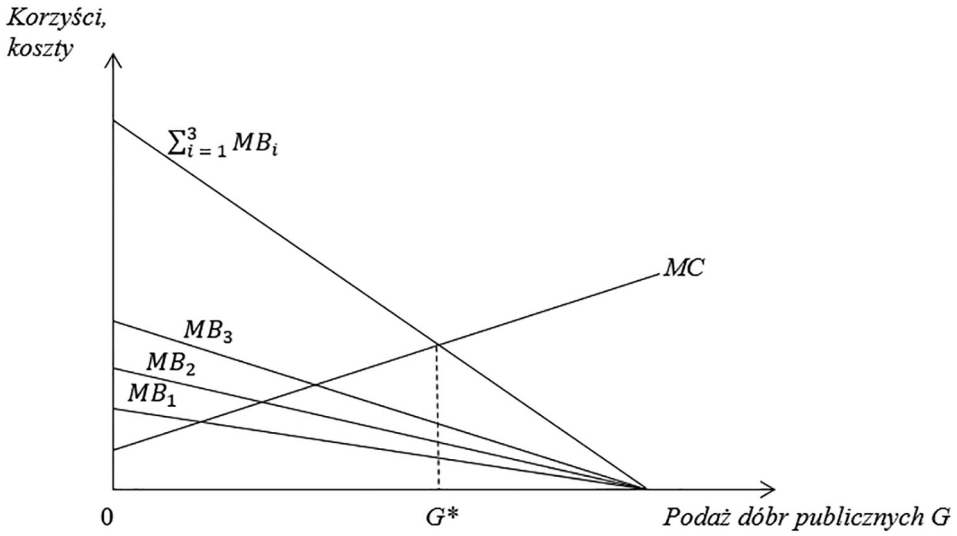
gdzie MB_i jest krańcową korzyścią, jaką odnosi konsument z wykorzystania jednostki dobra publicznego, a *MC* jest krańcowym kosztem dostarczenia tej jednostki. Warunek (1.5) oznacza, że dobro publiczne powinno być dostarczane aż do momentu, gdy suma korzyści odnoszonych przez konsumentów z konsumpcji tego dobra jest co najmniej równa kosztowi jego wytworzenia. Sumowanie indywidualnych korzyści z konsumpcji jednostki dobra publicznego jest możliwe, ponieważ dobra publiczne mają z reguły charakter nierywalizacyjny, co oznacza, że mogą być wykorzystywane jednocześnie przez wielu konsumentów. Gdy spełniony jest warunek (1.5), dalszy wzrost produkcji i dostarczania dóbr publicznych oznacza spadek społecznej funkcji użyteczności, czyli przynosił straty społeczne netto.

Warunek Samuelsona przedstawiony jest na rysunku 1.6, gdzie przyjęto założenie, że w gospodarce funkcjonuje trzech konsumentów. Odnoszone przez nich indywidualne korzyści MB_1, MB_2 i MB_3, związane z konsumpcją danego dobra publicznego, sumują się w postaci krzywej zagregowanych korzyści społecznych $\sum_{i=1}^{3} MB_i$. Ich suma jest równa łącznemu popytowi na dobro publiczne. Punkt przecięcia tej krzywej z krzywą kosztu krańcowego *MC* wyznacza optymalny poziom podaży danego dobra publicznego – G^*. Rysunek 1.6 dowodzi, że warunek

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Samuelsona można interpretować jako warunek łącznej równowagi podaży i popytu na rynku dóbr publicznych i prywatnych.

Rysunek 1.6. Warunek optymalności Samuelsona – interpretacja graficzna



Źródło: opracowanie własne.

Kolejną próbę udzielenia odpowiedzi na pytanie o optymalny poziom produkcji dóbr publicznych podjął w latach 50. ubiegłego wieku amerykański ekonomista Charles Tiebout. Zauważył on, że zaspokajanie potrzeb społecznych w zakresie dóbr prywatnych dokonuje się dzięki mechanizmowi konkurencji, która pozwala na wybór najtańszego dostawcy (tzw. *shopping*). Czy i na ile można ten mechanizm zastosować w odniesieniu do dóbr publicznych?

W przypadku niektórych dóbr publicznych, takich jak obrona narodowa, autostrady, stabilność finansowa czy publiczny system emerytalny nie ma możliwości wyboru dostawcy w ramach danego kraju. Dlatego też odpowiedzialność za dostarczanie tego typu dóbr powinna spoczywać na władzy publicznej szczebla centralnego. W przypadku wielu innych dóbr publicznych funkcjonuje jednak konkurencja, dzięki czemu istnieje możliwość dokonania wyboru dostawcy. Dotyczy to na przykład usług w zakresie szkolnictwa, usług komunalnych (takich jak wywóz śmieci, transport publiczny czy oświetlenie ulic), a także takich dóbr publicznych jak parki,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

tereny rekreacyjne i ścieżki rowerowe. We wszystkich tych przypadkach obywatele, jeśli nie są zadowoleni z poziomu i jakości dostarczanych dóbr publicznych, mogą przenosić się do innych dzielnic lub miast. Oznacza to, że – przynajmniej w pewnym zakresie – istnieje konkurencja w dostarczaniu dóbr publicznych. Możliwość przeprowadzki oznacza, że obywatele, którzy chcą otrzymywać więcej dóbr publicznych (w zamian za wyższe podatki) lub są zainteresowani otrzymywaniem innych dóbr publicznych, będą osiedlać się w takich miejscowościach, w których te oczekiwania zostaną spełnione (głosowanie „nogami" – Tiebout 1956).

Tiebout podkreślał, że możliwość przeprowadzenia się obywateli (zwłaszcza tych zamożniejszych) z danej miejscowości do innej oznacza dla władz lokalnych groźbę utraty dochodu z podatków, głównie dochodowych[7], co powinno skłaniać rządzących do dostarczania większej ilości dóbr publicznych o lepszej jakości. Wynika z tego również, że te dobra publiczne, w przypadku których działa konkurencja i istnieje możliwość wyboru dostawcy, powinny być dostarczane przez szczebel lokalny władzy publicznej. Zgodnie z rozumowaniem Tiebouta, gdy wielu obywateli ma możliwość przenoszenia się z miejsca na miejsce i gdy dostępnych jest wiele miejscowości oferujących różny poziom dostarczania rozmaitych dóbr publicznych, obywatele zaczną osiedlać się w taki sposób, że w każdej z tych miejscowości znajdą się ludzie reprezentujący zbliżony poziom i strukturę popytu na dobra publiczne. Wówczas cała społeczność danej miejscowości może płacić jednakowy podatek ryczałtowy (*lump-sum tax*). W ten sposób model Tiebouta rozwiązuje problemy związane z ujawnianiem (nie zachodzi potrzeba oszukiwania) i agregowaniem preferencji (w każdej miejscowości struktura i poziom popytu pozostają zbliżone).

Choć model Tiebouta stanowi niewątpliwy krok naprzód w porównaniu z modelem Lindahla, jego praktyczne zastosowanie napotyka także na szereg trudności. Mobilność ludzi (w sensie zmiany miejsca zamieszkania) jest w rzeczywistości ograniczona, a niezadowolenie z powodu zbyt małego poziomu dostaw jednego czy kilku dóbr publicznych może nie być wystarczającym bodźcem do przeprowadzki. Wiedza o dobrach publicznych dostępnych w alternatywnych miejscach zamieszkania jest niepełna. Nie w każdym miejscu zresztą wszystkie dobra publiczne mogą być dostarczane (np. w bardzo małych miejscowościach nie ma szkół, teatrów czy basenów). Wreszcie też dobra publiczne dostępne w danej miejscowości, takie jak

[7] W Polsce w 2015 roku udział władz lokalnych (gmin, powiatów, województw) w dochodach z podatku PIT wyniósł łącznie 49,27%, a w dochodach z podatku CIT 22,86%.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

parki, stadiony czy baseny, mogą być wykorzystywane przez mieszkańców z sąsiednich miejscowości, co generuje efekty zewnętrzne lub ciągnione (*spillovers*). To powoduje, że korzyści społeczne są wyższe niż prywatne, co z kolei staje się uzasadnieniem w przypadku zaangażowania wyższych szczebli władzy publicznej w dostarczanie tego rodzaju dóbr.

Istnieje jednak ważna dziedzina, w której model Tiebouta znajduje potwierdzenie w praktyce – mowa o rynku nieruchomości. Rynkowe ceny mieszkań i domów wahają się w zależności od miejsca i lokalizacji – pozostają wyższe tam, gdzie poziom dostarczanych dóbr i usług publicznych jest wysoki (dobry transport publiczny, a zwłaszcza możliwość korzystania z metra, dostępność żłobków i przedszkoli, dobre szkoły, bezpieczne ulice). Oznacza to, że lokalne inwestycje warunkujące dostarczanie dóbr publicznych „kapitalizują się" w cenach lokalnych nieruchomości.

W świetle opisanych trudności z wyceną zapotrzebowania społecznego na dobra publiczne władze zmuszone są sięgać do innych metod ujawniania i szacowania preferencji społecznych. Powszechnie uważa się, że funkcję tę mogą spełniać demokratyczne wybory. Głosując na określone partie polityczne, wyborcy dają jednocześnie wyraz swemu poparciu dla ich programów wyborczych. Ponieważ w ramach tych strategii poszczególne partie określają swoje preferencje i zamiary co do wielkości oraz struktury podaży dóbr publicznych, demokratyczne wybory (a także referenda oraz innego typu głosowania) pozwalają szacować społeczne zapotrzebowanie na dobra publiczne.

1.2.6. Głosowanie większościowe i paradoks głosowania*

Demokratyczne głosowanie nie zawsze zapewnia jednak społecznie optymalny wynik i trwałą równowagę. Przykładowo można założyć, że rząd ma możliwość zwiększenia wydatków na dostarczanie jednego z trzech określonych dóbr publicznych – edukacji, ochrony zdrowia i obrony narodowej. Przed podjęciem decyzji władza chce jednak wiedzieć, jakie są preferencje społeczeństwa, i w tym celu postanawia przeprowadzić głosowanie z udziałem przedstawicieli trzech grup społecznych: X, Y i Z, reprezentujących mniej więcej jednakowe liczebnie elektoraty. Oznaczmy wymienione dobra publiczne odpowiednio jako A, B i C, i przyjmijmy, że indywidualne preferencje obywateli układają się w sposób następujący:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

– grupa X woli A od B, i B od C;
– grupa Y woli C od A, i A od B;
– grupa Z woli B od C, i C od A.

Uporządkowaną kolejność preferencji indywidualnych przedstawia tabela 1.2.

Tabela 1.2. Preferencje indywidualne w odniesieniu do trzech dóbr publicznych

	A (edukacja)	B (ochrona zdrowia)	C (obrona narodowa)
Grupa X	1	2	3
Grupa Y	2	3	1
Grupa Z	3	1	2

Źródło: opracowanie własne.

Wychodząc z założenia, że rząd chce najpierw rozstrzygnąć alternatywę: A lub B, dostrzega się następującą prawidłowość: ponieważ grupy X i Y wolą A od B, a tylko grupa Z woli B od A, można przyjąć, że większość woli A niż B. W kolejnym kroku rząd przechodzi do rozstrzygnięcia alternatywy: A lub C. Wówczas okazuje się, że grupy Y i Z wolą C niż A, a tylko grupa X woli A niż C, w związku z czym można przyjąć, że większość woli C niż A. Na podstawie tych dwóch głosowań przychodzi stwierdzić, że C jest bardziej preferowane niż A, podczas gdy A bardziej preferowane niż B. Jeśli rząd zechce jednak sprawdzić ten rezultat i rozstrzygnąć w odrębnym głosowaniu alternatywę: B lub C, wynik okaże się inny, bo grupy X i Z chętniej opowiadają się za B niż C. Jak widać, przy takim układzie preferencji indywidualnych żaden z wariantów nie jest jednoznacznie najlepszy. Zjawisko to nazywane jest paradoksem głosowania (*voting paradox*). Dowodzi ono, że kolejne głosowania obalają poprzednie wyniki, przy czym w każdym głosowaniu wynik zależy wyłącznie od porządku, w jakim rozstrzygane są kolejne warianty. W rezultacie system demokratyczny nie zapewnia stabilnej równowagi. Może się zatem okazać, że rząd nie będzie w stanie określić społecznej funkcji preferencji w odniesieniu do różnych opcji polityki na podstawie zaobserwowanych indywidualnych funkcji preferencji.

Czy można w jakiś sposób przełamać ten niekończący się cykl głosowań i znaleźć społeczną funkcję preferencji, odzwierciedlającą preferencje indywidualne i zapewniającą stabilne rozwiązanie? Zjawisko to badał amerykański ekonomista Kenneth Arrow, wykazując, że jeśli społeczna funkcja preferencji ma spełniać pewne pożądane właściwości, to większościowe głosowania nie zapewniają stabilnej równowagi politycznej, ponieważ możliwość tworzenia zmiennych koalicji

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

może prowadzić do ciągłych zmian większości parlamentarnej, a w efekcie – do ciągłych zmian w zakresie polityki gospodarczej (zob. ramka 1.5).

Ramka 1.5. Twierdzenie Arrowa o niemożliwości

Kenneth Arrow (1950, 1951) badał możliwość zbudowania ogólnospołecznej funkcji preferencji na podstawie zaobserwowanych indywidualnych funkcji preferencji poszczególnych osób w ramach systemu demokratycznego. Zakładając, że wyborcy mogą wybierać spośród trzech lub więcej opcji polityki, a społeczne preferencje mają pewne określone właściwości[8] oraz że nie ma „dyktatora" narzucającego własne preferencje całemu społeczeństwu, Arrow wykazał, że nie istnieje system polityczny, który byłby w stanie określić ogólnospołeczną funkcję preferencji, spełniającą te wymagania na podstawie preferencji indywidualnych. W szczególności zaś takiej stabilnej społecznej funkcji preferencji nie zapewnia demokratyczny system głosowania większościowego. Twierdzenie to nazywa się twierdzeniem Arrowa o niemożliwości (*Arrow's impossibility theorem*).

Ilustracją twierdzenia Arrowa jest przykład liczbowy przedstawiony w tabeli 1.3. Punkt wyjścia stanowi założenie, że w parlamencie reprezentowane są trzy grupy społeczne: biedni (45% miejsc), klasa średnia (35% miejsc) i bogaci (20% miejsc), przy czym ich udziały parlamentarne są równe udziałom dochodów poszczególnych grup w dochodzie narodowym. Jak widać, żadna z grup nie dysponuje bezwzględną większością w parlamencie. Rozważane są różne warianty opodatkowania dochodów poszczególnych grup. Dla uproszczenia przyjmuje się, że zmiany opodatkowania muszą być neutralne względem budżetu, co oznacza, że suma dochodów podatkowych w każdym wariancie musi być stała. W wariancie początkowym (A) wszyscy płacą jednolitą stawkę podatkową 20%. Ta sytuacja nie zadowala biednych, którzy zawiązują większościową koalicję z klasą średnią, dążąc do obniżenia płaconych przez siebie podatków (odpowiednio do 18% i 19%) kosztem podniesienia podatku wśród bogatych (do 26,25%). W wyniku tego przegłosowany zostanie wariant B. W tej sytuacji bogaci, chcąc obniżyć swoje podatki, mogą jednak zawiązać

[8] Społeczne preferencje powinny spełniać warunki kompletności, przechodniości i niezależności od innych wariantów. Kompletność (*completeness*) oznacza, że zawsze można określić, czy A jest lepsze od B lub że B jest lepsze od A bądź też że A jest takie samo jak B. Przechodniość (*transitivity*) wskazuje na to, że jeśli A jest lepsze od B i zarazem B jest lepsze od C, to A jest lepsze od C. Niezależność (*independence*) od innych wariantów polityki oznacza, że wybór między A i B nie powinien zależeć od C.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

koalicję z biednymi, proponując im dalsze obniżenie podatków kosztem wzrostu opodatkowania klasy średniej. W wyniku powstania nowej koalicji zostanie przegłosowany wariant C, w którym stawki wynoszą odpowiednio 17%, 21,57% i 24%. Lecz ta sytuacja także nie jest stabilna. Klasa średnia może bowiem zaproponować koalicję bogatym, aby obniżyć swoje opodatkowanie kosztem wzrostu opodatkowania biednych. Ilustruje to wariant D, identyczny jak wariant początkowy A. Jak można zauważyć, gdy żadna z grup nie dysponuje absolutną większością w parlamencie, zawsze istnieją bodźce skłaniające przeważającą część grup do zmiany obciążeń podatkowych na swoją korzyść, w efekcie czego żadne głosowanie nie zapewnia stabilnego rozwiązania.

Tabela 1.3. Ilustracja twierdzenia o niemożliwości Arrowa: zmienne koalicje podatkowe (dane w %)

Grupa społeczna	Liczba miejsc w parlamencie	Stawka podatku (wariant A)	Stawka podatku (wariant B)	Stawka podatku (wariant C)	Stawka podatku (wariant D)
Biedni	45	20	18	17	20
Klasa średnia	35	20	19	21,57	20
Bogaci	20	20	26,25	24	20
Suma dochodów podatkowych[a]	–	20	20	20	20

a) W relacji do dochodu narodowego.

Źródło: opracowanie własne na podstawie Stiglitz (2004), s. 194–201.

W praktyce twierdzenie o niemożliwości nie musi być spełnione w każdym momencie i w każdej sytuacji. Wpływ na wynik głosowania w parlamencie mogą mieć też inne czynniki, a nie wyłącznie egoistyczna kalkulacja. Wśród tych dodatkowych uwarunkowań największą rolę odgrywają asymetria oraz niedostatek informacji wśród głosujących, wpływ określonej ideologii lub religii na wybór dokonywany w trakcie głosowania, poziom środków finansowych dostępnych podczas kampanii wyborczej, a także wpływ na głosowanie rozmaitych grup nacisku i lobbingu.

1.2.7. Dobra społecznie pożądane

Dobra społecznie pożądane (*merit goods*) służą zaspokajaniu potrzeb, które są zgłaszane przez ogół obywateli i społecznie uzasadnione (*merit wants*), ale

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

na zaspokojenie których sektor prywatny z różnych powodów nie zgłaszałby dostatecznie dużego popytu. Dóbr takich jest bardzo wiele, zalicza się do nich np. pasy bezpieczeństwa montowane obowiązkowo w samochodach, kaski motocyklowe, foteliki samochodowe dla dzieci, uziemienia w gniazdkach elektrycznych, obowiązkowe szczepienia przeciw chorobom epidemicznym, zakaz spożywania narkotyków, powszechny obowiązek szkolny czy obowiązek odprowadzania składki emerytalnej. Dobra te zaspokajają istotne potrzeby związane z zapewnieniem bezpieczeństwa i poprawą dobrobytu. Pozyskanie tych dóbr wymaga poniesienia dodatkowych kosztów, co powoduje, że wiele osób zrezygnowałoby z nich, gdyby przyszło im płacić za nie cenę rynkową. Dobra te zaspokajają jednak społecznie uzasadnione potrzeby, co przynosi dodatnie efekty zewnętrzne, i dlatego ich użytkowanie powinno znajdować wsparcie ze strony państwa, niezależnie od indywidualnych preferencji poszczególnych członków społeczeństwa.

Istnienie pozytywnych efektów zewnętrznych sygnalizuje, że mechanizm rynkowy nie zapewnia dostatecznie dużej produkcji dóbr społecznie pożądanych oferowanych po odpowiednio niskich cenach, tak aby ich konsumpcja osiągnęła rozmiary oczekiwane przez państwo. W takiej sytuacji istnieje uzasadnienie dla ingerencji państwa w proces alokacji zasobów w celu zwiększenia podaży dóbr społecznie pożądanych i konsumpcji tych dóbr przez obywateli.

Interwencja w zakresie dóbr społecznie pożądanych oznacza w istocie zastąpienie wyborów konsumpcyjnych, dokonywanych samodzielnie przez jednostki, decyzjami innego podmiotu – w tym przypadku decyzjami władzy publicznej. Taka praktyka, zwana paternalizmem, najczęściej znajduje uzasadnienie właśnie w istnieniu dodatnich efektów zewnętrznych. Paternalizm może jednak niekiedy budzić sprzeciw zwolenników nieograniczonej wolności gospodarczej, zwłaszcza gdy dotyczy wymogów narzucanych obywatelom w celu ochrony ich życia lub zdrowia, ograniczających zarazem swobodę ich postępowania. Istnieje także ryzyko tego, że zakres decyzji podejmowanych w sposób paternalistyczny może przekroczyć granice uzasadnione wielkością dodatnich efektów zewnętrznych, co przełoży się na nadmierne ograniczenie wolność wyboru wśród poszczególnych jednostek.

1.2.8. Niesprawność rynku na szczeblu makroekonomicznym

Wymienione wyżej niesprawności rynku mają charakter mikroekonomiczny, ponieważ odnoszą się do zjawisk zachodzących na poszczególnych, indywidualnych rynkach. Interwencja państwa prowadząca do zmiany struktury alokacji

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zasobów produkcyjnych może być skutecznym sposobem korekty tych niesprawności i skierowania gospodarki bliżej optimum Pareto. Gospodarka rynkowa jest jednak podatna również na inne niesprawności mechanizmu rynkowego, występujące na szczeblu makroekonomicznym. Odnosi się to w szczególności do zjawiska cyklu koniunkturalnego i reakcji gospodarki na wstrząsy podażowe oraz popytowe.

Szczególnym przejawem zawodności rynku jest bezrobocie. Neoklasyczna teoria ekonomii zakładała, że konkurencyjna gospodarka rynkowa powinna w każdym momencie zapewniać pełne zatrudnienie. Keynes zwrócił jednak uwagę na to, że bezrobocie jest permanentną cechą gospodarki rynkowej, ponieważ przyjęte przez (neo)klasyków założenia dotyczące doskonale giętkich cen i płac oraz pełnej mobilności siły roboczej w rzeczywistości nie są spełnione, przynajmniej w krótkim okresie. Sztywność płac skierowana w dół, będąca wynikiem instytucjonalnych ograniczeń na rynku pracy (kontrakty płacowe, umowy zbiorowe), oraz ograniczona mobilność zasobów pracy, częściowo na skutek przepisów prawa pracy chroniących miejsca jej wykonywania (*Employment Protection Legislation* – EPL), utrudniają dostosowanie podaży pracy do zmian popytu wynikających z przebiegu cyklu koniunkturalnego i postępu technologicznego. W rezultacie mechanizm rynkowej autokorekty (samoregulacji) przestaje działać lub działa w ograniczonym zakresie. Powoduje to, że w każdym momencie część dostępnych w gospodarce zasobów pracy pozostaje nieaktywna.

Ponadto Keynes zwracał także uwagę na krótkowzroczność oraz zmienność zachowań uczestników rynku, a w szczególności na skłonność do przesadnego optymizmu w okresach *prosperity* i nadmiernego pesymizmu w okresach recesji. Jego zdaniem, tego typu postawy są przejawem funkcjonowania „instynktów zwierzęcych" (*animal spirits*)[9], charakteryzujących zachowania uczestników rynku. W wyniku odchylania się – w górę oraz w dół – indywidualnych oczekiwań od rzeczywistości dochodzi do zjawiska nadmiernego inwestowania i spekulacji, co kończy się pęknięciem bańki spekulacyjnej i recesją. Cykliczny charakter rozwoju gospodarki rynkowej stanowił, w opinii Keynesa, uzasadnienie dla interwencji ze strony państwa, z założenia mającej zmierzać do stabilizowania wzrostu gospodarczego wokół pewnej linii długookresowego trendu, a w szczególności wyrównywania brakującego popytu w okresach słabej koniunktury, tak aby unikać okresowych załamań poziomu produkcji i zatrudnienia.

[9] Keynes użył tego terminu w *Ogólnej teorii zatrudnienia, procentu i pieniądza* (1936, s. 161–162), aby opisać spontaniczne zachowania przedsiębiorców i konsumentów, uznane przez niego za charakterystyczne dla ludzkiej natury.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Keynesowskie uzasadnienie w przypadku polityki stabilizacyjnej zyskało wielką popularność w okresie następującym bezpośrednio po II wojnie światowej. Dzięki polityce pobudzania popytu rządom krajów rozwiniętych udało się zapewnić przez wiele lat stabilny wzrost gospodarczy i wysoki poziom zatrudnienia. Wyzwaniem dla koncepcji Keynesa stały się dopiero zjawiska stagflacji z lat 70. i 80. XX wieku oraz rozwój monetaryzmu, teorii racjonalnych oczekiwań i teorii realnego cyklu koniunkturalnego (*real business cycle theory*). Zwłaszcza ostatnia z nich, wyjaśniająca cykliczne wahania produkcji i popytu zmianami zachodzącymi w technologiach produkcyjnych oraz reakcjami na te zmiany ze strony racjonalnie zachowujących się uczestników rynku, stała się instrumentem obrony tezy o skuteczności funkcjonowania rynkowego mechanizmu samoregulacji i argumentem na rzecz krytyki dotyczącej polityki stabilizacyjnej. Tymczasem, mimo prowadzenia bardzo obszernych badań, teoria realnego cyklu koniunkturalnego pozostaje jednak nadal obszarem sporów, a polityka stabilizacyjna, mimo licznych zastrzeżeń, jest wciąż popularnym narzędziem polityki gospodarczej.

1.2.9. Polityka gospodarcza jako reakcja władzy publicznej na szoki zewnętrzne

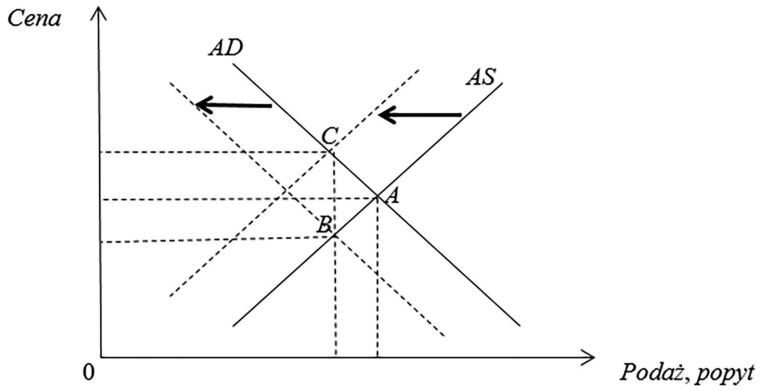
Stabilizacja makroekonomiczna w ramach cyklu koniunkturalnego polega na zmianach łącznego (zagregowanego) popytu w odpowiedzi na autonomiczne zmiany popytu prywatnego. Głównymi narzędziami zmian łącznego popytu są polityka fiskalna i polityka pieniężna. Powinny one mieć charakter antycykliczny, aby niwelować spadek zagregowanego popytu w okresach słabej koniunktury i hamować wzrost zagregowanego popytu w okresach zbyt szybkiego wzrostu. Oprócz wygładzania cyklu koniunkturalnego zadaniem polityki stabilizacyjnej powinno być jednak także reagowanie na jednorazowe wstrząsy (szoki) zarówno popytowe, jak i podażowe. Można wykazać, że właściwe przeciwdziałanie zależy od charakteru wstrząsu oraz jego pochodzenia i że bynajmniej nie we wszystkich przypadkach polityka stabilizacyjna może być skutecznym sposobem reagowania na wstrząsy.

Kwestię tę można łatwo wyjaśnić przy pomocy prostego modelu zagregowanej podaży i zagregowanego popytu, zwanego modelem AD-AS (model został opisany szczegółowo w rozdziale 2). Na rysunku 1.7 łączna podaż i łączny popyt są odpowiednio rosnącą i malejącą funkcją cen. W krótkim okresie zagregowana podaż reaguje pozytywnie na wzrost ceny, ponieważ wzrost cen w warunkach sztywnych płac nominalnych powoduje spadek płac realnych, co oznacza spadek

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kosztów produkcji i wzrost zysków. W tej sytuacji producenci zwiększą produkcję i podaż. Z kolei zagregowany popyt reaguje negatywnie na wzrost ceny, ponieważ wzrost cen zmniejsza realną wartość aktywów nominalnych (zasoby pieniężne), co pociąga za sobą spadek konsumpcji. Krzywa *AS* jest więc krzywą rosnącą, podczas gdy krzywa *AD* krzywą opadającą.

Rysunek 1.7. Szok podażowy i szok popytowy w modelu AS-AD



Źródło: opracowanie własne.

W ramach modelu AS-AD należy rozróżnić dwa rodzaje szoków zewnętrznych, czyli egzogenicznych. Szok podażowy (*supply shock*) oznacza zmianę dotychczasowej relacji pomiędzy produktem potencjalnym a ceną w wyniku zmian po stronie kosztów produkcji. Przykładem negatywnego (niekorzystnego) szoku podażowego jest wzrost cen paliw na rynkach światowych, powodujący spadek produktu potencjalnego ze względu na wzrost kosztów produkcji i spadek zysków. W podobny sposób oddziałuje egzogeniczny wzrost płac lub nałożenie podatku od wykorzystywanych zasobów. Przykładem pozytywnego (korzystnego) szoku podażowego jest wykorzystanie wynalazku i zastosowanie nowej, wydajniejszej technologii produkcji, pozwalającej obniżyć koszty i w ten sposób zwiększyć poziom produktu potencjalnego. Z kolei szok popytowy (*demand shock*) oznacza zmianę dotychczasowej relacji pomiędzy popytem i ceną. Przykładem negatywnego szoku popytowego jest spadek dochodów wśród gospodarstw domowych w wyniku podniesienia podatków dochodowych, co zmniejsza konsumpcję przy danych cenach. I odwrotnie, trwałe obniżenie podatków jest przykładem pozytywnego szoku popytowego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W każdym przypadku efektem tych wstrząsów jest zmiana położenia krzywej podaży lub krzywej popytu (nie należy tego mylić z przesuwaniem się wzdłuż krzywych przy niezmienionym ich położeniu). W zależności od tego, która krzywa ulega przesunięciu, stosowana polityka musi mieć jednak odmienny charakter. Negatywny szok popytowy przesuwa krzywą popytu w lewo w dół, podczas gdy położenie krzywej podaży nie ulega zmianie. Punkt równowagi przesuwa się z *A* do *B*, co oznacza spadek podaży oraz spadek cen. Z kolei negatywny szok podażowy przesuwa krzywą podaży w lewo w górę, podczas gdy położenie krzywej popytu nie ulega zmianie. Punkt równowagi przesuwa się z *A* do *C*, co oznacza spadek podaży i wzrost cen. A więc tym, co odróżnia oba szoki od siebie, jest różnokierunkowe zachowanie się cen przy spadku produkcji.

Dokonane rozróżnienie pozwala zrozumieć granice skuteczności polityki stabilizacyjnej. Ekspansywna polityka fiskalna lub ekspansywna polityka pieniężna oznaczają egzogeniczny wzrost popytu. Powoduje to przesunięcie krzywej popytu w prawo górę do poprzedniego położenia i w ten sposób eliminuje skutki pierwotnego szoku popytowego. Taka była właśnie reakcja polityki makroekonomicznej wśród większości krajów rozwiniętych w następstwie kryzysu finansowego w 2008 roku. Kryzys, związany ze spadkiem cen nieruchomości w warunkach wysokiego zadłużenia gospodarstw domowych, spowodował spadek wartości aktywów majątkowych i w konsekwencji doprowadził do ostrego spadku popytu konsumpcyjnego. Z kolei, gdy przyczyną naruszenia początkowej równowagi jest egzogeniczny wzrost popytu, powodujący przesunięcie krzywej popytu w prawo w górę, właściwą reakcją staje się zaostrzenie polityki fiskalnej i/lub pieniężnej, co przekłada się na przesunięcie krzywej popytu w lewo w dół do poprzedniego położenia. Mimo to ani polityka fiskalna, ani polityka pieniężna nie wpływają na położenie krzywej podaży, co oznacza, że nie są one w stanie neutralizować skutków szoków podażowych.

W szczególności, gdy ma się do czynienia z negatywnym szokiem podażowym, takim jak na przykład wzrost cen surowców czy paliw, pobudzanie popytu przy pomocy polityki fiskalnej i/lub pieniężnej prowadzi do jeszcze większego wzrostu cen i jedynie przejściowego wzrostu produkcji. Wzrost produkcji będzie tylko przejściowy, ponieważ w dłuższym okresie krzywa podaży przyjmuje kształt pionowy w związku z powrotem stopy bezrobocia do poziomu naturalnego. Oznacza to, że w dłuższym okresie polityka zmian zagregowanego popytu jest całkowicie nieskuteczna w neutralizowaniu szoków podażowych.

Powracając do krótkiego okresu, skuteczność polityki zmian popytu staje się zależna od stopnia nachylenia krzywej zagregowanej podaży. W warunkach niskiej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

elastyczności rynku pracy, gdy wynagrodzenia pozostają sztywne, krótkookresowa krzywa podaży jest płaska, co oznacza, że wzrost produkcji nie pociąga za sobą wzrostu płac i kosztów jednostkowych. W takich warunkach polityka zmian popytu okazuje się bardzo skuteczna. Z kolei w warunkach giętkich wynagrodzeń i elastycznego rynku pracy krótkookresowa krzywa podaży jest stroma, a wzrost produkcji powoduje wzrost płac i kosztów jednostkowych. W takich warunkach polityka zmian popytu staje się mało skuteczna w przypadku pobudzania wzrostu produkcji. W rezultacie wybór instrumentu stabilizacyjnego oraz skuteczność polityki zależy nie tylko od charakteru szoku zewnętrznego, lecz także od innych czynników, w tym od instytucjonalnej charakterystyki rynku pracy.

Wstrząs podażowy może być też skutkiem zmian instytucjonalnych, jakie zachodzą w otoczeniu gospodarczym firm. Obniżenie barier administracyjnych i biurokratycznych (na przykład poprzez wprowadzanie ułatwień w uzyskiwaniu zgody na inwestycje, zmniejszanie liczby kontroli skarbowych czy ograniczanie obowiązków sprawozdawczych oraz dokumentacyjnych w przypadku małych i średnich przedsiębiorstw) powoduje przesunięcie krzywej *AS* w prawo i zwiększa poziom produktu potencjalnego. Z kolei nałożenie dodatkowych obowiązków administracyjnych podnosi koszty produkcji i przesuwa krzywą *AS* w lewo.

1.2.10. Redystrybucja i drugie podstawowe twierdzenie ekonomii dobrobytu

Jak wspomniano wyżej, równowaga gospodarcza, która spełnia warunek optymalności Pareto, nie zapewnia automatycznie realizacji postulatu sprawiedliwego podziału. Innymi słowy, nawet gdyby spełnione były rygorystyczne założenia konkurencji doskonałej, osiągnięty stan równowagi – mimo że ekonomicznie super-efektywny – mógłby prowadzić do takiego podziału wytworzonego produktu, który byłby uznany za społecznie niesprawiedliwy. W danej sytuacji społeczeństwo dążyłoby do jego zmiany. Zmierzanie do osiągnięcia innej niż pierwotna dystrybucji dochodów, dyktowane chęcią uwzględnienia wymogów sprawiedliwości społecznej, może więc także stanowić uzasadnienie dla interwencji władzy publicznej w gospodarce.

Redystrybucja, czyli wtórny podział dochodów w społeczeństwie, polega na przesuwaniu zasobów finansowych od jednych grup społecznych do innych. Wielkość wytworzonego w gospodarce produktu zależy przede wszystkim od wielkości zasobów produkcyjnych – takich jak siła robocza, kapitał ludzki, kapitał fizyczny

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(wraz z technologiami) i zasoby naturalne, a także od efektywności, z jaką zasoby te są wykorzystywane w produkcji. Władza publiczna może przyczyniać się do zwiększenia poziomu tego produktu poprzez wpływ na efektywniejszą alokację zasobów i usprawnienie działania rynków. Może także oddziaływać na szybszy wzrost niektórych zasobów, np. poprzez zwiększanie nakładów na edukację w celu przyspieszenia budowy kapitału ludzkiego.

Niemal każdy rząd będzie jednak na ogół również chciał wpływać na podział wytworzonego produktu. Gdyby bowiem powstrzymał się od jakiejkolwiek interwencji w zakresie redystrybucji, z podziału tego produktu wykluczone byłyby grupy społeczne niepracujące, czyli nieuczestniczące – z różnych powodów – bezpośrednio w tworzeniu produktu społecznego. Chodzi o takie grupy jak ucząca się młodzież, osoby w podeszłym wielu (emeryci), chorzy i niepełnosprawni, osoby bezrobotne. Taki stan rzeczy byłby jednak nie do zaakceptowania przez członków zbiorowości, gdzie respektowana jest generalnie zasada solidarności międzyludzkiej i międzypokoleniowej. Jeśli więc przyrównać produkt wytworzony w ramach gospodarki w danym przedziale czasu (na przykład w ciągu roku) do bochenka chleba, to należy zauważyć, że społeczeństwo oczekuje od władzy publicznej, że ta zechce zapewnić taki podział tego „bochenka", który – przynajmniej w pewnym zakresie – uwzględni społeczne potrzeby i oczekiwania, a także będzie zgodny z powszechnie wyznawanymi wartościami, takimi jak solidarność, spójność i sprawiedliwość społeczna.

W realizacji tych oczekiwań władza publiczna musi wyjść poza ramy określone analizą optymalności Pareto. Kryterium to zapewnia wprawdzie ekonomiczną efektywność, ale w żaden sposób nie przesądza o tym, czy osiągnięty stan równowagi jest społecznie pożądany. Można wykazać, że optimum Pareto będzie także spełnione w pewnych skrajnych sytuacjach, nierealizujących w najmniejszym stopniu elementarnego poczucia sprawiedliwości społecznej, np. gdy najuboższa część członków społeczeństwa pozostanie w ogóle pozbawiona środków na przeżycie albo gdy państwo będzie rządzone przez dyktaturę faworyzującą tylko jedna grupę społeczną. Dzieje się tak dlatego, że stan optymalny w rozumieniu Pareto zapewnia jedynie ogólną efektywność w relacji do określonego początkowego rozkładu zasobów produkcyjnych (Sen 1970; Coles i Hammond 1995; Acocella 2002, s. 95), przesądzającego z kolei o ostatecznym podziale dochodów. Takich początkowych rozkładów zasobów pomiędzy poszczególnych członków społeczeństwa może być teoretycznie bardzo wiele, a każdy z nich – w warunkach doskonałej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

konkurencji oraz kompletności rynków – będzie prowadził do osiągnięcia stanu optymalnego w sensie Pareto.

Aby zapewnić jednak bardziej sprawiedliwy podział i zarazem utrzymać wysoką efektywność, społeczeństwo (lub reprezentująca je władza publiczna) może dokonać najpierw odpowiedniego, tj. bardziej sprawiedliwego, podziału zasobów produkcyjnych między poszczególne jednostki, a dopiero potem zezwolić na produkcję i wymianę w ramach konkurencyjnego mechanizmu rynku. Można w ten sposób osiągnąć optimum Pareto, lecz przy innym początkowym rozkładzie zasobów, spełniającym określone wymogi co do ostatecznej redystrybucji dochodów. To właśnie stanowi istotę drugiego podstawowego twierdzenia ekonomii dobrobytu.

Zgodnie z nim, gdy mechanizm rynku jest wspierany w niezbędnym zakresie przez interwencję państwa, można połączyć efektywność w produkcji i sprawiedliwość w podziale. Warunkiem utrzymania optymalności Pareto staje się dokonanie początkowego podziału zasobów w taki sposób, aby nie zakłócić funkcjonowania rynkowego mechanizmu alokacji zasobów w procesie produkcji.

Problem z praktycznym wykorzystaniem drugiego twierdzenia dotyczy tego, że jakakolwiek redystrybucja ma z reguły wpływ na ekonomiczną efektywność i z reguły jest to wpływ negatywny[10]. Wymienność (*trade-off*) zachodząca między efektywnością w produkcji a sprawiedliwością w podziale wynika stąd, że im wyższa jest skala redystrybucji, tym słabsze są bodźce do podejmowania i rozwijania działalności gospodarczej oraz tym większa w efekcie staje się strata w zakresie produkcji i dochodu z powodu zmniejszonej podaży czynników produkcji – pracy i kapitału.

Co więcej, sam mechanizm redystrybucji jest kosztowny, ponieważ wymaga osobnych nakładów na jego utrzymanie (administracja podatkowa, ubezpieczenia społeczne), co dodatkowo zmniejsza zasoby zainwestowane w produkcję. Koszt redystrybucji będzie odpowiednio większy, jeśli podatki i transfery dochodu mają być właściwie – czyli indywidualnie – adresowane. Stawia to przed politykami istotny problem. Przykładowo można wspierać słabsze finansowo grupy społeczne, obniżając stawki podatku VAT na artykuły pierwszej potrzeby, takie jak podstawowa żywność, stanowiące dominującą część wydatków w przypadku osób najuboższych. Zaletą takiego rozwiązania jest jego prostota, ale wadą pozostaje to, że wówczas

[10] Oczywiście, redystrybucja może poprawić efektywność ogólnospołeczną, gdy straty powodowane podatkami są z nadwyżką kompensowane przez wytwarzane dzięki podatkom dobra publiczne (edukacja, opieka zdrowotna, pomoc społeczna).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

z ulgi tej korzystają wszyscy, w tym także najbogatsi obywatele, co przekłada się na dodatkowe koszty i staje społecznie niesprawiedliwe. Bardziej uczciwe byłoby zastąpienie obniżonej stawki VAT bezpośrednimi subsydiami dochodowymi dla najuboższych członków społeczeństwa – wymaga to jednak funkcjonowania rozbudowanej administracji, będącej w stanie rzetelnie oceniać sytuację dochodową poszczególnych obywateli. Taki „sprawiedliwy" system może mimo wszystko okazać w praktycznym zastosowaniu się znacznie kosztowniejszym przedsięwzięciem. Te i tym podobne wątpliwości trzeba rozstrzygać na podstawie bilansu korzyści i strat, z uwzględnieniem społecznych ograniczeń.

1.3. Normatywna teoria polityki gospodarczej

Omówione wyżej niesprawności rynku wskazują na to, że istnieją obszary gospodarki, w przypadku których interwencja władzy publicznej może okazać się pożądana. Na tej podstawie da się ustalić, jakie działania powinna podejmować władza publiczna, aby korygować istniejące niesprawności mechanizmu rynkowego lub łagodzić ich negatywne skutki. Tego typu zagadnieniami zajmuje się normatywna teoria polityki gospodarczej, której podstawy będą omówione w niniejszym podrozdziale. Z kolei pozytywna teoria polityki gospodarczej zajmuje się tym, jak w rzeczywistości funkcjonuje władza publiczna i czy istotnie realizuje założenia wynikające z teorii normatywnej. Zagadnienia odnoszące się do pozytywnej teorii polityki gospodarczej zostaną omówione w dalszej części rozdziału[11].

Polityka gospodarcza to część polityki państwa, polegająca na takim kształtowaniu działalności gospodarczej przez władzę publiczną, które zapewni możliwie szybki i zrównoważony rozwój gospodarczy kraju. W ramach polityki gospodarczej należy więc początkowo zdefiniować cele, a następnie zastosować określone metody, środki i sposoby służące osiąganiu tych celów (zob. Sulmicki 1978, Ćwikliński 2001, Marczewski 2008). Teoria normatywna ma służyć właśnie prawidłowemu formułowaniu celów polityki oraz sposobów i metod ich realizacji. W celu przeprowadzenia niniejszej analizy przyjęto dość abstrakcyjne założenie, że państwo jest obiektywnym, bezstronnym i racjonalnym podmiotem, reprezentującym

[11] Podział na pozytywną i normatywną teorię ekonomii oraz polityki gospodarczej jest powszechnie stosowany w literaturze (zob. Friedman 1953, Sen 1979, Caplin i Schotte 2008). Niekiedy wyróżnia się również deskryptywne podejście do polityki gospodarczej (Kaja 2014), ale jest ono, zdaniem autora niniejszej publikacji, odmianą pozytywnej teorii polityki gospodarczej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w swym działaniu interes ogólnospołeczny. Założenie to zostanie uchylone w dalszej części rozważań.

1.3.1. Funkcje polityki gospodarczej

W pierwszej kolejności należy zacząć od określenia ogólnych funkcji polityki gospodarczej, wynikających z omówionych wcześniej niesprawności rynku. Pozwoli to na zarysowanie głównych obszarów interwencji państwa. Najczęściej mówi się o trzech zasadniczych funkcjach polityki gospodarczej, określanych też czasem jako tzw. triada musgraviańska od nazwiska amerykańskiego ekonomisty Richarda Musgrave'a, który jako pierwszy sformułował je w odniesieniu do polityki finansów publicznych (zob. Musgrave 1959, Musgrave i Musgrave 1989):

a) **funkcja alokacyjna** – polityka gospodarcza ma na celu kształtowanie wielkości i struktury alokacji zasobów produkcyjnych między różne rodzaje produkcji w taki sposób, aby dostarczać dobra publiczne i dobra społecznie pożądane w wymaganej ilości oraz jakości, a także spełniać kryterium efektywności ekonomicznej poprzez maksymalizację wartości produkcji przy danych zasobach (wymiar mikroekonomiczny);
b) **funkcja stabilizacyjna** – polityka gospodarcza ma na celu łagodzenie cyklicznych wahań produkcji i zatrudnienia oraz łagodzenie skutków jednorazowych szoków popytowych lub podażowych (wymiar makroekonomiczny);
c) **funkcja redystrybucyjna** – polityka gospodarcza ma na celu zmianę pierwotnego podziału dochodów za pomocą regulacji podatków i transferów, tak aby spełnić kryterium sprawiedliwości społecznej (wymiar społeczny – pozaekonomiczny).

Dwie pierwsze funkcje dotyczą sfery gospodarki i służą zapewnieniu efektywności jej funkcjonowania. Funkcja alokacyjna zmierza zasadniczo do zapewnienia społecznego optimum na poszczególnych rynkach produktów i zasobów, co w dłuższym okresie pozwala maksymalizować tempo wzrostu gospodarczego. Funkcja stabilizacyjna minimalizuje odchylenia od stanu równowagi przy danych zasobach. Chodzi tu głównie o minimalizację odchyleń bieżącej produkcji od produktu potencjalnego, rzeczywistego bezrobocia od bezrobocia naturalnego i bieżącej inflacji od celu inflacyjnego, zapewniającego stabilność cen. Funkcja redystrybucyjna ma z kolei inny zasięg oddziaływania niż dwie poprzednie funkcje, ponieważ dotyczy podziału wytworzonego dochodu między członków danego społeczeństwa. Oznacza to, że jej działanie wychodzi poza zakres polityki gospodarczej i wkracza w domenę wyboru politycznego oraz polityki społecznej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

1.3.2. Cele

Każda świadoma działalność wymaga określenia celów, które chce się osiągnąć. Polityce gospodarczej wyznacza się różne cele, wśród których najczęściej wymienia się pełne zatrudnienie, stabilne ceny, sprawiedliwy rozkład dochodów, zrównoważony wzrost gospodarczy, równowaga bilansu płatniczego, równowaga budżetowa, konkurencyjność. Niekiedy są one bardziej szczegółowe i odnoszą się do wybranych sektorów (na przykład rozwój eksportu lub osiągnięcie samowystarczalności energetycznej), w innych przypadkach dotyczą przemian społecznych, takich jak ograniczenie ubóstwa czy wzrost zatrudnienia wśród kobiet. Mnogość celów wymaga ustalenia ich hierarchii, a także rozpoznania występujących miedzy nimi zależności. Nie wszystkie pożądane cele da się realizować jednocześnie, lecz zarazem w wielu przypadkach można mieć do czynienia z wymiennością (zastępowalnością) celów.

Cele polityki gospodarczej są na ogół publikowane w oficjalnych dokumentach. Jest to ważne pod względem zapewnienia stabilnych i przewidywalnych warunków gospodarowania, ponieważ uczestnicy rynku powinni wiedzieć, do czego dążą władze publiczne, aby móc stosownie do ich założeń planować własne działania. Ogłaszanie celów polityki jest rodzajem zobowiązania ze strony władzy państwowej do realizowania określonego kierunku działań i w ten sposób zmniejsza niepewność wobec przyszłości. Cele polityki gospodarczej realizowane w Unii Europejskiej oraz w Polsce przedstawione są w ramce 1.6.

Ramka 1.6. Cele polityki: art. 3 TUE i polski Program Konwergencji

Twórcy Unii Europejskiej zatroszczyli się o to, aby jej cele zostały przedstawione w dokumentach założycielskich. Określa je art. 3 Traktatu o Unii Europejskiej:

1) Celem Unii jest wspieranie pokoju, jej wartości i dobrobytu jej narodów.
2) Unia zapewnia swoim obywatelom przestrzeń wolności, bezpieczeństwa i sprawiedliwości bez granic wewnętrznych, w której zagwarantowana jest swoboda przepływu osób, w powiązaniu z właściwymi środkami w odniesieniu do kontroli granic zewnętrznych, azylu, imigracji, jak również zapobiegania i zwalczania przestępczości.
3) Unia ustanawia rynek wewnętrzny. Działa na rzecz trwałego rozwoju Europy, którego podstawą jest zrównoważony wzrost gospodarczy oraz stabilność cen, społeczna

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gospodarka rynkowa o wysokiej konkurencyjności zmierzająca do pełnego zatrudnienia i postępu społecznego oraz wysoki poziom ochrony i poprawy jakości środowiska naturalnego. Wspiera postęp naukowo-techniczny. Zwalcza wyłączenie społeczne i dyskryminację oraz wspiera sprawiedliwość i ochronę socjalną, równość kobiet i mężczyzn, solidarność między pokoleniami i ochronę praw dziecka. Wspiera spójność gospodarczą, społeczną i terytorialną oraz solidarność między Państwami Członkowskimi. Szanuje swoją bogatą różnorodność kulturową i językową oraz czuwa nad ochroną i rozwojem dziedzictwa kulturowego Europy.

4) Unia ustanawia unię gospodarczą i walutową, której walutą jest euro.
5) W stosunkach zewnętrznych Unia umacnia i propaguje swoje wartości i interesy oraz wnosi wkład w ochronę swoich obywateli. Przyczynia się do pokoju, bezpieczeństwa, trwałego rozwoju Ziemi, do solidarności i wzajemnego szacunku między narodami, do swobodnego i uczciwego handlu, do wyeliminowania ubóstwa oraz do ochrony praw człowieka, w szczególności praw dziecka, a także do ścisłego przestrzegania i rozwoju prawa międzynarodowego, w szczególności zasad Karty Narodów Zjednoczonych.
6) Unia dąży do osiągnięcia swoich celów właściwymi środkami odpowiednio do zakresu kompetencji przyznanych jej na mocy Traktatów.

Ogólny, opisowy charakter użytych sformułowań wskazuje na to, że są to cele polityczne. Określają one raczej kierunki działania Unii Europejskiej, nie zaś jakiś stan docelowy, a tym bardziej wyrażony w kategoriach ilościowych. Tego typu cele stanowią ogólną wytyczną polityki gospodarczej.

Zupełnie inny charakter mają cele wyznaczające kierunki polityki gospodarczej w poszczególnych krajach i w określonym czasie. Przykładowo, w odróżnieniu od ogólnej, politycznej formuły założeń opisanych w artykle 3 TUE, cele polityki gospodarczej w Polsce, zawarte w dokumentach rządowych, pozostają bardziej konkretne, dotycząc przede wszystkim obszaru gospodarki, i obrazują pewien stan, jaki rząd chce osiągnąć w perspektywie średniookresowej[12]. W polskim *Programie Konwergencji. Aktualizacja 2012* można przeczytać, miedzy innymi, co następuje:

Cele polityki gospodarczej

Głównym celem polityki gospodarczej rządu w średnim okresie pozostaje stworzenie warunków do szybkiego i zrównoważonego wzrostu przy zapewnieniu optymalnego tempa konsolidacji finansów publicznych (…).

[12] *Program Konwergencji. Aktualizacja 2012*, Rada Ministrów, Warszawa, kwiecień 2012, s. 6–7.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

1) **Konsolidacja fiskalna**

Celem rządu jest zlikwidowanie nadmiernego deficytu sektora instytucji rządowych i samorządowych w 2012 r., a następnie kontynuowanie konsolidacji fiskalnej i osiągnięcie średniookresowego celu budżetowego w 2015 r. (…). W długim okresie celem rządu pozostaje zapewnienie stabilności finansów publicznych. Realizacja tych zadań wymaga m.in. dalszego wzmocnienia ram instytucjonalnych prowadzenia polityki fiskalnej. Kluczową zmianą w stosunku do obecnych ram fiskalnych będzie położenie większego nacisku na zapewnienie trwałej równowagi sektora instytucji rządowych i samorządowych poprzez wprowadzenie stabilizującej reguły wydatkowej obejmującej cały sektor (…).

2) **Wzrost aktywności zawodowej**

Zmiany w strukturze demograficznej ludności i dążenie do zachowania równowagi systemu ubezpieczeń społecznych wymagają reformy systemu emerytalnego. Od 2013 r. wiek przechodzenia na emeryturę kobiet i mężczyzn będzie stopniowo podwyższany i zrównany na poziomie 67 lat. Proces ten zakończy się w 2040 r. Zreformowany zostanie także system zabezpieczenia emerytalnego dla służb mundurowych poprzez wydłużenie okresu pracy niezbędnego do nabycia uprawnień emerytalnych oraz podwyższenie wieku uprawniającego do świadczenia emerytalnego. Ponadto nabywanie uprawnień do emerytur górniczych zostanie ograniczone do osób pracujących bezpośrednio przy wydobyciu (…). Podniesieniu poziomu zatrudnienia służyć będzie także m.in. zwiększenie opłacalności podejmowania pracy poprzez racjonalizację systemu zasiłków i świadczeń społecznych. Dla osób w wieku 60+ zostanie przygotowany program działań prozatrudnieniowych.

3) **Poprawa warunków prowadzenia działalności gospodarczej i zwiększenie poziomu innowacyjności gospodarki**

Działania na rzecz dalszego usuwania barier rozwoju przedsiębiorczości będą przede wszystkim obejmowały zmniejszenie obciążeń administracyjnych i obowiązków informacyjnych oraz usprawnienie procedury rozpoczynania działalności gospodarczej. W celu poprawy warunków prowadzenia działalności gospodarczej, w br. zostanie opracowany projekt założeń do projektu tzw. trzeciej ustawy deregulacyjnej. Inne działania obejmują m.in. reformę prawa budowlanego, która doprowadzi do usprawnienia przebiegu procesu inwestycyjnego, a także zmiany w pracy sądów, mające na celu usprawnienie i znaczne skrócenie czasu postępowań sądowych (...). Powyższe działania będą sprzyjać tworzeniu środowiska przyjaznego w podejmowaniu działalności innowacyjnej przez podmioty gospodarcze. Głównym wyzwaniem związanym z podniesieniem poziomu innowacyjności jest jednak zwiększenie ogólnych nakładów na działalność badawczo-rozwojową, a w szczególności nakładów prywatnych. Realizacji tego celu, poza wspomnianym już zmniejszaniem obciążeń administracyjnych, służyć będą m.in. zwiększanie jakości regulacji gospodarczych, ich upraszczanie i dostosowywanie do potrzeb nowoczesnej gospodarki, rozwój kompetencji kadr dla przedsiębiorstw innowacyjnych, wspieranie sektora nauki i jego współpracy z biznesem oraz koordynacja polityki innowacyjnej i proeksportowej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

4) **Prywatyzacja i poprawa zarządzania aktywami Skarbu Państwa**

Działania w tym obszarze mają na celu usprawnienie i podniesienie efektywności procesów prywatyzacyjnych, zakończenie przekształceń własnościowych w części branż i sektorów oraz zwiększenie efektywności nadzoru właścicielskiego Skarbu Państwa. Rola państwa zostanie ograniczona w tych obszarach gospodarki, w których sprawowanie nadzoru właścicielskiego przez Skarb Państwa nie jest konieczne. Obecność Skarbu Państwa nie jest natomiast wykluczona w spółkach o istotnym znaczeniu strategicznym dla kraju. Podejmowane wysiłki będą prowadzić do przyspieszenia rozwoju i modernizacji przedsiębiorstw oraz przyczynią się do wzmocnienia konkurencyjności gospodarki polskiej.

Tak zarysowane cele, choć nadal ujęte w formie jakościowej, a nie ilościowej, stanowią już bardzo konkretny zapis zamierzeń rządu realizowanych w zakresie polityki gospodarczej. Ukazują nie tylko kierunki planowanych działań, lecz także elementy stanu docelowego, jaki rząd chce osiągnąć.

Z kolei kilka lat później w *Programie konwergencji. Aktualizacja 2015* rząd sformułował cele polityki gospodarczej w sposób bardziej ogólny[13]:

Przyjęta przez Radę Ministrów we wrześniu 2012 r. Średniookresowa Strategia Rozwoju Kraju 2020[14] (dalej jako Strategia) formułuje główny cel rządu w średnim horyzoncie czasowym jako **wzmocnienie i wykorzystanie gospodarczych, społecznych i instytucjonalnych potencjałów zapewniających szybszy i zrównoważony rozwój kraju oraz poprawę jakości życia ludności**. Strategia wskazuje trzy obszary strategiczne, na których koncentrować się powinny główne działania w celu przyspieszenia procesów rozwojowych: **Sprawne i efektywne państwo, Konkurencyjna gospodarka oraz Spójność społeczna i terytorialna**, stanowiące również bazę dla przygotowania kolejnych aktualizacji KPR, w tym KPR 2015/2016.

Podkreśla się przy tym złożoność wyzwania, jakie stanowi konieczność równoważenia finansów publicznych i zwiększania oszczędności przy jednoczesnym wdrażaniu wizji rozwoju opartego na likwidowaniu największych barier rozwojowych i w coraz większym stopniu opartego na edukacji, cyfryzacji i innowacyjności. Przyjęte w Strategii założenia dotyczące konsolidacji finansów publicznych (prowadzonej zgodnie z kolejnymi edycjami Programu) oznaczają, że warunek niezbędny dla wdrożenia wspomnianej wyżej wizji rozwoju stanowi racjonalizacja struktury wydatków i ich reorientacja na działania kluczowe dla rozwoju społeczno-gospodarczego i przestrzennego kraju.

---

[13] *Program Konwergencji. Aktualizacja 2015*, Rada Ministrów, Warszawa, kwiecień 2015, s. 11.

[14] *Strategia Rozwoju Kraju 2020 – Aktywne społeczeństwo, konkurencyjna gospodarka, sprawne państwo*, przyjęta przez Radę Ministrów 25 września 2012 r., dostępna na: http://www.mrr.gov.pl/rozwoj_regionalny/Polityka_rozwoju/SRK_2020/Documents/SRK_2020_112012_1.pdf (20.02.2015).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Zrównoważone finanse publiczne są niezbędnym warunkiem dla stabilności makroekonomicznej i, w konsekwencji, długofalowego wzrostu gospodarczego. Wzmocnieniu stabilności makroekonomicznej w Polsce sprzyjać też będą m.in. utworzenie organu nadzoru makroostrożnościowego oraz stworzenie ram prawnych dla przymusowej restrukturyzacji banków. (…).

Cele mogą być przedstawiane i zapisywane w rozmaity sposób. Raz nadaje się im bardziej jakościowy, opisowy charakter, a innym razem wyrażane są w postaci konkretnych działań lub liczb. Przykładem celu sformułowanego w sposób opisowy jest zapowiedź walki z bezrobociem jako głównej wytycznej działań rządu bądź też zapowiedź zwiększenia nakładów na edukację. Cel ilościowy stanowi zapowiedź obniżenia liczby bezrobotnych o wskazany odsetek (także stopy bezrobocia do określonego poziomu) lub sygnalizowanie osiągnięcia określonego udziału osób z wykształceniem wyższym w danym społeczeństwie. Ogólnie rzecz biorąc, cele polityki gospodarczej mogą przyjmować postać zmiennych o ściśle ustalonych wartościach, zmiennych bez ustalania określonej wartości bądź pewnej funkcji agregującej różne zmienne wyrażające cele cząstkowe.

Ustalając cele produkcyjne w postaci konkretnych wielkości ilościowych, rząd musi być przekonany, że są one możliwe do zrealizowania w danych warunkach, w danym czasie i w ramach znajdujących się w dyspozycji zasobów produkcyjnych. Nie da się sensownie zaplanować konkretnych wielkości celów możliwych do zrealizowania, jeśli rząd nie wie, jakimi zdolnościami wytwórczymi w danym okresie dysponuje gospodarka. Gdy rząd takiej wiedzy nie posiada albo wiedza ta jest niepełna, musi ograniczyć się do bardziej opisowego formułowania przynajmniej niektórych celów na przykład ustalając konkretnie cele priorytetowe, a pozostałe określając w sposób opisowy lub kierunkowy.

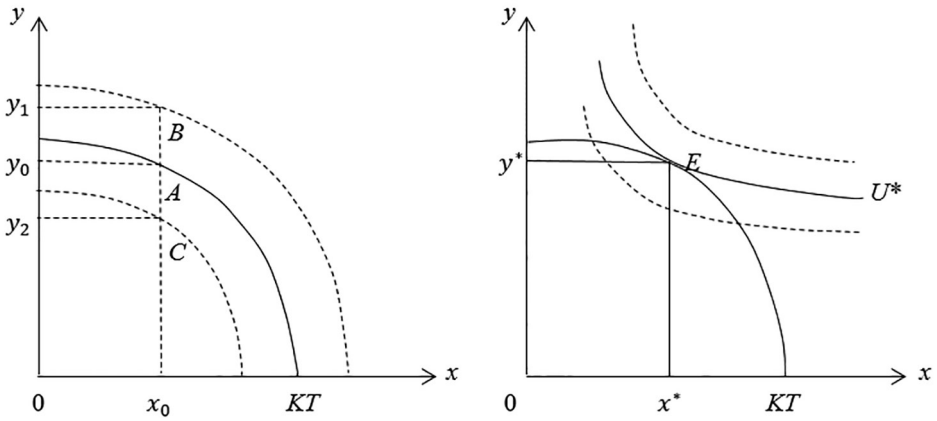
W lewej części rysunku 1.8 maksymalne możliwości produkcyjne w odniesieniu do dwóch produktów x i y opisuje krzywa transformacji KT (zwana też krzywą możliwości produkcyjnych). Zakłada się, że rząd dąży do zapewnienia określonej podaży dóbr na poziomie x_0 i y_0, i przyjmuje chwilowo, że oba dobra są równorzędne. Wówczas wyznaczenie celów zależy od kształtu i położenia krzywej transformacji.

Jeśli rząd zna położenie i kształt krzywej transformacji, a zatem wie, jakimi zasobami dysponuje i przy pomocy jakich technologii może przetwarzać zasoby w produkty, ma możliwość ustalenia zakładanych celów w postaci produkcji na określonym

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

poziomie x_0 i y_0, tak, jak opisuje to punkt *A* w lewej części rysunku 1.8. Jeśli rząd nie posiada pełnej wiedzy na temat położenia *KT*, a w szczególności nie wie, czy leży ona powyżej czy poniżej punktu *A*, może uznać jedno dobro za priorytetowe i ustalić pożądany zakres jego produkcji na poziomie x_0. Poziom produkcji drugiego dobra *y* będzie wówczas wynikowy, rezydualny, i będzie zależał od tego, gdzie w rzeczywistości znajdzie się krzywa *KT*. Jeśli krzywa *KT* leży powyżej punktu *A*, produkcja drugiego dobra wynosi y_1, co obrazuje punkt *B*. Z kolei jeśli krzywa *KT* leży poniżej punktu *A*, produkcja drugiego dobra osiąga wartość y_2, na co wskazuje punkt *C*.

Rysunek 1.8. Wyznaczanie celów o stałej wartości (lewa strona) i celów o zmiennej wartości (prawa strona)



Źródło: opracowanie własne.

Inną metodą wyznaczania celów jest skonstruowanie określonej społecznej funkcji użyteczności $-U = f(x, y)$, której wartość zależy od wielkości jednoczesnej konsumpcji jednego i drugiego dobra, a następnie znalezienie maksimum tej funkcji przy danych ograniczeniach dotyczących możliwości produkcyjnych (w ramach określonych zasobów i technologii). Wartości x^* i y^*, dla których $U^* = max$, stają się wówczas wyznaczonymi celami produkcji. Ilustruje to prawa strona rysunku 1.8, gdzie punkt styczności *E* wklęsłej krzywej transformacji *KT* i najdalej położonej wypukłej krzywej użyteczności U^* wyznacza pożądane i możliwe do zrealizowania cele x^* i y^*. Krzywe *U* leżące poniżej krzywej U^* zapewniają poziom użyteczności niższy od U^*, a krzywe *U* leżące powyżej krzywej U^* znajdują się poza obszarem dopuszczalnych rozwiązań.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

1.3.3. Instrumenty

Instrumentami polityki gospodarczej są określone zmienne (ekonomiczne, techniczne, prawne), które rząd może kontrolować i kształtować w taki sposób, by realizować założone cele. Aby wskazana zmienna mogła pełnić funkcję instrumentu, niezbędnym wymogiem jest jej skuteczność, co oznacza że dany instrument wywiera wpływ na cele. Im silniejsze jest oddziaływanie w przypadku danej zmiany wartości instrumentu, tym skuteczniejszy jest instrument. Ponieważ na ogół istnieje wiele instrumentów, które wpływają na kształtowanie się danej zmiennej będącej celem polityki, rząd powinien dążyć do wybierania najskuteczniejszych z nich oraz takich, nad którymi ma największą kontrolę. Jeśli instrumenty różnią się pod względem siły i kierunku wpływu wywieranego na zmienną celu oraz pod względem możliwości ich kontrolowania, uznaje się, że są one niezależne. Instrumenty zależne, czyli takie, które mają taki sam lub zbliżony wpływ na dany cel, nie są z punktu widzenia polityki gospodarczej odmiennymi narzędziami. Przykładowo, podniesienie stopy procentowej przez bank centralny i ograniczenie skali zasilających operacji rynku otwartego w podobny sposób oddziałują na wielkość podaży pieniądza, a więc są to instrumenty wzajemnie zależne. Z kolei wykorzystanie polityki fiskalnej i polityki pieniężnej w zakresie zwalczania bezrobocia wywołuje różne skutki[15], co pozwala uznać politykę fiskalną i politykę pieniężną za instrumenty niezależne.

Instrumenty polityki gospodarczej można klasyfikować w różny sposób. Holenderski ekonomista Jan Tinbergen (1956) wprowadził podział na politykę ilościową (zmiana wartości istniejących instrumentów), politykę jakościową (wprowadzanie nowych instrumentów) i politykę reform (zmiany systemowe). Tradycyjnie dzieli się też wskazane instrumenty na środki bezpośredniego oddziaływania i środki pośredniego oddziaływania. Do tych pierwszych zalicza się nakazy i zakazy (np. ceny maksymalne lub minimalne, limity emisji CO_2, kontyngenty ilościowe w handlu zagranicznym), podczas gdy druga z tych kategorii obejmuje zmienne, które wpływają na decyzje uczestników rynku poprzez bodźce skłaniające ich do przejawiania określonych zachowań (na przykład poprzez wprowadzanie podatków czy przyznawanie subsydiów).

[15] W szczególności wzrost wydatków rządowych w celu poprawy sytuacji na rynku pracy może nie przynieść oczekiwanego rezultatu, jeśli skutkiem tego wzrostu będzie wzrost stóp procentowych i w konsekwencji spadek wydatków prywatnych, w tym zwłaszcza inwestycji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Najczęściej stosowana w praktyce klasyfikacja instrumentów polityki gospodarczej polega na wyodrębnieniu obszaru oddziaływania danego instrumentu. Z tego punktu widzenia instrumenty dzieli się na makroekonomiczne i mikroekonomiczne. Wśród tych pierwszych wyróżnia się politykę fiskalną, rozumianą jako ustalanie poziomu wydatków i dochodów publicznych, politykę pieniężną (czasem wyodrębnia się z niej politykę kursową), polegającą na regulowaniu ilości pieniądza w obiegu, oraz – od niedawna – politykę makroostrożnościową, której celem jest ograniczenie ryzyka systemowego w sektorze finansowym. Niekiedy uzupełnia się ten podział o politykę kształtowania dochodów, politykę cen, zewnętrzną politykę ekonomiczną, obejmującą politykę handlową (w tym celną) i politykę wobec inwestycji zagranicznych, oraz politykę przyspieszania wzrostu gospodarczego. Z kolei do instrumentów mikroekonomicznych polityki gospodarczej zalicza się podatki, subsydia i subwencje, regulacje dotyczące funkcjonowania określonych rynków, konkurencji, praw własności i umów gospodarczych, pomocy publicznej, bankructwa przedsiębiorstw, a także wiele innych przepisów regulujących bieżące życie gospodarcze w danym kraju.

1.3.4. Związki między celami a instrumentami. Modele gospodarcze

Jak wspomniano, kluczową cechą instrumentu polityki jest skuteczność. Oznacza to, że między celem a instrumentem musi istnieć określone powiązanie, tak aby zmiana instrumentu powodowała pożądaną zmianę celu. Tego typu zależności pomiędzy instrumentami a celami często przedstawia się w postaci modeli ekonomicznych. Gdy dostatecznie wiernie odzwierciedlają one rzeczywiste relacje panujące w gospodarce, stają się wygodnym narzędziem polityki gospodarczej, ponieważ ukazują siłę i kierunek związku zachodzącego między instrumentami a celami, co z kolei pozwala precyzyjnie sterować procesami ekonomicznymi. Trzeba jednak pamiętać, że rzeczywistość gospodarcza jest nieskończenie złożona. Dlatego każdy model – niezależnie od swoich rozmiarów i stopnia szczegółowości – pozostaje zawsze uproszczonym obrazem zależności w gospodarce. Nie musi to jednak stanowić wady modelu, a wręcz przeciwnie – może być jego zaletą, jeśli uzna się, że celem modelowania jest wychwycenie najistotniejszych zależności, a zatem „oczyszczenie" badanej rzeczywistości ze związków nieistotnych lub przypadkowych[16].

[16] Warto przypomnieć, że interpretację taką zaproponował w latach 70. profesor Paweł Sulmicki.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W pierwszym etapie modelowania zjawisk na potrzeby polityki gospodarczej buduje się model ekonometryczny, opisujący zależności istniejące między różnymi zmiennymi i wyrażający je przy pomocy określonych funkcji. Modele takie mogą przyjmować postać tzw. modeli strukturalnych, wyjaśniających zmiany wartości danej zmiennej objaśnianej (zmiennej zależnej, zmiennej endogenicznej), będącej celem polityki gospodarczej, za pomocą zmian ograniczonej liczby zmiennych objaśniających (niezależnych, egzogenicznych). Chodzi więc o wyrażenie pewnej wybranej zmiennej y jako funkcji zbioru zmiennych x. Ponieważ w praktyce często zdarza się, że wśród tych ostatnich mogą występować także zmienne będące jednocześnie celem polityki gospodarczej, nierzadko zachodzi potrzeba przekształcenia modelu w taki sposób, aby zmiennymi objaśniającymi pozostawały wyłącznie zmienne niebędące celami. Taką postać modelu nazywa się postacią zredukowaną.

Różnicę zachodzącą między modelem w postaci strukturalnej a modelem w postaci zredukowanej można przedstawić na prostym przykładzie. Początkowo wychodzi się z założenia, że celem polityki gospodarczej jest wytwarzanie dwóch produktów y_1 i y_2, które są jednocześnie zużywane w produkcji jako surowiec. Oprócz tego w produkcji zużywa się (przenośne i podzielne) zasoby pracy Z i kapitału K. Model strukturalny produkcji będzie wyglądał następująco:

$$\left.\begin{aligned} y_1 &= f(Z, K, y_1, y_2) \\ y_2 &= g(Z, K, y_1, y_2). \end{aligned}\right\} \tag{1.6}$$

Jak widać, zmienne y występują w równaniach produkcji zarówno jako zmienne objaśniane, jak i zmienne objaśniające. Przekształcenie modelu do postaci zredukowanej polega na tym, że zmienne celu wyraża się wyłącznie jako zmienne objaśniane, czyli jako funkcje zasobów pracy Z i kapitału K:

$$\left.\begin{aligned} y_1 &= m(Z, K) \\ y_2 &= n(Z, K). \end{aligned}\right\} \tag{1.7}$$

Zaletą postaci zredukowanej pozostaje to, że cele wyrażone są wyłącznie jako funkcja zmiennych niezależnych. Jeśli rząd dąży do realizacji zmiennych celów na poziomie odpowiednio $\hat{y}_1$ i $\overline{y}_2$, to znając konkretną postać równań modelu zredukowanego, może określić wymagane wartości nakładów Z i K, potrzebne do osiągnięcia założonych celów:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$\left.\begin{aligned} Z &= \varphi(\overline{y}_1, \overline{y}_2) \\ K &= \mu(\overline{y}_1, \overline{y}_2). \end{aligned}\right\} \tag{1.8}$$

Taka postać modelu nosi nazwę postaci odwrotnej modelu zredukowanego – równania (1.8) określają niezbędną wielkość czynników produkcji, konieczną do osiągnięcia celów o ustalonej wysokości.

Różnicę między postacią strukturalną modelu a jego postacią zredukowaną można wyjaśnić na przykładzie prostego modelu makroekonomicznego, zwanego modelem IS-LM. Jest to jeden z najbardziej popularnych modeli wykorzystywanych do analiz makroekonomicznych. Jego szczegółowe omówienie zawiera następny rozdział. W tym miejscu wykorzystana zostanie tylko część modelu, ograniczona do równań opisujących równowagę na rynku produktów. W związku z tym przyjęto następujące oznaczenia: Y – dochód, C – konsumpcja, I – inwestycje, G – wydatki publiczne, r – stopa procentowa, natomiast a, b, c i g są parametrami modelu. Wówczas, w warunkach gospodarki zamkniętej, model ogranicza się do trzech równań:

$$\left.\begin{aligned} Y &= C + I + G \\ C &= a + cY \\ I &= b - gr. \end{aligned}\right\} \tag{1.9}$$

W modelu tym za cele polityki uznaje się poziom dochodu (produkcji) Y i poziom inwestycji I, a za jej instrumenty– poziom wydatków rządowych G i poziom stopy procentowej r. W postaci strukturalnej modelu zmienne celu występują w dwojakiej roli – zmiennych objaśniających i zmiennych objaśnianych: przykładowo zmienna celu Y występuje w pierwszym równaniu jako zmienna objaśniana, podczas gdy w drugim równaniu jako zmienna objaśniająca. Co więcej, w modelu występują dwie zmienne celu i dwa instrumenty, ale aż trzy równania – co oznacza, że z punktu widzenia poszukiwanych rozwiązań można zredukować układ (1.9) do dwóch równań. Podstawiając równanie drugie do pierwszego i przekształcając równania tak, aby Y i I znalazły się po lewej stronie, otrzymuje się postać zredukowaną modelu:

$$\left.\begin{aligned} Y &= \left[(b - gr) + G + a\right]/(1 - c) \\ I &= b - gr. \end{aligned}\right\} \tag{1.10}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Należy zauważyć, że w postaci zredukowanej cele polityki zależą wyłącznie od instrumentów (G, r) i od stałych parametrów (a, b, c, g). Model zredukowany można rozwiązać, przekształcając go do tzw. postaci odwrotnej, to znaczy przedstawiając instrumenty G i r jako funkcje celów Y oraz I, a także przyjmując następnie określone pożądane, docelowe wartości $Y = \bar{Y}$ oraz $I = \bar{I}$:

$$\left.\begin{aligned} G &= \bar{Y}(1-c) - a - \bar{I} \\ r &= \left(b - \bar{I}\right)/g. \end{aligned}\right\} \tag{1.11}$$

Rozwiązanie postaci odwrotnej modelu (1.11) pozwala znaleźć takie wartości instrumentów, które zapewnią osiągnięcie celów $\bar{Y}$ i $\bar{I}$.

Jan Tinbergen wykazał, że istnieje związek między liczbą celów a liczbą instrumentów polityki. Warunkiem skuteczności polityki jest spełnienie wymogu, aby liczba niezależnych instrumentów nie była mniejsza od liczby realizowanych celów. Innymi słowy, w przypadku m celów i n instrumentów warunkiem skuteczności polityki staje się to, aby $m \leq n$. Gdy $m > n$, część celów nie może być zrealizowana, co oznacza, że władza publiczna musi dokonywać wyboru celów, które chce zrealizować. Prawidłowość ta nazywana jest „złotą regułą" polityki gospodarczej (*golden rule*) lub po prostu regułą Tinbergena (Tinbergen 1952, Marczewski 2008)[17].

W sytuacji, gdy liczba instrumentów jest większa od liczby celów, władza publiczna może stosować więcej niż jeden instrument służący osiągnięciu określonego celu. Przy czym przy wyborze konkretnych instrumentów należy kierować się ich skutecznością, to znaczy wybierać takie narzędzia, które mają największy wpływ na zmienną celu. W praktyce można w takiej sytuacji oszacować regresję zmiennych celu względem wszystkich potencjalnych instrumentów, a następnie wykorzystywać w polityce te narzędzia, które są najsilniej skorelowane ze zmienną celu (Brainard 1967).

Jak wspomniano wyżej, wybór celów polityki niekoniecznie musi polegać na określeniu wartości liczbowych w odniesieniu do konkretnych zmiennych. Inną możliwością jest przyjęcie pewnej funkcji agregującej różne zmienne. Jej przykładem jest funkcja użyteczności społecznej (*social utility function*), zwana także funkcją społecznego wyboru lub funkcją społecznego dobrobytu (*social welfare*

[17] „Złota reguła" nie musi obowiązywać w modelu dynamicznym, w którym może zachodzić relacja $m > n$, ponieważ realizację większej liczby celów można rozłożyć na kolejne okresy.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

function), którą rząd pragnie maksymalizować przy danych ograniczeniach dotyczących zasobów produkcyjnych. W przypadku dwóch zmiennych celu, y_1 i y_2 problem optymalnej polityki gospodarczej można opisać następująco:

$$U = f\left(y_1, y_2\right) = max. \tag{1.12}$$

Przy ograniczeniach:

$$\left.\begin{aligned} y_1 &= m(Z, K) \\ y_2 &= n(Z, K) \end{aligned}\right\} \tag{1.13}$$

dla danych wartości Z i K.

Posługując się funkcją użyteczności, rząd nie zakłada z góry określonych wartości poszczególnych celów, a raczej maksymalizuje wartość funkcji użyteczności, będącej kombinacją różnych wartości celów, i na tej podstawie uzyskuje optymalne wartości zmiennych celu. W takim przypadku możliwe staje się uzyskanie rozwiązania nawet wtedy, gdy liczba instrumentów jest mniejsza od liczby celów, choć wartość funkcji użyteczności będzie odpowiednio niższa.

Kluczowym warunkiem praktycznej przydatności i skuteczności zastosowania modeli ekonomicznych pozostaje to, aby dany model stanowił dobre odzwierciedlenie rzeczywistych zależności. W celu dokonania oceny tego, czy i w jakim stopniu określony model właściwie przedstawia rzeczywistość, stosuje się wiele rozmaitych technik statystycznych. Nigdy nie ma jednak gwarancji, że dany model da prawidłową odpowiedź co do *przyszłego* zachowania się wybranej zmiennej objaśnianej. Wynika to stąd, że oszacowanie modelu opiera się zawsze na danych z przeszłości, co oznacza, po pierwsze, że nie wszystkie zmienne objaśniające, mające wpływ na zmienne celu, mogą w takiej sytuacji być brane pod uwagę, i, po drugie, że zależność między zmiennymi i celami w przyszłości będzie najprawdopodobniej taka sama jak w przeszłości. Wprawdzie obliczenia przeprowadzane na podstawie danych historycznych najczęściej podlegają weryfikacji ekspertów, lecz nie pozwala to na całkowite uniknięcie błędów. Ograniczeniem pozostaje także statyczny (z reguły) i deterministyczny (często) charakter przyjętej postaci modelu. Próbą przezwyciężenia tych barier było budowanie modeli dynamicznych i stochastycznych, co narzucało jednak konieczność przyjmowania kolejnych założeń, takich jak np. to o określonej ścieżce zmian w czasie i określonym rozkładzie losowym wartości badanych zmiennych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Konkretna analityczna postać modelu pozwala wprawdzie na badanie związku między instrumentami a celami, ale tylko przy założeniu, że charakter tego związku oraz parametry modelu są stałe. Tymczasem rzeczywiste zależności istniejące w gospodarce ulegają ciągłym przekształceniom – zwłaszcza uczestnicy rynku nie pozostają nigdy bierni wobec zmieniających się warunków działania i dostosowują do nich swoje postepowanie. Oznacza to, że polityka gospodarcza staje się w tym przypadku czymś w rodzaju podążania za ruchomym celem. Na to sprzężenie zwrotne między polityką a zachowaniami uczestników rynku jako pierwszy zwrócił uwagę w 1976 roku amerykański ekonomista Robert Lucas, a jego słynna krytyka polityki gospodarczej opartej na makromodelach ekonometrycznych nosi w ekonomii nazwę krytyki Lucasa (*Lucas critique*)[18]. Wieloletnie doświadczenia związane ze stosowaniem modeli strukturalnych do wyjaśniania zależności ekonomicznych potwierdziły, że charakter badanych uwarunkowań ulega zmianom (zarówno ewolucyjnym, jak i skokowym). Ponadto w wielu przypadkach bardzo trudno jest ustalić, które zmienne objaśniające mają w danym przedziale czasu największy wpływ na zmienne objaśniane, będące celami polityki.

Ekonomiści próbowali temu zaradzić, wykorzystując szybki rozwój technologii komputerowych do budowania coraz większych modeli z coraz większą liczbą zmiennych i coraz większą liczbą równań, ale rezultaty nie były zadawalające. Przekonano się, że modele składające się z kilkuset lub kilku tysięcy równań oraz zbliżonej liczby zmiennych wcale nie zapewniały bardziej precyzyjnego i dokładniejszego opisu nieskończenie złożonej rzeczywistości gospodarczej niż proste modele złożone z kilku równań. Dlatego też odchodzi się w badaniach od modeli deterministycznych na rzecz modeli stochastycznych, uwzględniających niepewność w postaci odpowiednio modelowanych składników losowych.

1.3.5. Reguły polityki

Cele i instrumenty często powiązane są ze sobą regułami polityki. Publicznie ogłaszane zasady postępowania mogą przyjąć postać dobrowolnie nałożonych ograniczeń na określone zmienne. W takiej sytuacji reguła polega na przestrzeganiu danego ograniczenia w bieżącej polityce, co oznacza konieczność odpowiedniego dostosowania instrumentów polityki gospodarczej. Przykładem tego typu norm jest obowiązujące w Unii Europejskiej ograniczenie wielkości deficytu finansów

[18] Szersze omówienie krytyki Lucasa znajduje się w rozdziale 5.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

publicznych do poziomu 3% PKB lub zapisane w artykule 216.5 Konstytucji RP ograniczenie wielkości długu publicznego do poziomu 3/5 PKB. W innych przypadkach reguła polityki określa funkcjonalny związek pomiędzy celem a instrumentem – można ją wówczas przyrównać do określonej funkcji reakcji władz publicznych. Funkcja reakcji opisuje pewien stały związek pomiędzy stopniem realizacji danego celu a zmianą danego instrumentu, wobec czego reguła polityki wprowadza pewien stopień automatyzmu do działań rządu. Istnienie takich reguł, w przeciwieństwie do ich braku, tj. uznaniowości (dyskrecjonalności), pozwala uczestnikom rynku lepiej przewidywać działania władz publicznych. Zmniejsza to niepewność odnośnie do przyszłego rozwoju sytuacji, obniża koszty związane z ryzykiem i pozwala lepiej planować działalność gospodarczą, a zwłaszcza inwestycyjną.

Przykładem popularnej reguły stosowanej w polityce pieniężnej jest ścisłe uzależnienie wysokości stóp procentowych banku centralnego od tempa wzrostu podaży pieniądza. Można założyć, że w opinii banku centralnego pomiędzy ilością pieniądza w obiegu a inflacją występuje stała relacja i że, gdy podaż pieniądza wzrasta szybciej niż, np. 6% rocznie, powstaje duże ryzyko, że ceny będą rosły w tempie szybszym niż potencjalnie bezpieczne, np. szybciej niż 2% rocznie. Jeśli celem polityki pieniężnej jest utrzymanie stabilności cen, bank centralny może ogłosić, że za każdym razem, gdy tempo wzrostu podaży pieniądza przekroczy 6%, stopy procentowe będą podnoszone. Bank centralny może nawet pójść o krok dalej, ogłaszając, że, przykładowo, będzie podnosił stopy procentowe o 0,25 punktu procentowego na każdy jeden punkt procentowy wzrostu podaży pieniądza powyżej 6%. Z kolei przykładem reguły zastosowanej w polityce fiskalnej jest uzależnienie wielkości deficytu budżetowego od poziomu długu publicznego. W Polsce przez wiele lat obowiązywała reguła mówiąca o tym, że gdy poziom długu publicznego przekroczy 50% PKB, rząd musi opracować kolejny budżet w taki sposób, aby relacja deficytu do dochodów nie uległa zmianie, gdy poziom długu publicznego przekroczy zaś 55% PKB, rząd będzie musiał przygotować budżet zrównoważony, czyli z deficytem równym zeru[19].

Celowość stosowania reguł w polityce gospodarczej była i jest szeroko dyskutowana. Z jednej strony dyskrecjonalność polityki zapewnia dużą swobodę działań i pozwala elastycznie oraz w sposób w danym momencie optymalny reagować na pojawiające się szoki. Taka elastyczność może być traktowana jako pożądana cecha polityki prowadzonej w świecie, w którym przyszłość jest zawsze niepewna

[19] Zob. ustawa o finansach publicznych z 2009 roku.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i nie można z góry przewidzieć, z jakimi wyzwaniami przyjdzie mieć do czynienia. Z drugiej strony dyskrecjonalność zwiększa niepewność co do przyszłości, ponieważ uczestnicy rynku nie wiedzą, czego można spodziewać się po polityce władzy gospodarczej (również sama władza często nie wie, w jaki sposób zareaguje na nieoczekiwany rozwój wydarzeń). Ta zwiększona niepewność przekłada się na wzrost premii za ryzyko w rynkowych stopach procentowych oraz w cenach kontraktów handlowych, co w rezultacie ogranicza skalę działalności gospodarczej i hamuje wzrost.

Postępowanie według z góry ustalonych i ogłoszonych publicznie reguł zwiększa natomiast przejrzystość oraz przewidywalność polityki rządu, co z kolei ułatwia planowanie działań i obniża prawdopodobieństwo podejmowania błędnych decyzji przez uczestników rynku. W związku z tym panuje ogólne przekonanie co do tego, że trzymanie się ustalonych reguł zwiększa wiarygodność oraz przewidywalność działań władzy publicznej i obniża koszty rozwoju poprzez spadek premii za ryzyko w stopach procentowych i w cenach kontraktów.

1.3.6. Krytyka Lucasa

Modele ekonomiczne (strukturalne i w postaci zredukowanej) opisują związek pomiędzy zmiennymi objaśniającymi a zmienną objaśnianą przy założeniu, że relacja ta ma stabilny charakter. Oznacza to, że zarówno postać funkcjonalna, jak i oszacowane parametry modelu nie ulegają zmianom. Innymi słowy przyjmuje się, że funkcje reakcji uczestników rynku nie zmieniają się w czasie, a więc ich zachowanie jest przewidywalne. Jak już wspomniano, Robert Lucas (1976) wykazał, że tego typu podejście nie uwzględnia sprzężenia zwrotnego między decyzjami polityki gospodarczej a zachowaniem uczestników rynku, którzy z reguły nie ignorują decyzji władzy, a raczej odpowiednio się do nich dostosowują. Przy czym działanie to polega nie tylko na prostym reagowaniu na zmiany zmiennych będących instrumentami polityki zgodnie z oszacowanym modelem ekonomicznym, lecz także na dostosowaniu wzorców zachowania na skutek zmiany oczekiwań, czyli przekształceniu zarówno samych zmiennych, jak i parametrów modelu, tj. kształtu funkcjonalnej zależności.

Najbardziej charakterystycznym przykładem polityki kształtującej zachowania uczestników rynku jest polityka reform strukturalnych, określana też jako wiarygodna zmiana systemowa (*credible regime change*). W wyniku jej oddziaływania dotychczasowe wzorce zachowań ulegają przekształceniom ze względu na zmianę

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

oczekiwań co do funkcjonowania systemu gospodarczego. Przykładem polityki spełniającej kryteria wiarygodnej zmiany systemowej, a zarazem trafnie opisywanej przez krytykę Lucasa, jest „przestrzelona" dewaluacja złotego w ramach programu terapii szokowej w Polsce z 1990 roku (zob. ramka 1.7).

Ramka 1.7. Polityka wiarygodnej zmiany systemowej (*credible regime change*): dewaluacja złotego w ramach polskiego programu terapii szokowej z 1990 roku jako przykład krytyki Lucasa

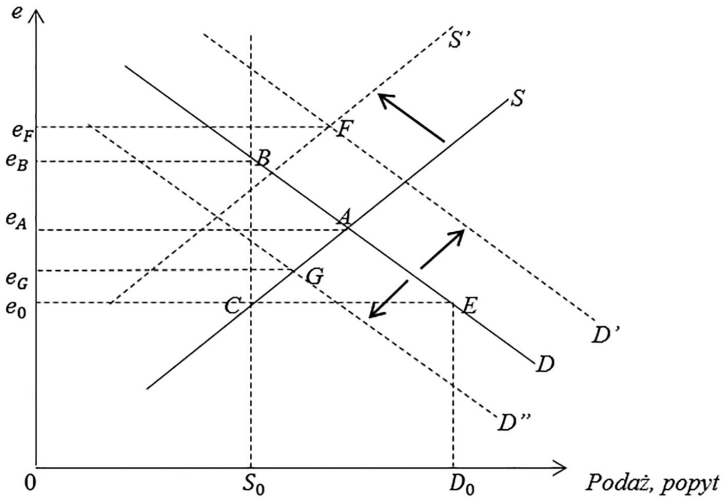
Ważnymi komponentami reform rynkowych wprowadzonych w Polsce w 1990 roku były: liberalizacja cen, liberalizacja handlu zagranicznego i wprowadzenie wymienialności waluty krajowej na waluty obce. Sytuacja w końcu lat 80. charakteryzowała się chronicznym brakiem walut wymienialnych, co było spowodowane m.in. sztucznie zawyżonym kursem złotego i dużym spekulacyjnym popytem ludności oraz przedsiębiorstw na waluty obce, wynikającym z kolei z „ucieczki" od złotego w warunkach galopującej inflacji i niedoborów na rynku towarów. Wprowadzenie i utrzymanie wymienialności oraz przywrócenie równowagi między podażą a popytem na rynku walutowym wymagało m.in. przeprowadzenia dewaluacji złotego. Przed ówczesnym polskim rządem stanęło więc pytanie dotyczące tego, o ile należy zdewaluować polska walutę, aby przywrócić trwałą równowagę na rynku walutowym w warunkach jednoczesnego uwolnienia cen krajowych i liberalizacji wymiany z zagranicą (Rosati 1994).

Problem ten udało się przedstawić na rysunku 1.9, gdzie na osi pionowej odłożono nominalny kurs walutowy (w PLN/USD), a na osi poziomej podaż i popyt z przełożeniem na waluty obce (zob. Rosati 1993, 1994). Krzywe podaży S i popytu D jako funkcje kursu mają kształt standardowy. Zakłada się, że w ramach systemu centralnie planowanego sztucznie zaniżony, oficjalny kurs dolara amerykańskiego wynosi e_0, co wyznacza wielkość podaży walut obcych na poziomie S_0. Niedobór między podażą a popytem przy danym kursie oficjalnym wynosi $(D_0 - S_0)$, co powoduje, że kurs czarnorynkowy – dla danej ograniczonej podaży – wynosi e_B. W tej sytuacji można przypuszczać, że dewaluacja e_0 do e_A dorowadzi do równowagi na rynku walutowym w warunkach wprowadzenia wymienialności liberalizacji handlu zagranicznego. Tak wyglądałaby równowaga „statyczna". Rząd polski zdecydował się jednak na głębszą dewaluację, ponieważ jednocześnie miało mieć miejsce uwolnienie cen krajowych, co budziło obawy przed tym, że wzrost cen

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

towarów musi również prowadzić do wzrostu kursu rynkowego. Ponieważ uznano, że wzrost cen w skali wyższej od wzrostu kursu spowoduje przesunięcie krzywej podaży w lewo (ze względu na wzrost kosztów) i krzywej popytu w prawo (ze względu na przesunięcie popytu z drożejących towarów), odpowiednio do *S'* i *D'*, to nowym punktem równowagi stanie się *F*. Ustalono więc, że kursem równowagi będzie e_F, wyższy nie tylko od e_A, lecz także od kursu czarnorynkowego e_B. Faktycznie zresztą, 31 grudnia 1989 roku zdecydowano się na dewaluację złotego do poziomu 9500 PLN/USD, znacznie powyżej kursu „statycznej" równowagi (szacowanego na około 6700 PLN/USD), a także powyżej kursu równowagi na czarnym rynku (około 7500 PLN/USD)[20].

Rysunek 1.9. Przestrzelenie dewaluacji przy wprowadzaniu wymienialności i uwolnieniu cen



Źródło: Rosati (1993), s. 262.

W rzeczywistości krzywa popytu nie przesunęła się jednak w górę, lecz w dół. Było to spowodowane częściowo spadkiem dochodów realnych w warunkach wzrostu cen i kontroli wzrostu płac, a częściowo także zmianą oczekiwań. Ten drugi czynnik odegrał kluczową rolę. Wprowadzenie w życie wiarygodnego programu

[20] Nie brakowało głosów przestrzegających przed nadmierną dewaluacją. Wskazywano na to, że kurs równowagi powinien mieścić się w przedziale miedzy kursem średnim a kursem czarnorynkowym, który wyznaczał kurs krańcowy powiększony o premię za ryzyko z tytułu nielegalności transakcji na czarnym rynku (Rosati 1989).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

stabilizacyjnego, którego celem miało być wyhamowanie inflacji, i obejmującego drastyczne podniesienie krajowych stóp procentowych, spowodowało obniżenie oczekiwań inflacyjnych, zwiększając jednocześnie popyt na pieniądz krajowy. W rezultacie popyt na waluty obce – w jego części spekulacyjnej – spadł, a nie wzrósł, podczas gdy krzywa popytu przesunęła się w dół do położenia *D*". Z kolei krzywa podaży *S* nie przesunęła się w lewo, ponieważ wzrost cen krajowych w wyniku uwolnienia był porównywalny ze skalą dewaluacji. W rezultacie kurs równowagi ukształtował się na poziomie e_G, odpowiadającym punktowi przecięcia krzywej *D*" i krzywej *S*' (punkt *G*). Ustalony w wyniku dewaluacji kurs e_F okazał się więc zdecydowanie za wysoki. Jest to przykład sytuacji, do której nawiązuje krytyka Lucasa – wiarygodna polityka zmiany systemowej powoduje zmianę zachowań uczestników rynku.

Można przyjąć, że popyt na waluty obce w warunkach transformacji jest funkcją realnego dochodu i niezabezpieczonego parytetu stóp procentowych:

$$D = a\left(\frac{Y}{P}\right) + b\left[\left(\frac{e^p - e_0}{e_0}\right) - i\right], \tag{1.14}$$

gdzie e^p jest oczekiwanym kursem złotego, a wyrażenie $\left(\frac{e^p - e_0}{e_0}\right)$ odpowiada stopie oczekiwanej dewaluacji złotego, na przykład w ciągu roku, z kolei *i* jest krajową stopą procentową (w warunkach wysokiej inflacji w kraju zakłada się, że zagraniczna stopa procentowa jest równa zeru). Można też przyjąć, że oczekiwana stopa dewaluacji jest funkcją różnicy między kursem czarnorynkowym a kursem oficjalnym:

$$\left(\frac{e^p - e_0}{e_0}\right) = c\left(\frac{e_B - e_0}{e_0}\right). \tag{1.15}$$

Podstawiając do równania (1.14) i różniczkując, otrzymuje się:

$$dD = a\left(\frac{dY}{P} - \frac{dP}{P}\frac{Y}{P}\right) + bc\left(\frac{de_B}{e_0} - \frac{de_0}{e_0}\frac{e_B}{e_0}\right) - b(di). \tag{1.16}$$

Pierwszy element po prawej stronie równania (1.16) mierzy efekt dochodowy (spadek dochodów realnych na skutek wzrostu cen), drugi zaś mierzy wpływ oczekiwań kursowych, a trzeci – wpływ wzrostu krajowych stóp procentowych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Następnie zakłada się, że nominalne dochody nie ulegają zmianie. Wówczas równanie (1.16) upraszcza się do postaci:

$$dD = -a\left(\frac{dP}{P}\frac{Y}{P}\right) - bc\left(\frac{de_B}{e_0} - \frac{de_0}{e_0}\frac{e_B}{e_0}\right) - b(di). \tag{1.17}$$

Łatwo zauważyć, że pierwszy i trzeci element prawej strony równania (1.17) są zawsze ujemne. Również drugi element będzie ujemny pod warunkiem, że:

$$\frac{de_B}{e_B} < \frac{de_0}{e_0}. \tag{1.18}$$

Warunek (1.18) jest spełniony, gdy stopa rzeczywistej dewaluacji jest większa od stopy zmiany kursu czarnorynkowego. W przypadku wiarygodnej polityki stabilizacyjnej kurs czarnorynkowy nie rośnie. W tej sytuacji $dD < 0$, popyt na waluty obce spada, a krzywa D przesuwa się w dół.

Dewaluacja złotego dokonana w końcu 1989 roku nie uwzględniała więc efektu spadku dochodów realnych, wzrostu stopy procentowej, a przede wszystkim zmiany oczekiwań uczestników rynku. Nadmierna dewaluacja przyczyniła się do dodatkowego wzrostu inflacji (poprzez wpływ na ceny, czyli tak zwany efekt *pass-through*) oraz do pogłębienia recesji w 1990 roku na skutek silnego wzrostu kosztów produkcji w odpowiedzi na wzrost cen towarów wymiennych.

Krytyka Lucasa stawia dość radykalne wyzwanie pod adresem polityki gospodarczej. Ponieważ, jak wspomniano, makroekonomiczne modele strukturalne są budowane i szacowane na podstawie obserwacji z przeszłości, ich wykorzystanie w celu kształtowania pożądanego zachowania podmiotów gospodarczych w przyszłości jest jednym z głównych sposobów prowadzenia polityki gospodarczej. Lucas twierdzi jednak, że samo zastosowanie określonej polityki wywołuje zmiany oczekiwań wśród uczestników rynku, powodujących z kolei przekształcenie dotychczasowych wzorców zachowań. Oznacza to, że zastosowana polityka gospodarcza nie będzie optymalna, a co więcej może okazać się nieskuteczna lub nawet przynieść efekt przeciwny do zakładanego.

Krytyka Lucasa pokazuje, że w polityce gospodarczej występuje problem niespójności czasowej (*time inconsistency, dynamic inconsistency*), tj. niedostosowania polityki do sytuacji w kolejnych okresach. Przykładowo można założyć, że rząd

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

po rozpoznaniu sytuacji gospodarczej i opisaniu jej za pomocą pewnego modelu strukturalnego decyduje się na prowadzenie określonej polityki, która przy danych wartościach parametrów postaci analitycznej modelu jest w danym okresie optymalna. Samo zastosowanie tej strategii wywołuje jednak zmianę bodźców i warunków funkcjonowania, co skłania racjonalnych uczestników rynku do zmiany oczekiwań wobec przyszłości, a w konsekwencji do dostosowania swojego postepowania do nowej sytuacji. Racjonalni uczestnicy rynku kierują się nie tylko bieżącą polityką, lecz także oczekiwaniami co do dalszego rozwoju sytuacji w następstwie stosowania określonej polityki. Wszystko to powoduje, że w następnym okresie uczestnicy rynku będą zachowywać się inaczej niż wynikałoby to z przyjętego modelu, wskutek czego prowadzona polityka przestaje być optymalna. Rząd może zatem nie osiągnąć założonych celów i będzie zmuszony dokonać na nowo diagnozy sytuacji oraz podjąć próbę zastosowania innej polityki. Oznacza to, że uczestnicy rynku wpływają tym razem na określone zmiany w zakresie polityki prowadzonej przez rząd – występuje tu więc klasyczne sprzężenie zwrotne. Lecz skutki kolejnych działań rządu będą podobne – uczestnicy rynku znowu zmienią swoje zachowania w reakcji na wprowadzenie danej polityki, a rząd ponownie może nie być w stania realizować założonych celów.

Krytyka Lucasa może mieć daleko idące konsekwencje. Chodzi nie tylko o niemożność osiągania celów wyznaczonych w ramach polityki, lecz także o to, że ciągłe zmiany dokonywane przez rząd powodują konieczność nieustannego dostosowywania się po stronie uczestników rynku, co zwiększa niepewność i pociąga za sobą dodatkowe koszty. Uwzględniając wszystkie te czynniki, Lucas był zdania, że polityka gospodarcza jest zasadniczo skazana na niepowodzenie i radził ograniczenie jej stosowania do minimum.

Tak radykalny wniosek nie musi być jednak zasadny Państwo, prowadząc swą politykę, może i powinno brać pod uwagę reaktywną naturę zachowań uczestników rynku i starać się je przewidywać. Statyczne modele strukturalne mogą być w takiej sytuacji mało przydatne. Większe możliwości dają modele dynamiczne, uwzględniające opóźnienia i przyspieszenia czasowe. Przykładem takich modeli są wektorowe modele autoregresyjne typu VAR i SVAR, które wychwytują – przynajmniej do pewnego stopnia – sprzężenia zwrotne pomiędzy celami i instrumentami (Sims 1980). Ponadto działania władzy publicznej powinny być zaplanowane i realizowane w taki sposób, aby nie rodzić oczekiwań skłaniających uczestników rynku do takiej zmiany zachowań, która uczyni działania władzy publicznej nieskutecznymi (Acocella 2002, s. 230–234).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

1.3.7. Instytucje

Instytucje definiuje się jako formalne i nieformalne reguły określające relacje między uczestnikami rynku i szerzej – pomiędzy ludźmi – oraz tworzące określoną strukturę bodźców kształtujących ich zachowania (North 1990, 1991). Przyjmują one postać struktur organizacyjnych (systemów) oraz regulacji prawnych o względnie trwałym charakterze, które skłaniają uczestników rynku do dostosowywania swoich zachowań do istniejących ram prawno-instytucjonalnych. W ten sposób instytucje tworzą określoną strukturę bodźców funkcjonujących w gospodarce i kształtujących postępowanie poszczególnych jednostek. Efektywność i skuteczność polityki gospodarczej zależy w dużej mierze od charakteru i jakości istniejących instytucji. Do najważniejszych i najbardziej typowych instytucji w gospodarce rynkowej należą rynki dóbr, usług i czynników wytwórczych, instytucje rządowe i samorządowe, systemy bankowe, płatnicze, system kursu walutowego, sądownictwo gospodarcze, związki zawodowe i związki pracodawców, a także procedury prawa handlowego, budżetowe i upadłościowe. Sprawne i budzące zaufanie instytucje są ważnym składnikiem tzw. kapitału społecznego, pozwalającego obniżać koszty prowadzenia działalności gospodarczej.

Instytucje różnią się od instrumentów polityki tym, że mają względnie trwały charakter, a więc mają służyć tworzeniu stabilnych ram w przypadku danej działalności. Co więcej, w przeciwieństwie do instrumentów, nie powinny być zmieniane zbyt często, ponieważ tracą wówczas swój wpływ na kształtowanie zachowań uczestników rynku. Częste zmiany instytucji powodują wstrzymanie procesów dostosowawczych i ograniczenie skali działalności ze względu na zwiększoną niepewność, a także skłaniają do podejmowania prób oddziaływania na proces tych przemian przez uczestników rynku. Niemniej, co jakiś czas albo w wyniku zasadniczych zmian w otoczeniu konieczna jest również zmiana instytucji. Działania te nazywa się często reformami strukturalnymi lub zmianami systemowymi (*regime change*), co ma podkreślać ich trwały charakter i odróżniać je od bieżącej polityki, prowadzonej w istniejących ramach instytucjonalnych.

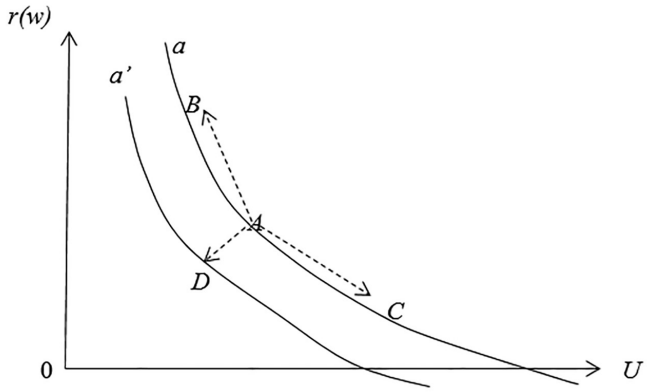
Rysunek 1.10 pozwala lepiej zrozumieć różnicę zachodzącą między wpływem instrumentów polityki a wpływem zmian instytucjonalnych na przykładzie sytuacji na rynku pracy. Ujemnie nachylona krzywa Phillipsa (*a*) prezentuje krótkookresową zależność między inflacją i bezrobociem. Początkowo zakłada się, że w punkcie wyjścia sytuację w gospodarce opisuje punkt A, w którym bezrobocie kształtuje się

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

na poziomie naturalnym. Zwiększając podaż pieniądza, władza publiczna ułatwia dostęp do kredytów, a wywołany tym wzrost produkcji oraz inwestycji powoduje szybszy wzrost płac *r*(*w*). Szybsze tempo wzrostu płac wpływa w krótkim okresie na spadek bezrobocia *U*, co odbywa się jednak kosztem wzrostu inflacji. Gospodarka przesuwa się z punktu *A* do *B*.

Rząd staje więc przed trudnym wyzwaniem: musi zdecydować, czy, dążąc do obniżenia bezrobocia, godzić się na wzrost inflacji, czy też, dążąc do ograniczania inflacji, godzić się na wzrost bezrobocia. W krótkim okresie dostępne instrumenty polityki pozwalają jedynie przesuwać się wzdłuż krzywej Phillipsa, na przykład z *A* do *B* lub z *A* do *C*. Sytuację tę można jednak *bezwzględnie* poprawić za pomocą wprowadzania reform strukturalnych (zmian instytucjonalnych) na rynku pracy. Przykładowo, zwiększenie przestrzennej i sektorowej mobilności pracowników poprzez podniesienie ich kwalifikacji umożliwia wzrost wydajności pracy, co z kolei pozwala zwiększać wysokość płac bez wywoływania inflacji. Podobne skutki przynosi usuwanie nadmiernie krępujących przepisów regulujących mechanizm zatrudniania i zwalniania pracowników. Alternatywnie można dzięki temu obniżać bezrobocie bez wywoływania zwiększonej inflacji. Krzywa Phillipsa przesuwa się w dół do położenia *a*', podczas gdy gospodarka przechodzi z punktu *A* do *D*, gdzie poziom zarówno bezrobocia, jak i inflacji jest niższy.

Rys 1.10. Ilustracja różnicy między wykorzystaniem instrumentów polityki w obecnych ramach instytucjonalnych a wykorzystaniem zmian instytucjonalnych (reformą strukturalną) – przesuwanie się wzdłuż krzywej Phillipsa a przesunięcie krzywej Phillipsa na zewnątrz



Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Powszechnie uważa się, że reformy strukturalne w krótkim okresie pociągają za sobą koszty społeczne, a korzyści pojawiają się dopiero po pewnym czasie. Międzynarodowy Fundusz Walutowy przeprowadził badania, z których wynika, że na przykład reformy rynku pracy oraz – w mniejszym stopniu – reformy rynku produktów wywołują w krótkim okresie negatywne skutki w odniesieniu do produkcji i zatrudnienia. Z kolei reformy dotyczące podatków i sektora finansowego wywierają od razu pozytywny wpływ na gospodarkę (IMF 2004). Obawa przed utratą społecznego poparcia w związku z kosztami reform jest często argumentem przemawiającym na rzecz ich odwlekania lub zaniechania. Na ogół jednak, jeśli reformy okazują się konieczne, ich odłożenie w czasie nie tylko nie zmniejsza kosztów wdrożenia, lecz pociąga za sobą także dodatkowe straty bieżącego dochodu, wynikające z pogłębiania się istniejących zniekształceń rynkowych.

Skalę dolegliwości społecznych związanych z reformami można zmniejszyć dzięki ich dobremu zaplanowaniu i przygotowaniu. Po pierwsze, wprowadzanie reform w całościowych pakietach jest najczęściej skuteczniejsze i mniej kosztowne dzięki osiąganym efektom synergii niż sekwencyjne wdrażanie reform krok po kroku. Po drugie, liczne badania dowodzą, że reformy strukturalne, takie jak na przykład reformy rynku pracy, dają lepsze efekty i niosą niższe koszty społeczne, gdy są wdrażane w okresach dobrej koniunktury, a nie w czasie recesji[21]. Po trzecie, ponieważ określony rozkład wydatków i korzyści w czasie pociąga za sobą zmiany w międzyokresowym i międzypokoleniowym rozkładzie dochodów w społeczeństwie, można łagodzić ciężar kosztów reform, jakie bardziej niż proporcjonalnie ponoszą określone grupy społeczne (lub pokolenia), poprzez zastosowanie mechanizmu transferu dochodów od grup korzystających z reform na rzecz grup ponoszących ich bezpośrednie koszty. Stworzenie skutecznego i sprawiedliwego systemu kompensacyjnego jest jednak zawsze wielkim wyzwaniem dla polityków gospodarczych.

1.3.8. Polityka gospodarcza jako sztuka wyboru w przestrzeni i w czasie (sekwencja wyborów)

Ponieważ zasoby są zawsze ograniczone, nie da się zrealizować wszystkich celów jednocześnie. Oznacza to, że trzeba dokonywać wyboru między różnymi celami lub między stopniem ich realizacji. Niekiedy taka selekcja jest dość oczywista. Na

[21] IMF, *World Economic Outlook*, Washington, April 2016, chapter 3.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przykład, w okresie recesji walka z bezrobociem jest z reguły ważniejsza niż walka z inflacją. W innych przypadkach można w procesie wyboru bazować na subiektywnych ocenach oraz intuicji. Przykładowo, budowa szkoły w miejscowości, w której nie ma żadnej placówki oświatowej, jest zapewne ważniejsza niż budowa basenu pływackiego. W ogromnej większości przypadków intuicja nie wystarcza jednak do określenia relatywnej wagi celów. Warto posługiwać się przy tym ustaloną *sekwencją wymienności* – czyli sposobem pomiaru i porównywania realizacji różnych celów. Uzgodniona i przyjęta sekwencja wymienności pozwala odpowiedzieć na pytanie o to, z realizacji jakich założeń opłaca się rezygnować na rzecz realizacji innych celów (*trade-offs*). Pomocna staje się w tej sytuacji funkcja celu (*objective function*), zwana też funkcją wyboru, w której poszczególnym zmiennym celu przypisane są określone wagi (Sulmicki 1978). Przykład takiej funkcji został przedstawiony po prawej stronie rysunku 1.8. W procesie jej konstrukcji konkretne, pożądane wartości celów nie są z góry ustalone, lecz pozostają raczej końcowym wynikiem określonego postępowania optymalizacyjnego.

Podejście alternatywne polega na zastosowaniu funkcji straty (*loss function*), nakazującej minimalizować odchylenia rzeczywistych wartości od założonych z góry wartości zmiennych celów, przy spełnieniu pewnych ograniczeń. Taki sposób postępowania staje się uzasadniony, gdy władza gospodarcza z góry wie, jakie konkretne wartości celów chce osiągnąć, i zarazem jest w stanie określić wielkość społecznej straty jako rosnącą funkcję dystansu pomiędzy założonym celem a rzeczywiście osiągniętą wartością zmiennej celu. W ramce 1.8 opisano sposób wyprowadzania funkcji straty oraz wskazano niezbędne warunki optymalności.

Ramka 1.8. Funkcja straty*

Typowa funkcja straty dla n różnych z góry ustalonych celów $\bar{Y}(i)$ ma postać:

$$L = L\left[\left(Y_1 - \bar{Y}_1\right), \left(Y_2 - \bar{Y}_2\right), \ldots, \left(Y_n - \bar{Y}_n\right)\right] \Rightarrow \min. \tag{1.19}$$

Jeśli rząd ma zarazem do dyspozycji m niezależnych instrumentów polityki $X_1, X_2, \ldots, X_m$, które wpływają na wartości zmiennych Y, wektor zmiennych celów $\boldsymbol{Y}$ może być przedstawiony jako funkcja wektora instrumentów $\boldsymbol{X}$:

$$\boldsymbol{Y} = \boldsymbol{H}(\boldsymbol{X}). \tag{1.20}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Polityka gospodarcza polega w takiej sytuacji na takim doborze instrumentów X, aby zminimalizować wartość funkcji straty L. Gdy $n \leq m$, znalezienie rozwiązania w postaci określonego wektora $\boldsymbol{X}$, dla którego $Y_i = \overline{Y}_i$ dla wszystkich i nie nastręcza trudności, wówczas wartość funkcji straty jest równa zeru ($L = 0$)[22]. Gdy jednak $n > m$, rząd musi dokonywać selekcji wartości instrumentów X w taki sposób, aby funkcja L osiągnęła minimum. Oznacza to, że pochodna funkcji celu musi być równa zeru. Musi więc zachodzić:

$$dL = \sum_{i=1}^{n} \sum \left[\partial L / \partial Y_i \right] \partial Y_i = 0, \tag{1.21}$$

czyli dla dowolnej pary zmiennych celu zachodzi prawidłowość:

$$\frac{dY_i}{dY_j} = -\left(\frac{\frac{\partial L}{\partial Y_j}}{\frac{\partial L}{\partial Y_i}} \right). \tag{1.22}$$

Oznacza to, że krańcowa stopa substytucji w przypadku dowolnej pary zmiennych celu (Y_i, Y_j) jest równa odwrotności pochodnych cząstkowych funkcji straty. Z warunku tego wynika, że dla L = minimum poprawa wykonania jakiegokolwiek celu jest równoważna z pogorszeniem się realizacji innego założenia w proporcji odwrotnej do relacji wpływu tej zmiany na funkcję straty[23].

Polityka gospodarcza to także sztuka wyboru najlepszego instrumentu służącego realizacji danego celu. W tym kontekście najczęściej przywołuje się słynną zasadę efektywnego przyporządkowania (*the principle of effective market classification*), sformułowaną przez kanadyjskiego ekonomistę Roberta Mundella w 1960 roku (Mundell 1960). Efektywne przyporządkowanie oznacza dobór takiego instrumentu, który wywiera na daną zmienną celu najsilniejszy bezpośredni wpływ (więcej na ten temat zob. Rosati 1990, rozdział 4). Przykładowo można wykazać, że w systemie stałego kursu walutowego polityka fiskalna jest generalnie bardziej skuteczna przy przywracaniu równowagi wewnętrznej (pobudzanie wzrostu

[22] Przy założeniu, że funkcja L (.) jest wklęsła i dwukrotnie różniczkowalna, i że L (0,0,…0) = 0.
[23] Zob. Benassy-Quéré i in. 2010, s. 16–17.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

produkcji i zatrudnienia), podczas gdy polityka pieniężna pozostaje generalnie bardziej skuteczna w procesie przywracania równowagi zewnętrznej (równowaga bilansu płatniczego).

Wreszcie polityka gospodarcza to także sztuka dokonywania rozkładu korzyści i kosztów między członków społeczeństwa w danym momencie i w dłuższej perspektywie, między pokoleniami. Szczególnie trudne jest znalezienie pożądanej równowagi między ekonomiczną efektywnością a społeczną sprawiedliwością. Wybór ten zależy od relacji krańcowych użyteczności dochodu w odniesieniu do różnych grup społeczeństwa oraz ich przemian w czasie, ale, ponieważ wielkości te nie są bezpośrednio obserwowalne, rozstrzyganie o rozkładzie korzyści między poszczególnych członków społeczeństwa – teraz i w przyszłości – pozostaje zawsze decyzją polityczną.

1.3.9. Planowanie

Uzasadnieniem w przypadku planowania gospodarczego staje się potrzeba koordynacji działań różnych podmiotów w ramach bardzo złożonych systemów, jakimi są współczesne gospodarki. Brak koordynacji (*coordination failure*) pociąga za sobą straty wynikające z nieoptymalnych decyzji, związanego z tym marnotrawstwa ograniczonych zasobów i konieczności korygowania błędów w działaniu. Wprawdzie idea planowania gospodarczego straciła na znaczeniu po nieudanych doświadczeniach krajów, w których wprowadzono gospodarkę centralnie planowaną, ale żadne racjonalne działanie w gruncie rzeczy nie może obejść się bez projektowania. Koordynacja w skali makroekonomicznej w formie planowania może mieć charakter bardziej lub mniej scentralizowany.

W formie scentralizowanej planowanie ma charakter dyrektywny, nakazowy. Polega na wyznaczaniu pewnych celów strategicznych i operacyjnych oraz na pełnym podporządkowaniu tym celom działań różnych podmiotów zarówno wśród organów władzy publicznej, jak i uczestników rynku. Cechą planowania scentralizowanego jest istnienie prawnie obowiązujących nakazów co do celów działania oraz administracyjny sposób rozdziału zasobów (Sulmicki 1978). Procedura planistyczna w tym przypadku ma charakter odgórny (*top-down*), a zadania i zasoby są przydzielane przedsiębiorstwom w postaci obowiązkowych zadań. Wpływ mechanizmu rynkowego na alokację i dystrybucję jest w takim systemie zredukowany do minimum. Ta forma planowania makroekonomicznego została zupełnie zdyskredytowana w czasach tzw. realnego socjalizmu. Do jej głównych wad

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

należą: arbitralny charakter wyznaczanych celów, marnotrawstwo środków, brak przestrzeni w przypadku inicjatywy niższych szczebli zarządzania, zniekształcone motywacje uczestników życia gospodarczego.

W formie zdecentralizowanej planowanie ma charakter indykatywny, orientacyjny. Wyznaczone cele nie są obowiązujące, stanowią raczej pożądane wskaźniki. W podejmowaniu decyzji co do alokacji zasobów oraz poziomu i struktury produkcji i konsumpcji istotną rolę odgrywa mechanizm rynkowy. Procedura planistyczna ma często charakter oddolny (*bottom-up*). Wyniki ekonomiczne tej formy planowania są z reguły znacznie lepsze, ponieważ wykorzystuje ona bodźce ekonomiczne do poprawy zarówno racjonalności zachowań uczestników rynku, jak i efektywności gospodarowania.

W wielu krajach, w których funkcjonuje gospodarka rynkowa lub mieszana, wykorzystuje się elementy planowania w formie indykatywnej w celu ograniczenia niepewności związanej z przyszłością. Uczestnicy rynku, znając cele władzy publicznej, są w stanie lepiej planować własne decyzje. Dzięki temu maleje ryzyko i obniżają się koszty związane z prowadzeniem działalności gospodarczej. Nie usuwa to jednak całkowicie arbitralności i uznaniowości w podziale zadań i środków oraz w ocenie działalności przedsiębiorstw.

1.4. Wzrost roli państwa w gospodarce w XX wieku

Niesprawności rynku, ujawniające się w ponad dwustuletnim okresie funkcjonowania gospodarki rynkowej, dostarczały uzasadnienia dla interwencji władzy publicznej i powodowały przejmowanie przez państwo szeregu zadań gospodarczych. Proces rozbudowywania funkcji państwa i poszerzania zakresu działalności władzy publicznej w gospodarce przebiegał najbardziej intensywnie w okresie po II wojnie światowej, a zwłaszcza w trzydziestoleciu od końca lat 40. do końca lat 70. Proces ten, dotyczący wzrostu roli państwa w kształtowaniu gospodarki, był wynikiem oddziaływania kilku czynników.

Po pierwsze, pod przemożnym wpływem myśli Keynesa i jego następców rozwinęło się i ugruntowało przekonanie o potrzebie oraz celowości interwencji rządu w zakresie gospodarki w celu łagodzenia przebiegu cyklu koniunkturalnego i stabilizowania poziomu produkcji oraz zatrudnienia. Keynes nie tylko stworzył przekonujące uzasadnienie w przypadku podejmowania działań stabilizacyjnych w postaci polityki makroekonomicznej w celu ograniczania bezrobocia i podtrzymywania

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

aktywności gospodarczej, lecz także wskazywał na potrzebę dokonywania redystrybucji dochodu na większą skalę w celu osiągania lepszych rezultatów w zakresie urzeczywistniania postulatów sprawiedliwości społecznej oraz ograniczania rozpiętości dochodowych i majątkowych.

Po drugie, tragiczne doświadczenia I i II wojny światowej uświadomiły opinii publicznej w krajach rozwiniętych, że bieda, wyzysk, bezrobocie i niesprawiedliwość społeczna sprzyjają konfliktom oraz agresji, a więc warunkiem utrzymania pokoju i uniknięcia kolejnych wojen jest zapewnienie pracy oraz godziwych warunków życia szerokim rzeszom obywateli.

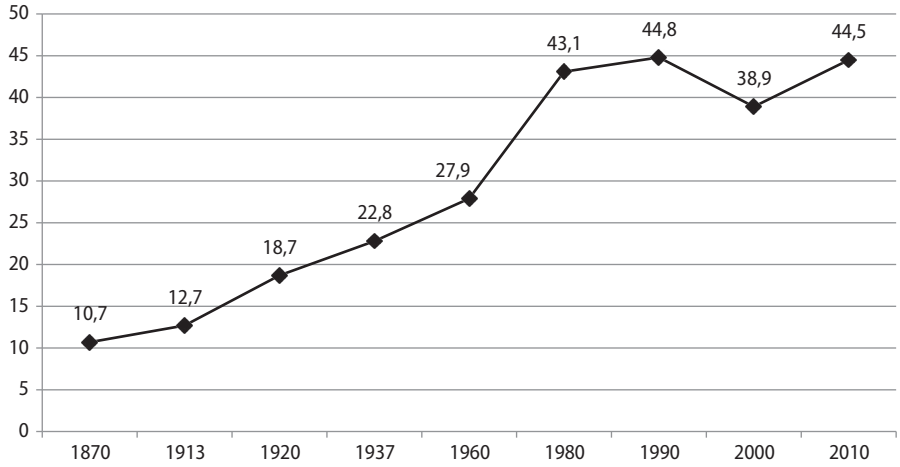
Po trzecie, budowa różnych odmian systemu centralnego planowania, zwłaszcza w ówczesnym Związku Radzieckim i krajach Europy Środkowo-Wschodniej, skłaniała kraje kapitalistyczne do poszukiwania lepszych metod planowania i koordynacji działalności gospodarczej. Początkowe sukcesy krajów, w których funkcjonowała gospodarka centralnie planowana, w procesie forsownej industrializacji i realizacji szerokiego zakresu celów społecznych, w tym eliminacji (jawnego) bezrobocia, sugerowały, że interwencja państwa może być skuteczna w nadrabianiu zacofania gospodarczego i usuwaniu niesprawności rynkowych. Przykład pochodzący z Europy Środkowo-Wschodniej stanowił potężny impuls do podjęcia w krajach zachodnich wysiłków w celu budowy nowego typu państwa – państwa opiekuńczego lub państwa dobrobytu (*welfare state*). Jego cechą charakterystyczną było przejęcie przez władze publiczne odpowiedzialności za realizację szeregu ważnych funkcji społecznych, takich jak edukacja, ochrona zdrowia, ochrona praw pracowniczych, zabezpieczenie emerytalne i opieka społeczna[24].

Wzrost roli państwa w gospodarce po II wojnie światowej był ściśle powiązany z dążeniem do przejmowania przez państwo coraz to nowych funkcji i zadań. Rosnące wydatki publiczne wymagały nakładu dodatkowych środków w celu ich finansowania, co powodowało wzrost skali i zakresu opodatkowania. Niemiecki ekonomista Adolf Wagner już na początku XX wieku trafnie przewidział kierunek tej tendencji. Tak zwane prawo Wagnera, określane też mianem prawa rosnących wydatków publicznych, mówi o tym, że udział państwa w PKB rośnie w miarę rozwoju gospodarczego. Mechanizmem, który napędza wzrost wydatków publicznych, jest system demokratycznych wyborów, zmuszający polityków do oferowania wyborcom coraz nowszych świadczeń (Musgrave 1959).

[24] Więcej na ten temat zob. Rosati (2008).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 1.1. Udział wydatków publicznych w PKB krajów uprzemysłowionych w latach 1870–2010 (dane w %)



Źródło: opracowanie własne na podstawie V. Tanzi, L. Schuknecht, *Public Spending in the 20th Century. A Global Perspective*, Cambridge University Press, Cambridge 2000, s. 6 – dla danych z lat 1870–1990, oraz dane OECD z roku 2000 i 2010.

Szybki rozrost zakresu interwencjonizmu i funkcji państwa w gospodarce uległ zahamowaniu w latach 90. ubiegłego stulecia. Wykres 1.1 ilustruje zmiany, jakie dokonały się w zakresie udziału wydatków publicznych w PKB krajów OECD po II wojnie światowej. Po okresie silnego wzrostu w latach 1950–1990, nastąpiło wyhamowanie, a następnie spadek[25] tej wartości. Zmiana trendu była wynikiem uświadomienia sobie dwóch ważnych ograniczeń. Po pierwsze, wraz ze wzrostem udziału wydatków publicznych w PKB malała ich krańcowa efektywność. Po drugie, pojawiły się kłopoty z finansowaniem wydatków – ich systematyczny wzrost wymagał albo ciągłego zwiększania bieżących obciążeń podatkowych, co miało ujemny wpływ na wzrost gospodarczy, albo zaciągania pożyczek, co z kolei powodowało narastanie długu publicznego i oczekiwanie wzrostu podatków w przyszłości. W rezultacie od drugiej połowy lat 80. wśród większości krajów rozwiniętych dostrzega się zjawisko wycofywania się państwa z niektórych obszarów gospodarki.

[25] Wzrost udziału wydatków publicznych w PKB w 2010 roku był częściowo wynikiem spadku PKB w latach 2008–2010 w następstwie kryzysu finansowego oraz przejściowego wzrostu wydatków publicznych na cele stymulacji fiskalnej w okresie kryzysu. Na przykład w krajach UE27 udział wydatków publicznych w PKB spadł do 49,1% w 2011 roku z 50,6% w 2010 roku, a w strefie euro odpowiednio do 49,4% z 51,0%.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Proces ten przebiegał w dwóch głównych postaciach – prywatyzacji i deregulacji. Po fali nacjonalizacji wielu gałęzi produkcji w latach 1950–1980 rządy poszczególnych państw zaczęły stopniowo wycofywać się z roli aktywnego właściciela zasobów produkcyjnych. Powodem były przede wszystkim wysokie koszty i niska efektywność zarządzania majątkiem produkcyjnym przez instytucje państwowe. Prywatyzacji poddano zwłaszcza przedsiębiorstwa w sektorach o wysokim udziale własności państwowej, takich jak przemysł ciężki, górnictwo, energetyka, telekomunikacja. Z kolei deregulacja, czyli liberalizacja obrotu gospodarczego, oznaczała ogólnie zmniejszenie zakresu administracyjnych ograniczeń i obniżenie poziomu intensywności kontroli państwa w zakresie gospodarki. Zjawisko to było spowodowane rosnącym przekonaniem, że biurokratyczne bariery i administracyjne ograniczenia nałożone na funkcjonowanie mechanizmu rynkowego niosą szereg szkodliwych skutków pod względem innowacyjności i konkurencyjności gospodarki.

W obu przypadkach odwrót od dominującej roli państwa w kształtowaniu gospodarki był wynikiem uświadomienia sobie ograniczeń i negatywnych skutków, jakie pociąga za sobą nadmierny interwencjonizm. Władza publiczna wkraczała w obszar gospodarki z zamiarem korygowania i usuwania niesprawności rynkowych. Państwo nie jest jednak ani nieomylne, ani bezstronne, ani wszechmocne. Sprawuje swe funkcje w określonym układzie prawno-instytucjonalnym, realizując cele i priorytety, które są wynikiem przetargu politycznego, oraz przy pomocy urzędników (biurokratów), których kompetencje i wiedza są ograniczone, a motywy działania często partykularne. W rezultacie często dochodziło do sytuacji, w których próba usunięcia skutków niesprawności rynku powodowała pojawienie się nowych negatywnych zjawisk, tym razem wynikających z ułomności państwa. Mogło się więc okazać, że w ostatecznym rozrachunku interwencja władzy publicznej przynosiła społeczne straty.

1.5. Pozytywna teoria polityki gospodarczej: niesprawności państwa

Normatywna teoria polityki gospodarczej stara się odpowiedzieć na pytanie o to, jakie cele państwo powinno sobie wyznaczać w zakresie gospodarki oraz w jaki sposób należy je realizować. Odpowiedź jest formułowana przy (często milczącym) założeniu, że państwo pozostaje bezstronnym, obiektywnym, efektywnym

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i racjonalnym reprezentantem interesu ogólnospołecznego. Istnieje jednak wiele powodów pozwalających uznać to założenie za bardzo mało realistyczne.

Po pierwsze, władza publiczna – czyli rząd – działa w danym kraju zawsze w ramach określonego systemu politycznego i reprezentuje najczęściej interesy oraz poglądy jedynie wybranych, określonych grup społecznych i partii politycznych. Funkcjonowanie systemu demokratycznego polega na tym, że władzę sprawują przedstawiciele partii i ugrupowań, które otrzymały w wyborach największe poparcie. W warunkach istnienia różnych grup społecznych i zawodowych, posiadających różne, często rozbieżne, cele oraz interesy trudno jest określić właściwy kształt interesu publicznego, ogólnospołecznego. Po drugie, każda władza działa w warunkach niepełnej informacji, a jej przedstawiciele mają ograniczone kompetencje merytoryczne i dysponują niepełną wiedzą na temat zjawisk gospodarczych. Brak dostatecznej wiedzy o gospodarce może wynikać z obiektywnych przesłanek – przykładowo, prognozowanie przyszłości, niezbędne w przypadku prowadzenia skutecznej polityki pieniężnej, jest zawsze obarczone niepewnością. Brak dostatecznej wiedzy może też wynikać z czynników subiektywnych, takich jak ignorancja oraz brak kwalifikacji do wykonywania pewnych działań. Po trzecie, władza gospodarcza nie jest monolitem – na najbardziej podstawowym poziomie klasyfikacji składa się z dwóch kategorii osób – polityków i urzędników, którzy posiadają własne, odrębne i często sprzeczne interesy oraz dążenia i kierują się w swoim postępowaniu właściwą sobie strukturą bodźców. Te zróżnicowane indywidualne kompetencje i motywacje sprawiają, że władza publiczna niemal nigdy nie pozostaje całkowicie bezstronna, obiektywna i racjonalna, a zatem relatywnie rzadko działa wyłącznie w interesie ogólnospołecznym.

1.5.1. Czy istnieje interes publiczny?

Czy w ogóle istnieje jednak coś takiego jak interes ogólnospołeczny, publiczny? A jeśli tak, to czy możliwe jest jego zdefiniowanie w przypadku celów praktycznych? Na najbardziej ogólnym poziomie rozważań odpowiedź na te pytania jest oczywiście twierdząca. Zapewne wszyscy zgodzą się, że takie działania jak egzekwowanie obowiązującego prawa, bezpieczeństwo i ochrona zdrowia konsumentów czy też ochrona środowiska naturalnego leżą w interesie ogólnospołecznym. Podobnie w interesie publicznym jest też podnoszenie poziomu wykształcenia społeczeństwa. Lecz na poziomie bardziej szczegółowym lub w odniesieniu do konkretnych decyzji pojawiają się już wątpliwości. Przykładowo, egzekwowanie prawa może być

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wykonywane przy pomocy różnych instrumentów prawnych, w tym instrumentów przymusu indywidualnego, ograniczających swobody obywatelskie, co sprzyja powstawaniu obaw co do zgodności tych metod z prawami podstawowymi, takimi jak wolność słowa czy wolność zgromadzeń. W dodatku niektóre przepisy prawne mogą być uważane powszechnie za nieuzasadnione. Z kolei polityka ochrony zdrowia konsumentów bądź środowiska naturalnego może narzucać tak ostre ograniczenia na prowadzenie działalności gospodarczej, że będzie to wiodło do utraty konkurencyjności i bankructwa firm.

Społeczeństwo nie jest zbiorem osób identycznych, składa się z jednostek reprezentujących rozmaite poglądy, potrzeby i interesy. Wśród takiej zbiorowości zawiązują się grupy społeczne składające się z jednostek wyznających jednakowe lub zbliżone poglądy, potrzeby lub interesy. Posiadają one swoich instytucjonalnych reprezentantów, takich jak partie polityczne, związki zawodowe, związki pracodawców, organizacje konsumentów, korporacje zawodowe, wspólnoty lokalne i regionalne czy organizacje pozarządowe (Axelrod 1984). Choć niekiedy grupy te współpracują ze sobą, to częściej dochodzi między nimi do rywalizacji o zakres kontroli sprawowanej nad zasobami i udział w wytworzonym produkcie. Zagadnienia związane z podziałem majątku oraz dochodu między poszczególne grupy społeczne w danej gospodarce stanowią przedmiot zainteresowania ekonomii politycznej.

Zderzanie się i ścieranie różnych interesów grupowych w obrębie danego społeczeństwa jest zjawiskiem normalnym. Spontaniczny, nieunormowany sposób ich rozstrzygania prowadziłby jednak do tyranii grup silniejszych nad słabszymi. Dlatego też w demokracji systemowym sposobem rozstrzygania konfliktu interesów są okresowe wybory, w których wyborcy udzielają poparcia określonym partiom politycznym i konkretnym politykom. Większość otrzymuje mandat do sprawowania władzy, ale mniejszość podlega ochronie. Wyniki wyborów odzwierciedlają preferencje wyborców w zakresie celów polityki gospodarczej i sposobów ich realizacji. W takich warunkach wybrany rząd będzie skłonny realizować w ramach polityki gospodarczej przede wszystkim interesy partykularne tych grup społecznych, które są jego politycznym zapleczem i zapewniają mu poparcie w wyborach, w mniejszym stopniu będzie zaś kierował się interesem ogólnospołecznym. Przykładem takich różnic w polityce gospodarczej, wynikających z odmiennych oczekiwań elektoratów różnych partii politycznych, jest polityka podatkowa w Stanach Zjednoczonych. Gdy rządzą republikanie, administracja dąży z reguły do obniżania podatków, a gdy rządy obejmują demokraci, administracja jest bardziej skłonna podnosić podatki, zwłaszcza wśród najbogatszych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Zagadnienia związane z wpływem interesów poszczególnych grup społecznych na decyzje z dziedziny polityki gospodarczej są przedmiotem zainteresowania teorii wyboru publicznego (*public choice theory*). Jej przedstawiciele (Buchanan, Tullock, Arrow, Niskanen) dowodzą, że w warunkach demokracji mechanizm wyborów staje się instrumentem realizowania interesów prywatnych przez określone grupy nacisku. Autorzy ci wskazują także na to, że prowadzona polityka gospodarcza odzwierciedla na ogół preferencje tych grup, które posiadają większość w systemie sprawowania władzy. Nawet jeśli w tej sytuacji dochodzi do uwzględniania interesów innych grup społecznych, wypadkowa tych uzgodnień rzadko reprezentuje wyłącznie interes publiczny.

Interesującą perspektywę badawczą w tym zakresie przedstawia teoria politycznego cyklu koniunkturalnego, związana z pracami takich autorów jak Nordhaus (1975), Lindbeck (1976), a wcześniej także Kalecki (1943)[26]. Dowodzą oni, że istnieje określony związek między polityką gospodarczą a demokratycznymi wyborami. Celem polityków jest z reguły zdobycie lub utrzymanie władzy, co skłania ich do składania takich obietnic wyborczych, i w konsekwencji do prowadzenia takiej polityki gospodarczej, które spełniają oczekiwania określonych grup wyborców i gwarantują im uzyskanie ich poparcia w wyborach. Polityka gospodarcza, będąca wynikiem tego typu motywacji ze strony polityków, nie musi jednak być zgodna z interesem ogólnospołecznym. Dzieje się tak dlatego, że sami wyborcy mają skłonność do dokonywania nieracjonalnych, krótkookresowych ocen i wyborów, a zatem są skłonni popierać polityków których obietnice rokują doraźne korzyści, ale mogą pociągać za sobą długookresowe straty. Ta nieracjonalność obecna w zachowaniach wyborców wynika głównie z asymetrii informacji, powodującej, że wyborcy są na ogół pozbawieni pełnej wiedzy na temat konsekwencji określonych decyzji politycznych (Rogoff i Sibert 1988). W rezultacie pojawia się tendencja do preferowania korzyści doraźnych, przejściowych kosztem korzyści trwałych, ale głównie takich, które pojawią się dopiero w dłuższym okresie. Alesina, Cohen i Roubini (1992) wykazali, że cykle wyborcze są jedną z przyczyn systematycznego

[26] Kalecki wprowadził jako pierwszy pojęcie „politycznego cyklu koniunkturalnego" w 1943 roku. Zwrócił wówczas uwagę na to, że, dążąc do utrzymania pełnego zatrudnienia i finansowania szerokiego zakresu funkcji społecznych państwa (państwo „opiekuńcze"), rząd musi zwiększać deficyt i pożyczać pieniądze, co prędzej czy później doprowadzi do wystąpienia presji inflacyjnej. Zmusi to rząd do powrotu do bardziej restrykcyjnej polityki fiskalnej i pieniężnej, co spowoduje osłabienie koniunktury. Te naprzemienne okresy dobrej i złej koniunktury są rezultatem wyborów politycznych. W ten sposób ekonomiczny cykl koniunkturalny zostaje zastąpiony politycznym cyklem koniunkturalnym (Kalecki 1943, wydanie polskie 1979, s. 348–349).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wzrostu wydatków publicznych i narastania deficytu budżetowego (tzw. *deficit bias*), co może mieć w przyszłości groźne konsekwencje, np. w postaci kryzysów zadłużenia (więcej na ten temat zob. rozdział 3). Jeśli jednak polityka gospodarcza jest kształtowana w taki sposób, że może prowadzić do kryzysu, to pozostaje w oczywisty sposób sprzeczna z interesem publicznym. Teoria cyklu wyborczego potwierdza więc wcześniejsze spostrzeżenia dotyczące tego, że faktycznie realizowana w warunkach demokracji polityka gospodarcza w wielu przypadkach nie reprezentuje interesu ogólnospołecznego.

1.5.2. Problem reprezentacji

Teoria normatywna polityki gospodarczej pomija to, że termin „władza publiczna" jest daleko idącym skrótem myślowym. Sugeruje istnienie jednego homogenicznego podmiotu, składającego się z osób nieposiadających żadnych osobistych przekonań i preferencji, a których jedyną motywacją w działaniu jest troska o interes publiczny. W rzeczywistości sytuacja ta jest o wiele bardziej złożona. Jak wcześniej wspomniano, władza publiczna składa się z dwóch kategorii osób – polityków i urzędników (biurokratów). Ci pierwsi są wybierani przez wyborców, a ich motywacje określone zostają częściowo lub w dominującej mierze przez chęć zdobycia (elekcja) i utrzymania się przy władzy (reelekcja). Drudzy mianowani są przez polityków, a ich motywacje określone są generalnie przez chęć utrzymania stanowisk, wpływu i zapewnienie sobie dalszego awansu. Byłoby daleko idącym uproszczeniem gdyby uznać, że i jedni, i drudzy w sposób właściwy reprezentują wszystkich wyborców, czyli ogół społeczeństwa.

To, że politycy pozostają stronniczy w tym sensie, że reprezentują grupowy (partyjny) raczej, a nie ogólnospołeczny punkt widzenia, nie powinno być niespodzianką. Nie musi to zresztą wynikać z wąsko pojmowanego partykularyzmu rozumianego jako postawa nastawiona na zagarnięcie konkretnych korzyści przez członków określonej grupy społecznej czy partii politycznej. Najczęściej odbywa się to skutkiem przywiązywania różnej wagi do rozmaitych celów politycznych, w tym do określonych kosztów i korzyści wynikających z przyjęcia danego kursu polityki gospodarczej, a których rozkład jest w społeczeństwie zróżnicowany. Partie różnią się między sobą nie tylko pod względem bardziej abstrakcyjnej ideologii, lecz także bardziej praktycznej polityki. Przykładowo, w danej sytuacji gospodarczej partie lewicowe będą na ogół proponowały inny program działania niż partie konserwatywne czy liberalne. Różnice zachodzące między partiami dotyczą zresztą

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

nie tylko różnych propozycji w zakresie polityki gospodarczej, lecz sięgają znacznie głębiej – aż do rozmaitych i konkurencyjnych względem siebie teorii dotyczących funkcjonowania gospodarki i państwa.

Różnice te stają się szczególnie widoczne w odniesieniu do roli wydatków publicznych w gospodarce. W związku z tym w świetle jednej koncepcji polityki gospodarczej wzrost wydatków publicznych finansowanych długiem ma głównie skutki „keynesowskie", a zatem powoduje wzrost produkcji i zatrudnienia, podczas gdy związany z tym wzrost dochodów podatkowych pozwala spłacić zaciągnięty dług. W świetle innej koncepcji wzrost wydatków publicznych finansowanych długiem ma głównie skutki „niekeynesowskie", a więc powoduje spadek produkcji i zatrudnienia wskutek obaw o wzrost podatków w przyszłości w związku z koniecznością spłaty długu oraz z powodu rosnącego ryzyka niewypłacalności kraju. Częściowo wynika to z istnienia różnych teorii ekonomicznych odnoszących się do tych samych zjawisk, niedających się na podstawie badań empirycznych jednoznacznie potwierdzić lub odrzucić. Częściowo wynika to jednak także z odmiennych założeń ideologicznych oraz aksjologicznych, stanowiących podstawę programów politycznych poszczególnych ugrupowań.

Jeszcze inaczej zachowują się urzędnicy sprawujący funkcję władzy publicznej w państwie. Oczywiście nie można wykluczyć, że funkcjonariusze państwowi kierują się w swoim postępowaniu interesem publicznym, tj. starają się reprezentować ogół obywateli. Jednak, jak wskazali liczni autorzy (np. Niskanen 1971, 1973; Tullock 1965), częściej większe znaczenie mają osobiste aspiracje biurokratów, takie jak chęć zwiększenia dochodu, poszerzenie zakresu władzy, prestiż i pozycja społeczna. Zarówno te aspiracje, jak i szczególne miejsce instytucji publicznych w państwie powodują powstanie szeregu negatywnych zjawisk, które ogólnie są źródłem istotnych niesprawności państwa.

Wskazuje się przede wszystkim na to, że biurokracja generuje wysokie koszty funkcjonowania. Wynika to częściowo z charakteru realizowanych w jej ramach usług, które nie mają „ceny" w sensie rynkowym, a zatem jakikolwiek rachunek efektywności jest w tym przypadku poważnie utrudniony. W szczególności zaś pomiar jakości pracy urzędników okazuje się w praktyce bardzo skomplikowany lub wręcz niewykonalny. Prowadzi to często do marnotrawstwa zasobów. Wysokie koszty funkcjonowania biurokracji są jednak także spowodowane trwałą tendencją do nadmiernej ekspansji i rozrostu instytucji publicznych – zarówno pod względem zakresu działań, jak i pochłanianych środków. Powodowane jest to z kolei dążeniem poszczególnych urzędników do poszerzenia zakresu władzy i kompetencji,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

podnoszenia prestiżu wykonywanej funkcji i uzasadniania konieczności własnego istnienia. Podkreśla się w tym kontekście różnicę między efektywnością pracy urzędników w sektorze publicznym i w sektorze prywatnym. W tym ostatnim przypadku nadzór właścicielski, znacznie większa waga przywiązywana do kontroli kosztów i konkurencyjne zasady zatrudniania oraz awansowania przekładają się generalnie na niższe koszty działania.

Podkreśla się także, że ze względu na odmienną funkcję celu urzędnicy są bardziej skłonni do przejawiania zachowań pasywnych niż aktywnych, co może obniżać konkurencyjność danej gospodarki. Odzwierciedla to tendencja do unikania konieczności podjęcia decyzji, zwłaszcza tych, które są obciążone większym ryzykiem. Jest to strategia zapewniająca bezpieczeństwo, ponieważ odpowiedzialność za brak rozstrzygnięcia (zaniechanie) jest zawsze trudniej udowodnić niż odpowiedzialność za podjęte błędne decyzje. Choć trzeba uznać, że w warunkach takiej asymetrii ryzyka bierność i odwlekanie momentu podjęcia decyzji jest z punktu widzenia urzędników zachowaniem racjonalnym, nie można jednak uznawać tej sytuacji za coś nieuchronnego. Doskonalenie i upraszczanie procedur administracyjnych, przejrzyste, okresowe systemy oceny pracy i osiągnięć poszczególnych urzędników, zachęta do przejawiania aktywności i wprowadzanie elementów konkurencji między urzędnikami mogą znacznie ograniczyć skalę tych niekorzystnych konsekwencji.

Funkcjonowanie biurokracji może też stanowić sprzyjające podłoże do rozpowszechniania się innych szkodliwych zjawisk, takich jak nadużycia władzy, korupcja, faworytyzm i nepotyzm czy oligarchizacja instytucji państwowych (tj. ich nieformalne podporządkowanie interesom wąskiej grupy najbogatszych członków społeczeństwa). Brennan i Buchanan (1980) dodają, że biurokracja, i – szerzej – państwo jako takie, wykazuje tendencję do przechwytywania prywatnych zasobów produkcyjnych poprzez nadmierne opodatkowanie. Wszystko to powoduje straty społeczne z powodu nieoptymalnej alokacji zasobów i zwiększa koszty prowadzenia biznesu (koszty transakcyjne).

Nieefektywność działania sektora publicznego może także wynikać z konfliktów na linii politycy – urzędnicy. Jak już wspomniano, funkcje celu obu tych części składowych władzy publicznej są odmienne, co nieuchronnie przekłada się na rozbieżność dążeń i konflikt interesów. Mimo że urzędnicy są formalnie zależni od polityków, to politycy pozostają jednak w tym konflikcie słabszą stroną. Mają mniejsze rozeznanie w szczegółach nadzorowanych obszarów spraw, częściej podlegają rotacji na skutek wyborów, są też z reguły słabiej przygotowani merytorycznie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W takiej sytuacji punkt widzenia urzędników często przeważa w kształtowaniu bieżącej polityki.

Niesprawność państwa w zakresie funkcjonowania władzy publicznej można ograniczać, choć zapewne nie da się jej całkowicie wyeliminować. W odniesieniu do biurokracji dobre efekty przynosi ograniczenie dyskrecjonalności działania urzędników, upowszechnianie jasnych, przejrzystych procedur postepowania oraz stosowanie systemów rekrutacyjnych, motywacyjnych i oceny zbliżonych do tych, jakie funkcjonują w sektorze prywatnym (np. konkursy, premiowanie działań – zob. Stiglitz 1994). W odniesieniu do polityków warto dbać o przejrzyste procedury ich selekcji oraz wzmacnianie demokratycznych instrumentów kontroli ich działalności, w tym ze strony niezależnych mediów i organizacji obywatelskich.

1.5.3. Problem ograniczonej wiedzy i niepełnej informacji

Teoria normatywna ogranicza się do wskazywania celów interwencji władzy publicznej w zakresie gospodarki i ignoruje jej rzeczywistą zdolność do skutecznego działania. Tymczasem, nawet zakładając, że władza publiczna kieruje się wyłącznie interesem publicznym oraz jest bezstronna i obiektywna, jej interwencja może być nieskuteczna, spóźniona lub wręcz szkodliwa. Przyczyną takiego stanu rzeczy może być z jednej strony niedostatek informacji o rzeczywistym stanie gospodarki, z drugiej zaś strony niedostatek wiedzy (kompetencji), na podstawie której podejmuje się decyzje o celowości i sposobie interwencji.

Niedostatek informacji jest czymś powszechnym i naturalnym w świecie, który pozostaje nieskończenie złożony. Nikt i nigdy nie ma pełnej wiedzy o całej gospodarce. To właśnie ograniczenia informacyjne stały się jedną z zasadniczych przeszkód uniemożliwiających zrealizowanie w XX wieku utopijnego projektu stworzenia gospodarki w pełni scentralizowanej[27]. Władza publiczna musi podejmować decyzje na podstawie ograniczonego zbioru wybranych informacji, odnoszących się jedynie do zmiennych obserwowalnych, w tym zagregowanych, które na dodatek docierają do decydentów z opóźnieniem. Przekłada się to na postulat dbania o jakość i zakres informacji statystycznej oraz sprawozdawczej, która pozostaje ważnym czynnikiem określającym skuteczność polityki.

[27] Inną kluczową przeszkodą były zniekształcone motywacje, premiujące bierność oraz zniechęcające do podejmowania wysiłku i wprowadzania innowacji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Drugi czynnik stanowią merytoryczne kompetencje polityków i biurokratów. Brak wiedzy o funkcjonowaniu gospodarki, rządzących nią głównych prawidłowościach i zależnościach uniemożliwia praktycznie prowadzenie jakiejkolwiek sensownej polityki gospodarczej. W takich warunkach zmierza ona bowiem w stronę woluntaryzmu, a nawet awanturnictwa gospodarczego, co prowadzi do chaosu, marnotrawstwa środków i z reguły kończy się głębokim kryzysem, a w przyszłości staje się przyczyną zacofania. Dominacja celów politycznych oraz ignorowanie rzeczywistych prawidłowości ekonomicznych stanowiły charakterystyczną cechę systemów gospodarczych opartych na centralnym planowaniu, co doprowadziło je ostatecznie do upadku.

Niedostatek wiedzy i informacji sprawia, że politycy często nie są w stanie sięgnąć po rozwiązania optymalne w danej sytuacji i pozostają zmuszeni stosować rozwiązania suboptymalne, „następne po najlepszych" (*second best*). Przykładowo, w większości krajów stosuje się obniżoną stawkę podatku VAT na podstawowe artykuły żywnościowe. Jest to spowodowane troską o utrzymanie odpowiedniego poziomu siły nabywczej dochodów uboższych warstw społecznych, dla których wydatki na żywność stanowią dominującą pozycję w budżetach domowych. Nie jest to jednak rozwiązanie optymalne, ponieważ z obniżonych cen na żywność korzystają wszyscy, w tym także osoby zamożne, co oczywiście nie znajduje uzasadnienia. Rozwiązaniem optymalnym w tej sytuacji byłoby subsydiowanie żywności tylko dla osób najuboższych, przy czym stawka subsydium byłaby tym wyższa, im niższy byłby dochód danej osoby. W praktyce takie rozwiązanie wymagałoby jednak posiadania szczegółowej informacji o sytuacji dochodowej wszystkich obywateli i dopiero na tej podstawie należałoby zbudować system subsydiowania, uwzględniający indywidualną sytuację każdego obywatela. Musiałby to być system bardzo skomplikowany i kosztowny w utrzymaniu. Zastosowanie go w praktyce byłoby niesłychanie trudne. Dlatego też w jego miejsce stosuje się rozwiązanie następne po najlepszym, jakim jest obniżona stawka VAT na produkty żywnościowe nabywane przez wszystkich. Sięganie po rozwiązania typu *second best* pozostaje powszechną praktyką. Może to wynikać nie tylko z niedostatku informacji bądź wiedzy, lecz także z ograniczeń politycznych, społecznych lub prawno-międzynarodowych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

1.5.4. Rola niepewności

Niedostatek wiedzy na temat gospodarki i jej uczestników zmusza również do stosowania szczególnych metod podejmowania decyzji, uwzględniających ryzyko i niepewność. Jednym z najbardziej popularnych sposobów jest stochastyczne modelowanie procesów gospodarczych, którego istotę można przedstawić w postaci następującego równania opisującego funkcjonowanie gospodarki:

$$Y_t = f\left(X_t,\ Y_{t-1},\ Y_{t-2},\ \ldots.\ ,\ k,\ \varepsilon_t\right), \tag{1.23}$$

gdzie Y_t to wektor wartości celów polityki (np. dochód, konsumpcja, inwestycje, zatrudnienie), X_t to wektor wartości instrumentów polityki (np. stopa procentowa, podatki, wydatki rządowe), k to wektor parametrów (np. skłonność do konsumpcji, elastyczność cenowa podaży i popytu), ε_t to wektor wartości zmiennych losowych niezależnych od władzy publicznej, a t jest indeksem czasu. Funkcja f reprezentuje określone zależności między zmiennymi. O ile nie występują nieprzewidziane wydarzenia lub ich wpływ wzajemnie się neutralizuje, o tyle stan gospodarki w okresie t zależeć będzie od działań rządu X w tym czasie i od jej stanu w poprzednich okresach, przy czym wpływ działań rządu jest uwarunkowany wartością parametrów k. Lecz związek między parametrami a zmiennymi celu nigdy nie jest jednoznacznie określony. Przykładowo, wiadomo, że wpływ zmiany cen na wielkość popytu na określone dobro jest z reguły ujemny, ponieważ elastyczność cenowa popytu pozostaje ujemna, ale nikt nie jest w stanie precyzyjnie jej ocenić. Ponadto w danym momencie popyt na wskazane dobro może zależeć od wielu innych czynników, niekoniecznie uwzględnionych w tym modelu. W rezultacie ma się do czynienia z dwoma źródłami niepewności – niepewnością wynikająca z przyjęcia określonej specyfikacji modelu oraz niepewnością dotyczącą nieznajomości wartości parametrów modelu w danym momencie.

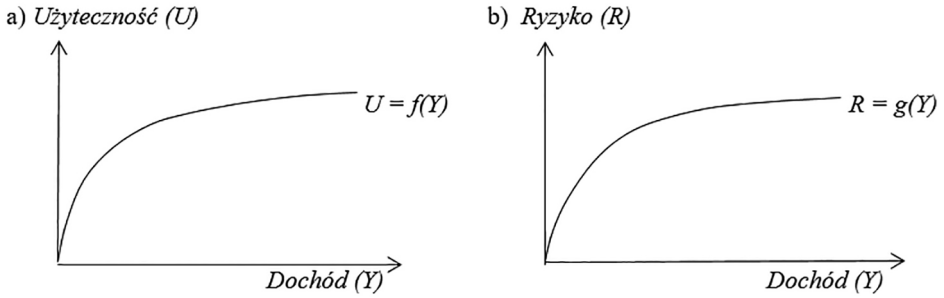
Ten rodzaj niepewności wstępuje szczególnie często w polityce makroekonomicznej. Na przykład „keynesowska" polityka pobudzania koniunktury przy pomocy zwiększania wydatków publicznych ma sens tylko wtedy, gdy mnożnik wydatków (mnożnik fiskalny – *fiscal multiplier*) jest wyraźnie większy od zera. Ale w momencie podejmowania decyzji o wydatkach nikt nie jest w stanie określić z całą pewnością, ile wynosi mnożnik. Jego wartość rzeczywistą poznaje się dopiero *ex post*, gdy ujawnią się skutki podjętej decyzji. Może bowiem stać się tak, że wzrost

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wydatków rządowych nie przyniesie oczekiwanego przyspieszenia wzrostu gospodarczego, dlatego że spowoduje on zaostrzenie polityki pieniężnej przez bank centralny, mogący obawiać się inflacyjnych skutków wzrostu wydatków rządowych.

Gdy niepewność co do przyszłości istnieje, ale da się zmierzyć, ma się wówczas do czynienia z ryzykiem. Uczestnicy rynku z reguły działają w warunkach ryzyka, starając się ocenić prawdopodobieństwo wystąpienia różnych stanów natury i odpowiednio dostosowując do tego strategie postepowania. Banki udzielają kredytów, biorąc pod uwagę nie tylko oczekiwany dochód, lecz także i poziom ryzyka związany z możliwością niewypłacalności dłużnika. Podobnie postępują inwestorzy decydujący się na zakup określonych aktywów, które mogą z różnych powodów stracić na wartości. Brak wiedzy o przyszłym rozwoju wydarzeń skłania do zachowywania ostrożności przy podejmowaniu decyzji inwestycyjnych. Miarą stopnia tej ostrożności jest awersja do ryzyka (*risk aversion*), określająca zależność między oczekiwanym dochodem a poziomem ryzyka. W szczególności zaś gdy ma się do czynienia z malejącą krańcową użytecznością dochodu (jest to jedno ze standardowych założeń teorii konsumpcji, potwierdzone wynikami badań empirycznych), w miarę wzrostu ryzyka inwestorzy będą skłonni do dalszego inwestowania tylko wtedy, gdy oczekiwany dochód wzrośnie bardziej niż proporcjonalnie (zob. lewa część rysunku 1.11).

Rysunek 1.11. Zależność między dochodem a użytecznością oraz między dochodem a ryzykiem



Źródło: opracowanie własne.

Poszczególni inwestorzy różnią się pod względem awersji do ryzyka – jedni są bardziej skłonni do podejmowania ryzyka, inni zaś bardziej ostrożni i zachowawczy. Na przykład uważa się, że uprzedzenie to rośnie z wiekiem. Formalnie stopień awersji do ryzyka zależy od stopnia wklęsłości krzywej awersji na prawej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

części rysunku 1.10 – im bardziej wklęsła krzywa, tym wyższy stopień awersji do ryzyka, a im mniejsze wygięcie krzywej, tym skłonność do ryzyka wyższa. Generalnie przyrostowi ryzyka o daną jednostkę musi towarzyszyć coraz większy przyrost dochodu, aby skłonić inwestora do podjęcia inwestycji. Zależność ta pozwala sformułować różne miary awersji do ryzyka[28].

Niepewność powoduje, że w niektórych sytuacjach korzystnie jest w ogóle nie podejmować decyzji i poczekać na dalszy rozwój wydarzeń. Opłacalność takiej strategii „czekania" wynika ze zjawiska nieodwracalności (*irreversibility*) skutków danych decyzji. Gdyby wszystkie decyzje podejmowane w ramach polityki gospodarczej odnosiły się do zmian krańcowych a ich skutki były odwracalne, polityka gospodarcza miałaby charakter adaptacyjny i dostosowywałaby się każdorazowo do bieżącego stanu gospodarki (*state-contingent*). W praktyce najczęściej jest jednak inaczej i czekanie może mieć sens wówczas, gdy nowe informacje są w stanie przesądzić o tym, czy dany projekt jest opłacalny. Rachunek strat i korzyści powinien w takiej sytuacji uwzględniać tzw. opcjonalną wartość czekania (*option value of waiting*), jeśli są podstawy, aby sądzić, że nowe informacje pozwolą lepiej oszacować przyszłe koszty i korzyści.

Opcjonalna wartość czekania ma znaczenie w odniesieniu do praktycznie niemal każdej decyzji inwestycyjnej. Powinny być one podejmowane tylko wtedy, gdy wewnętrzna stopa zwrotu z danej inwestycji jest wyższa od pewnego minimalnego poziomu (równego na przykład stopie oprocentowania kredytu). Jeśli rachunek kosztów i korzyści w danym momencie nie zapewnia wymaganej stopy zwrotu, decyzja o rozpoczęciu inwestycji powinna zostać odłożona do czasu, kiedy inwestor otrzyma nowe informacje na temat projektu. Szczególnie interesującym przypadkiem w kontekście opcjonalnego „czekania" są inwestycje dotyczące walki z ociepleniem klimatu. Wobec niepewności co do przyczyn i mechanizmu emisji gazów cieplarnianych, ich wpływu na ocieplenie klimatu oraz z uwzględnieniem ogromnej skali kosztów związanych z inwestycjami redukującymi emisję CO_2 w energetyce, przemyśle i transporcie, istnieje uzasadnienie dla wyboru strategii „czekania", by w miarę upływu czasu uzyskać więcej informacji na temat tego zjawiska i na tej podstawie wybrać skuteczniejszy oraz bardziej opłacalny wariant

[28] Bezwzględna awersja do ryzyka (*absolute risk aversion* – ARA) mierzona jest wyrażeniem: $ARA = -U''(Y)/U'(Y)$, gdzie U', U'' oznaczają odpowiednio pierwszą i drugą pochodną użyteczności będącej funkcją dochodu $U=f(Y)$. Dla wklęsłej krzywej U, $U'(Y) > 0$, a $U''(Y) < 0$. Względna awersja do ryzyka (*relative risk aversion* – RRA) mierzona jest wyrażeniem $RRA = -YU''(Y)/U'(Y)$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

polityki klimatycznej[29]. Należy jednak zauważyć, że „czekanie" może też wiązać się z ponoszeniem kosztów, co oznacza ujemną wartość strategii „czekania", którą z założenia należy uwzględnić w całościowym rachunku zysków i strat[30].

1.5.5. Rola oczekiwań

W tradycyjnym modelowaniu procesów gospodarczych zakładano stabilność związków między zmiennymi celów a instrumentami polityki, zgodnie z równaniem (1.23). Przyjmowano, że w tych warunkach polityka gospodarcza polega na ustaleniu wartości instrumentów X na takim poziomie (z reguły w ramach dopuszczalnych limitów), aby zminimalizować wartość pewnej funkcji straty (*loss function*), będącej sumą (ważonych) odchyleń osiągniętych rzeczywistych wartości celów Y od celów zaplanowanych $\bar{Y}$. W danym modelu wartości parametrów *k* były stałe, a czynnik losowy podlegał rozkładowi normalnemu z wartością oczekiwaną równą zeru. Taki sposób modelowania nie uwzględniał oczekiwań uczestników rynku, a jeśli je włączał, to jedynie w formie ekstrapolowania dotychczasowych tendencji.

W latach 60. i 70. ten sposób modelowania został poddany krytyce przez wielu ekonomistów. Wskazywano na to, że nie odzwierciedla on prawidłowo rzeczywistości w sytuacji rosnącej inflacji, spowodowanej wzrastającym deficytem budżetowym w USA i skutkami szoków naftowych. Podkreślano, że w takich warunkach zachowania uczestników rynku określone są w dużym stopniu przez oczekiwania co do przyszłości, które uwzględniają zarówno sygnały płynące z rynku, jak i zapowiedzi działań podejmowanych przez władze publiczne. Przykładowo, oczekiwania dotyczące przyszłego poziomu płac, ustalanego w ramach zbiorowych negocjacji, w rosnącym stopniu zależały od oczekiwanej raczej, a nie minionej inflacji. Amerykański ekonomista John Muth jako pierwszy sformułował koncepcję racjonalnych oczekiwań (*rational expectations* – Muth 1961). Podkreślał, że nie ma powodu

[29] Innym przykładem opcjonalnej wartości „czekania" może być decyzja o wejściu Polski do strefy euro, mająca zasadniczo charakter nieodwracalny. Wprawdzie Polska ma prawno-międzynarodowy obowiązek przyjęcia wspólnej waluty, ale data akcesji do strefy euro nie została określona. Wobec kryzysu zadłużenia w niektórych państwach członkowskich i niepewności co do tego, jaki będzie ostateczny rezultat reform podejmowanych w strefie euro, Polsce opłaca się czekać z decyzją o akcesji do momentu, kiedy będzie już wiadomo, jaki ma być przyszły kształt unii walutowej.

[30] Przygotowany dla rządu brytyjskiego raport o skutkach ocieplenia klimatu (tzw. raport Sterna) wskazuje na to, że zwlekanie może wiązać się z koniecznością poniesienia znacznie większych kosztów w przyszłości (Stern 2007).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zakładać, że uczestnicy rynku zachowują się nieracjonalnie w sposób systematyczny. Tymczasem postępowaliby nieracjonalnie, gdyby w swoich decyzjach uwzględniali wyłącznie dotychczas zaobserwowane trendy oraz ignorowali informacje dotyczące przyszłości. W istocie uczestnicy rynku zachowują się racjonalnie, ponieważ biorą pod uwagę wszystkie informacje dostępne w czasie podejmowania decyzji, w tym informacje dotyczące zapowiedzianych lub oczekiwanych przyszłych działań rządu. W świetle tego spostrzeżenia polityka gospodarcza przestaje być działaniem „jednokierunkowym", polegającym na manipulowaniu instrumentami w ramach statycznej struktury decyzyjnej w celu oddziaływania na uczestników rynku, a staje się działaniem interaktywnym, w ramach którego uczestnicy rynku kształtują swoje zachowania w reakcji na działania władzy publicznej, co z kolei zmusza władzę publiczną do zmiany polityki. Formalnie rzecz biorąc, hipoteza racjonalnych oczekiwań zakłada, że projekcje te uwzględniają całą dostępną informację, a zatem błędy w przewidywaniach wynikają wyłącznie z czynnika losowego.

Hipoteza racjonalnych oczekiwań pozwoliła podważyć dotychczas przyjmowane założenie o ekstrapolacyjnym charakterze oczekiwań. Z całą pewnością można uznać, że ludzie wykorzystują dostępne im informacje w podejmowaniu decyzji, co oznacza dostosowywanie ich zachowań do przewidywanej polityki. Hipoteza ta na pewno sprawdza się w znacznym stopniu w odniesieniu do takich uczestników rynku jak banki, biura maklerskie czy fundusze inwestycyjne. Instytucje te inwestują znaczne zasoby środków w badania rynku i analizy ekonomiczne w celu uzyskania wiarygodnych prognoz kształtowania się przyszłych wartości takich zmiennych jak ceny, kursy czy stopy procentowe. Można też zakładać, że, negocjując podwyżki płac, pracownicy będą uwzględniać w swoich żądaniach płacowych przyszły wzrost cen lub podatków. Lecz w odniesieniu do zachowania przeciętnego konsumenta bądź pojedynczego przedsiębiorcy rodzą się wątpliwości. Właściwe wykorzystanie ogółu dostępnych informacji wiąże się z koniecznością ich zbierania, przetwarzania i analizowania, co wymaga określonych umiejętności, odpowiednich nakładów czasu i dużych kosztów. Trudno zakładać, że indywidualny konsument będzie w stanie uwzględniać prawidłowo w swoich decyzjach wszystkie dostępne informacje. Przykładowo, gospodarstwa domowe w ogromnej większości nie są w stanie samodzielnie ocenić wpływu polityki zwiększania deficytu budżetowego na poziom płac i zatrudnienia.

Biorąc pod uwagę te wątpliwości, ekonomiści są obecnie skłonni zakładać, że w rzeczywistości przeważająca część uczestników rynku formułuje swe oczekiwania, uwzględniając wprawdzie wiele dostępnych informacji, ale w sposób nie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w pełni racjonalny. Składową procesu kształtowania się oczekiwań stanowi również mechanizm adaptacyjny, który polega na częściowym dostosowaniu oczekiwań do bieżących informacji. Z kolei w najprostszych sytuacjach decyzyjnych lub w obliczu braku wiedzy o przyszłości wielu uczestników rynku kształtuje swoje oczekiwania co do przyszłości, stosując prostą ekstrapolację dotychczasowych trendów.

1.5.6. Zawodne państwo?

Omówione wyżej przypadki niesprawności państwa wskazują, że polityka gospodarcza – z natury rzeczy – nie jest wolna od błędów, a przekonanie co do tego, że pozwala ona zawsze w sposób skuteczny usuwać lub niwelować skutki niesprawności rynku, nie jest do końca uzasadnione. W przeciwieństwie do standardowego wizerunku władzy publicznej jako podmiotu bezstronnego, obiektywnego, efektywnego i racjonalnego (choć czasami opresyjnego) w rzeczywistości władza publiczna jest na ogół stronnicza, generuje nadmierne koszty i popełnia błędy. Podsumowując wcześniejsze spostrzeżenia, można wskazać na cztery główne ograniczenia w funkcjonowaniu władzy publicznej. Po pierwsze, władza publiczna nie jest wszechmocna – respektując wolności obywatelskie i zasady demokracji, musi liczyć się z oczekiwaniami społeczeństwa wyrażanymi w okresowych wyborach i przestrzegać obowiązującego prawa. Oznacza to, że władza publiczna musi działać w ramach ograniczonych prawem i na podstawie prawa oraz że powinna uwzględniać głos opinii publicznej. Nie należy oczywiście uznawać tego za oznakę „słabości" władzy, a raczej za naturalne ograniczenie swobody działania, z którym każda władza musi się liczyć. Po drugie, władza publiczna nie jest wszechwiedząca – zarówno z uwagi na niepełną wiedzę o aktualnym stanie rzeczywistości i niepewność co do przyszłości, jak i ze względu na niedostateczne kompetencje oraz niepełne kwalifikacje. O ile tę drugą słabość można częściowo przezwyciężyć dzięki pogłębieniu wiedzy i zdobyciu lepszego wykształcenia, o tyle pierwsza z nich ma charakter obiektywny – działanie odbywa się zawsze w niepewnym świecie. Po trzecie, władza publiczna nie jest ani bezstronna, ani obiektywna – mechanizmy wyboru publicznego wpływają na to, że władzę sprawują politycy reprezentujący interesy określonych grup społecznych, a biurokracja (urzędnicy) ma swoje własne partykularne cele, które realizuje często ze szkodą dla interesu ogólnospołecznego. Po czwarte wreszcie, władza publiczna nie jest ekonomicznie efektywna – kieruje się odmienną strukturą bodźców niż przedsiębiorcy. W szczególności zaś

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

władza publiczna nie musi kierować się kryterium zysku oraz troszczyć się o wartość majątku, a ponadto posiada też skłonność do przechwytywania wytworzonej wartości przy pomocy podatków. W efekcie władza publiczna wykazuje z reguły większą tendencję do marnotrawstwa zasobów i generowania nadmiernych kosztów administracyjnych.

Zarówno rozwój pozytywnej teorii polityki gospodarczej w okresie ostatniego półwiecza, jak i doświadczenia rozwojowe wielu krajów pozwoliły ekonomistom oraz politykom lepiej uświadomić sobie rolę i znaczenie ograniczeń polityki państwa w zakresie gospodarki. Zakończyła się era naiwnej wiary w interwencjonizm państwowy jako niezawodną metodę okiełznania sił rynkowych oraz budowy pomyślnego i sprawiedliwego ładu w gospodarce. Okazało się, że państwo bywa omylne, ociężałe, stronnicze, nieskuteczne i niewydolne. Współczesna teoria polityki gospodarczej uwzględnia to, że polityk podejmuje decyzję na podstawie niepełnej informacji, a struktura jego preferencji może odbiegać od struktury preferencji ogólnospołecznych. Zważa się tu również na to, że uczestnicy rynku nie reagują na politykę gospodarczą w sposób mechaniczny, lecz ucząc się i przewidując przyszłe działania władzy publicznej oraz innych uczestników rynku, a także odpowiednio dostosowując swoje postępowanie do oczekiwań. Teoria ta uwzględnia wreszcie rosnącą zależność między indywidualnymi graczami, między różnymi podmiotami władzy publicznej oraz między poszczególnymi krajami. Uświadomienie sobie, że władza publiczna działa w warunkach niepełnej informacji, funkcjonowania „oczekiwań" i postępującej globalizacji wywarło wielki wpływ na kształt współczesnej polityki gospodarczej. Pozostało jednak przekonanie, że niezależnie od rozmaitych niedoskonałości interwencja państwa jest w wielu sytuacjach pożądana, a częstokroć niezbędna. Nie ma w istocie wątpliwości co do tego, czy tolerować niesprawności rynku w obawie przed tym, że niesprawność państwa będzie jeszcze gorsza bądź na odwrót – godzić się na niedoskonały interwencjonizm, zakładając, że jest on i tak mniejszym złem niż zawodny mechanizm rynkowy. We wszystkich krajach rozwiniętych mechanizm rynkowy uzupełniany jest i wspomagany przez politykę gospodarczą, choć jej cele, zakres, sposób działania i skuteczność mogą się istotnie różnić w poszczególnych przypadkach.

Jak wspomniano wyżej, immanentne niedoskonałości interwencji państwa nie powinny prowadzić do przekonania, że należy w ogóle zrezygnować z polityki gospodarczej i zdać się na łaskę oraz niełaskę rynków. Wręcz przeciwnie, czynniki te powinny zachęcać do tego, aby usuwać niesprawności w zakresie funkcjonowania państwa poprzez pogłębianie wiedzy o tym zjawisku i budowanie instytucji

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

chroniących gospodarkę przed błędnymi decyzjami władzy publicznej. Okoliczności te powinny też zachęcać do tego, aby doskonalić konkurencję i mechanizm rynkowy, ponieważ ogranicza to skalę niesprawności rynku i zmniejsza zapotrzebowanie na interwencję państwa.

1.6. Rodzaje polityki gospodarczej. Próba klasyfikacji

Polityka gospodarcza jest pojęciem popularnym i zarazem bardzo pojemnym. Dlatego ogólna definicja polityki gospodarczej podana na początku rozdziału wymaga często uszczegółowienia i doprecyzowania. Potrzeba taka pojawia się, gdy przychodzi określić, jakiemu celowi ma służyć dana polityka, wskazać na instrument, jaki ma być zastosowany, czy zarysować obszar gospodarki, którego polityka ma dotyczyć. We wszystkich tych przypadkach do rzeczownika „polityka" dodaje się najczęściej różne przymiotniki, które mają określać cele, instrumenty bądź obszary (dziedziny) oddziaływania polityki. Z tego punktu widzenia można dokonywać rozmaitej klasyfikacji rodzajów polityki gospodarczej.

Do najbardziej popularnych należy – wykorzystany także w niniejszej książce – podział rodzajów polityki gospodarczej ze względu na główny *instrument zastosowania*. W ramach tego ujęcia wyróżnia się politykę fiskalno-budżetową, politykę pieniężną, politykę kursową i politykę makroostrożnościową. Przykładowo, w przypadku polityki fiskalno-budżetowej głównym instrumentem pozostają podatki i wydatki budżetowe, a w przypadku polityki pieniężnej jest nim podaż pieniądza. Te rodzaje polityki odnoszą się do zmiennych makroekonomicznych, a więc dotyczą całej gospodarki. Makroekonomiczny wymiar mają także rodzaje polityki klasyfikowane z punktu widzenia ogólnego *kierunku wpływu* wywieranego na gospodarkę. Z tej perspektywy można wyróżnić politykę ekspansywną, mającą pobudzać gospodarkę do szybszego wzrostu, politykę restrykcyjną, służącą spowalnianiu zbyt szybkiego wzrostu gospodarczego w celu przywrócenia równowagi, oraz politykę neutralną. Alternatywnie można klasyfikować politykę pod względem wykorzystywanych *instrumentów mikroekonomicznych*, takich jak np. polityka cenowa, polityka płacowa, polityka podatkowa czy polityka pomocy publicznej. Te ostatnie rodzaje polityki odnoszą się z reguły do wyodrębnionych fragmentów systemu regulacyjnego gospodarki.

Inny jeszcze podział stanowi klasyfikacja rodzajów polityki biorąca pod uwagę główne *cele*, jakie chce się osiągnąć. Z tego punktu widzenia można mówić

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

np. o polityce antyinflacyjnej, proeksportowej, stabilizacyjnej, antyrecesyjnej i zapewne o wielu innych. Z perspektywy *obszaru oddziaływania* można mieć do czynienia z polityką odnoszącą się do poszczególnych *rynków* (np. polityka zatrudnienia, inwestycyjna, walutowa czy polityka wobec inwestycji zagranicznych), lub do poszczególnych *sektorów* gospodarki (np. polityka przemysłowa, handlowa, rolna czy transportowa). Zarówno polityka rynkowa, jak i polityka sektorowa mają wymiar mikroekonomiczny. W tabeli 1.4 przedstawiono najczęściej stosowane klasyfikacje rodzajów polityki gospodarczej. Trzeba oczywiście od razu zaznaczyć, że wszystkie klasyfikacje zawsze mają w jakimś zakresie charakter umowny i częściowo nakładają się na siebie.

Tabela 1.4. Przykładowa klasyfikacja rodzajów polityki gospodarczej

Nr	Kryterium klasyfikacji	Makro lub mikro	Rodzaj polityki
1	cel	makroekonomiczna	polityka stabilizacyjna polityka antyinflacyjna polityka antyrecesyjna polityka proeksportowa polityka antyimportowa polityka wzrostu gospodarczego
2	wpływ na gospodarkę	makroekonomiczna	polityka ekspansywna polityka restrykcyjna polityka neutralna
3	instrument makroekonomiczny	makroekonomiczna	polityka fiskalno-budżetowa polityka pieniężna polityka kursowa polityka makroostrożnościowa
4	instrument mikroekonomiczny	mikroekonomiczna	polityka cenowa polityka dochodowa (wynagrodzeń) polityka podatkowa polityka pomocy publicznej polityka konkurencji polityka mikroostrożnościowa
5	rynek oddziaływania	mikroekonomiczna	polityka zatrudnienia polityka inwestycyjna polityka walutowa polityka wobec inwestycji zagranicznych
6	sektor gospodarki	mikroekonomiczna	polityka przemysłowa polityka rolna polityka handlowa polityka transportowa polityka budownictwa

Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W niniejszej książce nie będzie jednak mowy o wszystkich wymienionych rodzajach polityki gospodarczej. Zamierzenie autora jest o wiele skromniejsze. Jak zaznaczono już w *Przedmowie*, przedmiotem rozważań będzie przede wszystkim polityka gospodarcza na szczeblu *makroekonomicznym*, z uwzględnieniem podziału na politykę fiskalno-budżetową, pieniężną, kursową i makroostrożnościową. Autor zdaje sobie przy tym sprawę, że jest to wybór arbitralny i że w żaden sposób nie umniejsza on znaczenia polityki mikroekonomicznej, w tym w szczególności polityki sektorowej.

Literatura cytowana

Acocella N., *Zasady polityki gospodarczej*, Wydawnictwo Naukowe PWN, Warszawa 2002.

Acocella N., *Economic Policy in the Age of Globalisation*, Cambridge University Press, Cambridge 2005.

Arrow K.J., *A Difficulty in the Concept of Social Welfare*, "Journal of Political Economy" 1950, vol. 58(4).

Arrow K.J., *Social Choice and Individual Values*, Wiley, New York 1951.

Axelrod R., *The Evolution of Cooperation*, Basic Books, New York 1984.

Coase R.H., *The Problem of Social Cost*, "Journal of Law and Economics" 1960, vol. 3.

Benassy-Quéré A., Qeuré B., Jacquet P., Pisani-Ferry J., *Economic Policy. Theory and Practice*, Oxford University Press, New York 2010.

Brainard W.C., *Uncertainty and the Effectiveness of Policy*, "American Economic Review" 1967, vol. 57, no. 2.

Brennan G., Buchanan J., *Towards a Tax Constitution of Leviathan*, "Journal of Public Economics" 1977, no. 8.

Brennan G., Buchanan J., *The Power to Tax*, Cambridge University Press, Cambridge 1980.

Caplin A., Schotter A., *The Foundations of Positive and Normative Economics: A Handbook* (*Handbooks of Economic Methodology*), Oxford University Press, Oxford 2008.

Coles J.L., Hammond P.J., *Walrasian Equilibrium Without Survival: Existence, Efficiency and Remedial Policy*, w: *Choice, Welfare, Development: A Festschrift in Honour of Amartya Sen*, red. P. Pattanaik, K. Suzumura, Clarendon Press, Oxford 1995.

Friedman M., *Essays in Positive Economics*, Chicago 1953.

Greenwald B.C., Stiglitz J.E., *Externalities in Economies with Imperfect Information and Incomplete Markets*, "The Quarterly Journal of Economics" 1986, vol. 101, no. 2.

Gruber J., *Public Finance and Public Policy*, Worth Publishers, New York 2007.

IMF, *GEM: A New International Macroeconomic Model*, 2004, mimeo (www.imf.org).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Kaja J., *Polityka gospodarcza. Wstęp do teorii*, wydanie 4 rozszerzone, Oficyna Wydawnicza SGH, Warszawa 2014.

Keynes J.M., *The General Theory of Employment, Interest and Money*, Macmillan, London 1936.

Lindahl E., *Die Gerechtigkeit der Besteuerung*, Lund 1919 (wydanie w przekładzie na język angielski: *Just Taxation: A positive solution*, przeł. E. Henderson, Auburn 1958).

Lucas R.E. Jr., *Econometric policy evaluation: A critique*, "Carnegie-Rochester Conference Series on Public Policy" 1976, vol. 1.

Marczewski K., *Normatywne podejście do polityki gospodarczej*, w: *Polityka gospodarcza. Teoria i realia*, red. J. Stacewicz, Oficyna Wydawnicza SGH, Warszawa 2008.

Mundell R.A., *The Monetary Dynamics of International Adjustment under Fixed and Flexible Exchange Rates*, "The Quarterly Journal of Economics" 1960, no. 4.

Musgrave R.A., *The Voluntary Exchange Theory of Public Economy*, "The Quarterly Journal of Economics" 1939, vol. 53, no. 2.

Musgrave R.A., *The Theory of Public Finance: A Study in Public Economy*, McGraw-Hill, New York 1959.

Musgrave R.A., Musgrave P.B., *Public Finance in Theory and Practice*, McGraw-Hill, New York 1973.

Muth J.F., *Rational Expectations and the Theory of Price Movements*, "Econometrica" 1961, vol. 29, no. 3.

Niskanen W.A., *Bureaucracy and Representative Government*, Aldine Atherton, New York–Chicago 1971.

Niskanen W.A., *Bureaucracy: Servant or Master? – Lessons from America*, Institute of Economic Affairs, London 1973.

North D., *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge 1990.

North D., *Institutions*, "The Journal of Economic Perspectives" 1991, vol. 5, no. 1.

Polityka gospodarcza, red. M. Ćwikliński, wydanie 3, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk 2001.

Polityka gospodarcza, red. B. Winiarski, Wydawnictwo Naukowe PWN, Warszawa 2004.

Polityka gospodarcza. Teoria i realia, red. J. Stacewicz, Oficyna Wydawnicza SGH, Warszawa 2008.

Rosati D.K., *Wpływ dewaluacji na równowagę wewnętrzną i zewnętrzną*, referat wygłoszony na seminarium IKC HZ, Warszawa, wrzesień 1989.

Rosati D.K., *Polityka proeksportowa*, PWE, Warszawa 1990.

Rosati D.K., *Poland: glass half empty*, w: *Economic Transformation in Central Europe. A Progress Report*, red. R. Portes, CEPR, London 1993.

Rosati D.K., *Output decline during transition from plan to market: A reconsideration*, "The Economics of Transition" 1994, vol. 2, no. 4.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Samuelson P.A., *The Theory of Public Expenditure*, "Review of Economics and Statistics" 1954, vol. 36.
Sen A., *Collective Choice and Social Welfare*, wydanie 1, Holden-Day, San Francisco 1970.
Sims C., *Macroeconomics and Reality*, "Econometrica" 1980, no. 48.
Stern N., *The Economics of Climate Change. The Stern Review*, Cambridge University Press, Cambridge 2007.
Stiglitz J.E., *Ekonomia sektora publicznego*, Wydawnictwo Naukowe PWN, Warszawa 2004.
Sulmicki P., *Planowanie i polityka gospodarcza*, wydanie 2, PWE, Warszawa 1978.
Tullock G., *The Politics of Bureaucracy*, Public Affairs Press, Washington 1965.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozdział 2

POLITYKA GOSPODARCZA W MODELACH MAKROEKONOMICZNYCH

2.1. Wprowadzenie

Poziom produkcji i zatrudnienia w gospodarce rynkowej ulega zmianom pod wpływem cyklu koniunkturalnego i jednorazowych wstrząsów. Cykl koniunkturalny jest charakterystyczną cechą gospodarki rynkowej. Fluktuacje gospodarcze pociągają za sobą zmiany w zakresie poziomu produkcji i zatrudnienia, wywołujące szereg niekorzystnych skutków gospodarczych i społecznych. Cykliczne spadki popytu, a także jednorazowe negatywne wstrząsy popytowe i podażowe przekładają się na spadek produkcji, ograniczenie inwestycji i zmniejszenie zatrudnienia. Utrata pracy wśród części pracowników oznacza zanik dochodów bieżących i gwałtowne pogorszenie warunków życia. Wzrasta też niepewność co do przyszłości przy jednoczesnym spadku poczucia bezpieczeństwa ekonomicznego. Pogorszenie sytuacji materialnej wśród znacznej części pracowników oraz wzrost niepewności co do losu własnego i rodziny mogą prowadzić do zaburzeń społecznych i destabilizacji politycznej sytuacji kraju. W dłuższej perspektywie bezrobocie powoduje także utratę kwalifikacji przez część siły roboczej, ograniczenie szans na rozwój zawodowy, osłabienie spójności społecznej i nasilenie się rozmaitych patologii, takich jak rozpad rodzin i wzrost przestępczości.

Ryzyko wystąpienia tych niekorzystnych zjawisk sprawia, że jednym z głównych celów polityki gospodarczej jest stabilizowanie popytu w celu ograniczenia skali wahań poziomu produkcji i zatrudnienia, a zwłaszcza unikanie cyklicznych załamań i niwelowanie skutków wstrząsów popytowych i podażowych. W krótkim

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

okresie – który w przypadku polityki gospodarczej mieści się na ogół w horyzoncie od kilku tygodni do 2–3 lat – moce wytwórcze traktowane są jako dane. W tych warunkach na wahania produkcji najczęściej wpływają cykliczne i jednorazowe zmiany popytu. Mogą jednak one też wynikać z egzogenicznych wstrząsów podażowych, związanych z nagłymi zmianami warunków produkcji i podaży. Zadaniem polityki gospodarczej jest łagodzenie skutków tych wahań, tak aby zapewnić stabilizację poziomu produkcji i zatrudnienia. Dlatego też polityka gospodarcza, postrzegana w ten sposób, nazywana jest też polityką stabilizacyjną.

Skuteczna polityka stabilizacyjna wymaga dobrego rozpoznania zależności ekonomicznych w gospodarce. Jest to konieczne, jeśli oczekuje się, by instrumenty polityki gospodarczej wywierały pożądany i oczekiwany wpływ na kształtowanie się takich zmiennych makroekonomicznych jak produkcja, popyt, zatrudnienie, konsumpcja czy inwestycje. Aby ułatwić analizę stanu gospodarki i wypracować właściwe decyzje w zakresie polityki gospodarczej, zależności te ujmuje się najczęściej w postaci modeli makroekonomicznych. Składają się one z pewnej ograniczonej liczby równań i nierówności, które – w dopuszczalnym uproszczeniu – mają symulować rzeczywiste zachowanie się poszczególnych zmiennych. Wyróżnia się wśród nich zmienne egzogeniczne – czyli dane pochodzące z „zewnątrz" lub określone poza modelem – oraz zmienne endogeniczne – tj. zmienne wynikające z zależności opisanych modelem. Do zmiennych egzogenicznych zalicza się najczęściej instrumenty polityki gospodarczej, takie jak oficjalna stopa procentowa, podatki lub wydatki publiczne, będące przedmiotem decyzji podejmowanych przez władzę gospodarczą, podczas gdy zmienne endogeniczne dotyczą takich czynników, które polityka gospodarcza ma kształtować – jak produkcja, popyt czy zatrudnienie.

Ekonomiści stworzyli bardzo wiele różnorodnych modeli opisujących zależności makroekonomiczne występujące w gospodarce. Różnią się one stopniem szczegółowości, liczbą zmiennych, założeniami co do charakteru powiązań między zmiennymi, rolą czynników losowych oraz sposobem ich praktycznego wykorzystania. Zapewne najbardziej popularnym modelem makroekonomicznym jest znany ze standardowych podręczników makroekonomii model IS-LM, sformułowany po raz pierwszy przez brytyjskiego ekonomistę Johna Hicksa w 1937 roku[1]. Model ten, rozwijany i poszerzany w następnych latach przez innych badaczy, ukazuje główne zależności pomiędzy najważniejszymi zmiennymi makroekonomicznymi.

[1] Zob. Hicks (1937). Nazwa modelu nawiązuje do oznaczeń głównych zmiennych wykorzystanych w modelu: *I* – inwestycje, *S* – oszczędności, *L* – popyt na pieniądz, *M* – podaż pieniądza.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Zależności te sformułowane są z uwzględnieniem typowo keynesowskich założeń, mówiących o tym, że w gospodarce istnieją wolne moce wytwórcze, a ceny w krótkim okresie pozostają względnie sztywne. Oznacza to, że przedsiębiorstwa dostosowują się do zmian popytu przede wszystkim poprzez zmiany bieżącego poziomu produkcji. Model zaproponowany przez Hicksa mimo upływu lat nie traci na swej aktualności i jest nadal powszechnie wykorzystywany ze względu na prostotę ujęcia oraz łatwość intuicyjnej interpretacji. Przyjdzie nam z niego korzystać także w dalszych częściach książki[2].

2.2. Podstawowe zależności makroekonomiczne w modelu IS-LM

Na wstępie należy przypomnieć najbardziej elementarną postać modelu IS-LM, w której zakłada się, że gospodarka jest zamknięta (brak handlu zagranicznego i transgranicznych przepływów kapitałowych)[3] oraz przyjmuje się najprostsze specyfikacje poszczególnych równań. W takim najprostszym modelu występują cztery zmienne endogeniczne, a więc zmienne zależne od polityki gospodarczej: dochód (produkcja) Y, konsumpcja C, inwestycje I oraz realna stopa procentowa r[4], a także trzy zmienne egzogeniczne, tj. zmienne określone poza modelem, które mogą być one kształtowane pod wpływem decyzji podejmowanych przez władzę publiczną (a więc w ramach polityki gospodarczej). Należą do nich wydatki rządowe G, podatki pomniejszone o transfery budżetowe T oraz nominalna podaż

[2] Jak już podkreślono, główną zaletą modelu IS-LM jest prostota ujęcia oraz łatwość ekspozycji. Lecz walor ten pozostaje jednocześnie zasadniczą wadą modelu. Pozwala on oszacować warunki równowagi tylko na dwóch rynkach – na rynku produktów i na rynku aktywów (a ściślej rzecz biorąc, na rynku tylko jednego spośród aktywów – pieniądza). Założenie sztywnych cen powoduje, że strona podaży nie jest opisywana przez model. Nie zawiera on również optymalizacji i brak jest mikroekonomicznych podstaw zachowania uczestników rynku. Ponadto model ten jest całkowicie statyczny – nie uwzględnia zmian zachodzących w poziomie kapitału i bogactwa, a oczekiwania pozostają zupełnie inercyjne. W tym sensie model IS-LM może być słusznie traktowany jako nadmiernie uproszczony i – z punktu widzenia współczesnych osiągnięć nauk ekonomicznych – anachroniczny. Lecz mimo tych wad nadal pozostaje najlepszym i najłatwiejszym wprowadzeniem do makroekonomii (Rose 2000).

[3] Alternatywnie można przyjąć założenie, że handel zagraniczny jest zrównoważony, czyli $Ex = Im$.

[4] Stopa realna oznacza stopę procentową nominalną pomniejszoną o stopę oczekiwanej inflacji. Ponieważ w modelu zakłada się niezmienność cen, stopa realna jest w tym przypadku równa stopie nominalnej. Trzeba też zauważyć, że w gospodarce występuje wiele stóp procentowych dotyczących różnych aktywów finansowych i różnych okresów, a stopa procentowa r może być traktowana jako średnia ważona realna stopa w odniesieniu do całej gospodarki.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pieniądza M. Zgodnie z tradycją keynesowską, poziom cen P jest w tym przypadku stały (jako zmienna z góry ustalona). W swej najprostszej wersji model opisuje kształtowanie się podaży i popytu na dwóch zasadniczych rynkach – na rynku produktów (dóbr i usług) oraz na rynku pieniądza krajowego. Ma to służyć ukazaniu, w jakich warunkach – czyli w przypadku jakich wartości poszczególnych zmiennych – w gospodarce występuje równowaga[5].

2.2.1. Rynek produktów

Punktem wyjścia niech będzie przypomnienie znanych tożsamości makroekonomicznych. Wielkość produkcji w gospodarce, mierzona na przykład jako produkt krajowy brutto (PKB), jest z definicji równa sumie wartości wszystkich dochodów uzyskanych przez przedsiębiorców i pracowników w procesie wytwórczym (suma zysków brutto i płac brutto) w danym okresie (z reguły jeden rok). Wszystkie uzyskane w produkcji dochody zostają ostatecznie rozdysponowane na trzy możliwe sposoby – na opłacenie podatków, na konsumpcję lub na oszczędności (trzeba w tym miejscu dodać, że dochody te są uzupełniane z budżetu w ramach bezzwrotnych transferów, takich jak emerytury, stypendia lub zasiłki). Ten sam PKB jest jednocześnie równy wszystkim wydatkom na różne dobra i usługi, poniesionym przez wszystkie podmioty gospodarcze – przedsiębiorstwa, gospodarstwa domowe oraz instytucje sektora rządowego (należy przypomnieć, że chwilowo zakładamy, że gospodarka nie prowadzi wymiany z zagranicą). Opisane zależności można przedstawić następująco:

$$Y \equiv C + S + T. \quad \text{(tożsamość dochodu)} \tag{2.1}$$

Równanie (2.1) jest tożsamością, co oznacza, że wynika z przyjętej definicji. Dochód całkowity Y, czyli wartość wytworzonej produkcji dóbr i usług, jest zawsze równy sumie konsumpcji, oszczędności i zapłaconych podatków. W skali całej gospodarki dochód całkowity Y pozostaje równy wartości całej produkcji, a zatem reprezentuje produkt krajowy brutto (PKB).

$$D \equiv C + I + G. \quad \text{(tożsamość wydatków)} \tag{2.2}$$

[5] Pełne omówienie modelu IS-LM znaleźć można w standardowych podręcznikach makroekonomii (zob. m.in. Hall i Taylor 2000, s. 185–200; Barro 1997, s. 600–604; Begg, Fischer i Dornbusch 2003, s. 164–172; Burda i Wyplosz 2000, s. 307–315; Blanchard 2011, s. 127–146).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Równanie (2.2) jest tożsamością wydatków. Wskazuje ono na to, że łączna suma wydatków w skali całej gospodarki składa się z wydatków sektora prywatnego na konsumpcję C oraz na inwestycje I, a także z wydatków sektora rządowego (publicznego) na zakupy dóbr i usług G[6]. Możliwa jest też inna interpretacja. Równanie (2.2) dowodzi również, że zagregowany – czyli łączny – popyt efektywny w gospodarce jest zawsze równy sumie poszczególnych składników popytu finalnego – popytu konsumpcyjnego, popytu inwestycyjnego oraz popytu sektora rządowego.

Równanie (2.3) pokazuje, że warunkiem równowagi na rynku produktów jest równość dochodu, tj. wartości wytworzonej produkcji, oraz wszystkich wydatków:

$$Y = D. \quad \text{(warunek równowagi na rynku produktów)} \tag{2.3}$$

Taka równość wynika *ex post* z definicji. Jeśli zdefiniujemy jednak D jako zamierzone (planowane) wydatki (*ex ante*), a nie jako wydatki rzeczywiste (*ex post*), równość (2.3) na ogół nie zostanie spełniona, choć polityka gospodarcza będzie dążyć do tego, aby warunek ten był zrealizowany.

W modelu IS-LM konsumpcja jest funkcją dochodu – im wyższy staje się dochód, tym wyższa jest konsumpcja. Funkcję tę można zapisać w następujący sposób:

$$C = a + c(1 - t)Y, \quad \text{(funkcja konsumpcji)} \tag{2.4}$$

gdzie
$a > 0$, $c > 0$, $t < 1$.

Równanie (2.4) pokazuje, że konsumpcja jest rosnącą funkcją dochodu po opodatkowaniu „netto", a więc funkcją dochodu powstałego po odjęciu należnych podatków i uzupełnieniu przez transfery budżetowe. Taki dochód nosi miano dochodu rozporządzalnego lub dochodu do dyspozycji (*disposable income*). Współczynnik c nazywany jest skłonnością do konsumpcji (*propensity to consume*), a jego wartość wskazuje na to, jaką część każdej jednostki dochodu rozporządzalnego wydaje się na konsumpcję. Współczynnik t oznacza stopę opodatkowania dochodu po uwzględnieniu transferów. Współczynnik a w równaniu (2.4) pozostaje dodatni, co odzwierciedla realistyczne założenie, że nawet przy zerowym poziomie dochodu poziom konsumpcji będzie większy od zera, ponieważ ludzie muszą coś jeść i zaspokajać

[6] W skład wydatków sektora rządowego G nie wchodzą wydatki transferowe, na przykład świadczenia socjalne (emerytury, renty, zasiłki). Stanowią one część wydatków konsumpcyjnych C.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

inne najbardziej podstawowe potrzeby, aby utrzymać się przy życiu. W przypadku braku bieżącego dochodu potrzeby te finansowane są ze zgromadzonych oszczędności lub ze sprzedaży posiadanych aktywów.

W modelu IS-LM inwestycje stanowią funkcję stopy procentowej – im wyższa jest stopa procentowa, tym niższe stają się inwestycje. Funkcję inwestycji można więc zapisać następująco:

$$I = b + g\,r, \quad \text{(funkcja inwestycji)} \tag{2.5}$$

gdzie
$b > 0$, $g < 0$.

Równanie (2.5) pokazuje, że poziom inwestycji pozostaje ujemnie skorelowany z poziomem realnej stopy procentowej, co oznacza, że wzrost inwestycji jest proporcjonalny do spadku stopy procentowej (i odwrotnie). Współczynnik g jest więc ujemny i pozwala zmierzyć bezpośrednią zależność inwestycji od stopy procentowej. Z kolei współczynnik b reprezentuje w równaniu inne, niewyszczególnione w najprostszej wersji modelu (np. poziom ryzyka czy tzw. współczynnik q Tobina[7]), czynniki wpływające na poziom inwestycji. Równanie pokazuje, że nawet w przypadku zerowej stopy procentowej poziom inwestycji nigdy nie będzie nieskończenie duży i zostanie ograniczony do poziomu b.

Przedstawione równania pozwalają zdefiniować warunki równowagi popytu i podaży w gospodarce. Rynek produktów, opisany równaniami (2.1) ÷ (2.2) i (2.4) ÷ (2.5), znajdzie się w równowadze, gdy planowane wydatki (czyli wydatki *ex ante*) będą dokładnie równe wytworzonej produkcji. W przeciwnym razie można mieć do czynienia albo z nadwyżką produkcji, albo z niedoborem produkcji w stosunku do popytu. W pierwszym przypadku nadwyżka produkcji przełoży się na przyrost zapasów i w rachunkach narodowych (*national accounts*) zostanie zarejestrowana jako część inwestycji brutto (akumulacji brutto). W drugim przypadku nastąpi spadek zapasów. Tak więc wydatki *ex post* zawsze pozostaną równe wytworzonej produkcji, przy czym element dostosowawczy będą stanowiły zmiany poziomu zapasów.

Ponieważ całkowita produkcja wynosi Y, to warunek równowagi na rynku produktów (dóbr i usług) wymaga, aby *ex ante* spełnione było następujące równanie:

[7] Współczynnik q Tobina jest stosunkiem wartości rynkowej przedsiębiorstwa do jego wartości księgowej (odtworzeniowej). Gdy osiąga on wartość większą od jedności, opłaca się inwestować w majątek produkcyjny.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$Y = C + I + G, \tag{2.6}$$

czyli:

$$Y = [a + c(1 - t)Y] + [b - g\,r] + G. \tag{2.6a}$$

Równanie to nie jest już tożsamością, a warunkiem równowagi między podażą i popytem. Warunek równowagi na rynku produktów (równanie 2.6) wygodnie jest zilustrować graficznie. W górnej części rysunku 2.1 na osi poziomej odłożono produkcję (dochód) Y, a na osi pionowej planowane wydatki (popyt) D. Prosta $D_1 = D(Y, r_1)$ przedstawia wydatki jako funkcję dochodu w przypadku danej stopy procentowej r_1. Linia ta przecina oś pionową powyżej punktu zero, ponieważ konsumpcja (a także wydatki rządowe) przyjmuje wartość dodatnią dla $Y = 0$. Linia o nachyleniu 45° odzwierciedla punkty, dla których $Y = D$. W punkcie przecięcia prostej D_1 i linii 45° (punkt A) planowane wydatki są równe produkcji i warunek równowagi (2.6) pozostaje spełniony. Spadek stóp procentowych z r_1 do r_2 powoduje wzrost popytu (inwestycyjnego) przy danym poziomie dochodu i w efekcie prosta wydatków przesuwa się w górę do położenia $D_2 = D(Y, r_2)$. Nowym punktem równowagi staje się punkt B, zapewniający wyższy poziom produkcji i dochodu Y_2.

Wykorzystując równanie (2.1) i podstawiając (2.2), otrzymujemy alternatywną postać warunku równowagi na rynku produktów:

$$C + S + T = C + I + G$$

$$S = I + (G - T). \tag{2.7}$$

Gdy założymy dla uproszczenia, że wydatki państwowe są równe dochodom podatkowym po potrąceniu płatności transferowych (czyli $G - T = 0$), otrzymamy najprostszą wersję warunku równowagi (wyjaśniającą symbolikę zastosowaną po raz pierwszy przez Hicksa):

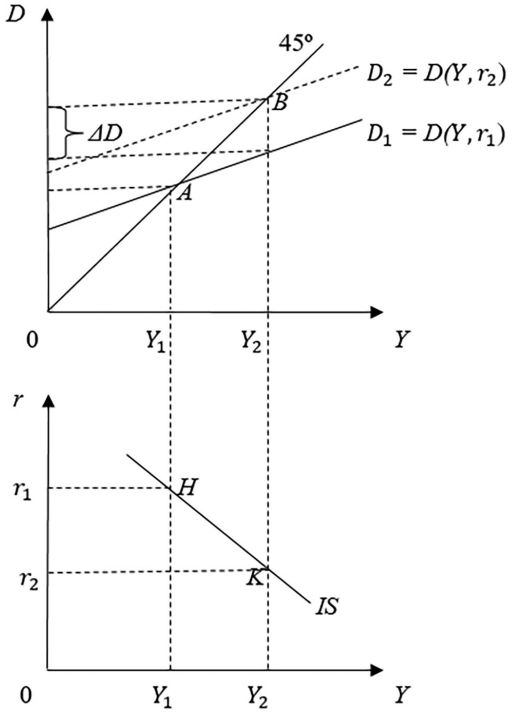
$$I = S. \tag{2.8}$$

Równanie (2.8) dowodzi, że warunkiem równowagi na rynku produktów w przepadku gospodarki zamkniętej jest równość inwestycji i oszczędności. Zauważmy, że równowaga na rynku produktów może być spełniona dla różnych kombinacji wartości zmiennych Y i r. Zależność tę można przedstawić graficznie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w układzie współrzędnych Y i r, czyli w przestrzeni dwóch zmiennych endogenicznych, na które oddziałuje polityka gospodarcza. Wyprowadzenie krzywej *IS* zostało ukazane na rysunku 2.1.

Rysunek 2.1. Graficzne wyprowadzenie krzywej IS



Źródło: opracowanie własne.

Początkowy stan równowagi, opisany punktem A w górnej części rysunku 2.1, znajduje się w punkcie przecięcia krzywej wydatków D i linii o nachyleniu 45°. Punkt ten wyznacza poziom produkcji zapewniający równowagę podaży i popytu na rynku produktów, przy danym poziomie stóp procentowych r_1. Punktowi temu odpowiada punkt H w dolnej części rysunku 2.1. Załóżmy teraz, że z jakiegoś powodu następuje spadek stóp procentowych z poziomu r_1 do poziomu r_2, pociągający za sobą wzrost popytu i wydatków łącznie o wielkość równą ΔD. Zmiana ta wpływa na przesunięcie krzywej popytu z D_1 do D_2, a nowym punktem równowagi staje się B. Utrzymanie równowagi wymaga zwiększenia produkcji z Y_1 do Y_2 – przy czym należy zauważyć, że łączny wzrost produkcji ($Y_2 - Y_1$) jest większy niż początkowy

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wzrost wydatków ΔD ze względu na działanie mnożnika. Nowemu punktowi równowagi B odpowiada punkt K znajdujący się w dolnej części rysunku 2.1. Łącząc punkty H i K, otrzymujemy poszukiwaną krzywą IS.

Krzywa IS odzwierciedla wszystkie kombinacje r i Y, spełniające tożsamość dochodu (2.1) oraz równanie wydatków (2.7), a zatem staje się reprezentacją punktów równowagi na rynku produktów. Interpretacja ekonomiczna jest w tym przypadku następująca. Dla danego poziomu produkcji (dochodu) Y istnieje tylko jedna wartość stopy procentowej r, zapewniająca poziom wydatków D równy poziomowi produkcji Y. Gdy produkcja z jakiegoś powodu wzrośnie, stopa procentowa musi ulec obniżeniu, aby umożliwić odpowiedni wzrost popytu i wydatków w celu zachowania równowagi. Z modelu wynika, że tym składnikiem popytu, który wzrośnie w wyniku spadku stóp procentowych, są inwestycje – tylko one zależą bowiem od zmian stopy procentowej[8].

Ze względu na odwrotnie proporcjonalną zależność inwestycji i stopy procentowej, krzywa IS ma nachylenie ujemne. Wynika to bezpośrednio z zapisu równania (2.6a), które dowodzi, że wzrostowi Y musi towarzyszyć spadek r, jeśli równowaga na rynku produktów ma być utrzymana. Stopień nachylenia krzywej IS zależy zaś od wrażliwości inwestycji (oraz ewentualnie innych składników popytu finalnego) na zmiany stopy procentowej – im silniej inwestycje reagują na zmiany stóp, tym mniejszy spadek ich poziomu jest niezbędny w celu uzyskania danego wzrostu popytu inwestycyjnego i tym bardziej płaska staje się krzywa IS.

Nachylenie krzywej IS można obliczyć, różniczkując całkowicie równanie (2.6) przy założeniu, że $dG = 0$:

$$dY = dC + dY = cdY + dr(\delta I/\delta r),$$

co po przekształceniu daje:

$$dr/dY = (1 - c)/\delta I/\delta r. \tag{2.9}$$

[8] Należy pamiętać, że jest to uproszczenie. Poziom stopy procentowej ma także wpływ na wydatki konsumpcyjne (poprzez wpływ na poziom oprocentowania kredytu konsumpcyjnego oraz na wartość składników majątku). Zależności te zostały jednak w modelu pominięte, aby nie zaciemniać głównego przesłania. Nie zmienia to zresztą wniosków, ponieważ wpływ stopy procentowej na konsumpcję (a także na eksport netto) pozostaje taki sam co do kierunku jak w przypadku inwestycji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ponieważ licznik po prawej stronie równania (2.9) jest zawsze dodatni (ze względu na to, że $c < 1$), a mianownik z założenia pozostaje ujemny, nachylenie krzywej *IS* będzie zawsze ujemne. Wielkość wyrażenia (2.9) określa stopień nachylenia krzywej *IS*. Łatwo zauważyć, że krzywa *IS* będzie bardziej stroma w przypadku małych wartości *c* oraz w przypadku małej wrażliwości inwestycji na zmiany stopy procentowej ($\delta I/\delta r$).

Położenie krzywej *IS* na płaszczyźnie (Y, r) jest określone także przez wartości pozostałych zmiennych egzogenicznych występujących w równaniu popytu, takich jak *G* czy *X*. Zmiany wartości tych zmiennych, wywołane decyzjami władzy gospodarczej, zmianami preferencji konsumentów czy szokami zewnętrznymi, będą powodować przesunięcia krzywej *IS*. Przykładowo, autonomiczny wzrost wydatków rządowych *G* przesuwa krzywą *IS* w górę i w prawo, ponieważ przy danym poziomie produkcji zachowanie równowagi na rynku produktów wymaga, aby wzrost popytu rządowego był skompensowany spadkiem pozostałych składników popytu finalnego. Oznacza to, że przy wybranym poziomie produkcji stopa procentowa musi wzrosnąć, aby spowodować spadek inwestycji o określoną wielkość. Wzrost *G* (przy danych podatkach) zmniejsza oszczędności krajowe *S*. W związku z tym, aby nadal spełniony był warunek (2.6) lub (2.7), inwestycje muszą odnotować spadek o dokładnie taką samą wielkość. Przy danym poziomie *Y* wyższa stopa procentowa jest potrzebna w celu ograniczenia innych wydatków (np. popytu konsumpcyjnego).

2.2.2. Rynek pieniądza

Drugim rynkiem opisywanym w modelu IS-LM jest rynek pieniądza krajowego. Popyt na pieniądz jest tym większy, im bardziej wzrasta dochód realny *Y* i im wyższy jest poziom cen *P*, oraz tym mniejszy – im wyższa staje się stopa procentowa *r*. W najprostszym ujęciu funkcja popytu na pieniądz przyjmuje więc następującą postać:

$$L = (hY - k\,r)P \quad \text{(funkcja popytu na pieniądz)}, \tag{2.10}$$

gdzie $h = \dfrac{\delta L}{\delta Y} > 0$ mierzy wrażliwość popytu na pieniądz względem zmian dochodu, podczas gdy $k = \dfrac{\delta L}{\delta r} > 0$ mierzy wrażliwość popytu na pieniądz względem zmian stopy procentowej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Równanie (2.10) stanowi funkcję popytu na pieniądz nominalny. W ramach modelu IS-LM pieniądz definiowany jest w wąskim znaczeniu jako zasób pieniądza utrzymywany przez uczestników rynku w celu dokonywania bieżących transakcji zakupu dóbr i usług, a także dla wygody, czyli na wypadek konieczności lub chęci zaspokojenia nieprzewidzianych potrzeb. Taka definicja pieniądza odnosi się do najbardziej płynnych elementów całego obiegu pieniężnego i nie obejmuje pozostałych składników szerszych agregatów pieniężnych, np. pieniężnych lokat terminowych zakładanych w bankach. Popyt na tak wąsko zdefiniowany pieniądz pozostaje dodatnio skorelowany z poziomem dochodu, ponieważ wzrost dochodu pozwala zwiększyć wydatki, podnosząc tym samym zapotrzebowanie na pieniądz transakcyjny. Popyt na pieniądz jest także skorelowany ujemnie z poziomem stopy procentowej, ponieważ jej wzrost przekłada się na wzrost kosztu alternatywnego trzymania pieniądza w postaci gotówki (lub wkładów na żądanie), ponieważ zwiększa się wówczas atrakcyjność lokat bankowych wskutek wzrostu oprocentowania. Pierwszy element zapisany w nawiasie w równaniu (2.10) reprezentuje popyt transakcyjny na pieniądz, podczas gdy drugi z nich odzwierciedla popyt „spekulacyjny" na pieniądz[9]. Równanie (2.10) pokazuje, że popyt na pieniądz nominalny zależy także od poziomu cen P.

Utrzymanie równowagi na rynku pieniądza wymaga, aby nominalna podaż pieniądza M była równa popytowi na pieniądz, opisanemu równaniem (2.10).

$$M = L = (hY - k\,r)P. \tag{2.11a}$$

Podczas gdy w kategoriach realnych (i z uwzględnieniem założenia o niezmienności cen P) równanie przyjmuje następującą postać:

$$M/P = L/P = (hY - k\,r). \tag{2.11b}$$

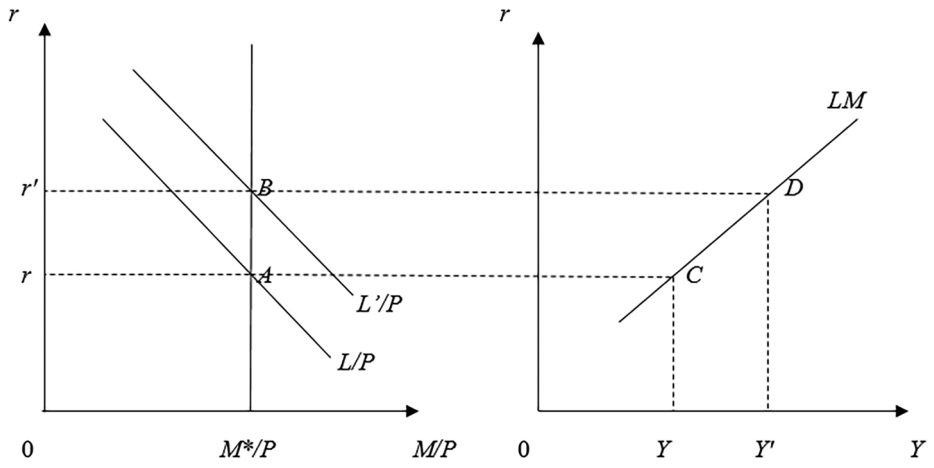
Warunek równowagi na rynku pieniądza można przedstawić graficznie w układzie współrzędnych Y i r (prawa strona rysunku 2.2). Linia LM reprezentuje kombinacje r i Y, które spełniają równanie (2.11a), a więc zapewniają równowagę na rynku

[9] Termin „spekulacyjny" w ekonomii nie ma wydźwięku pejoratywnego, związanego niekiedy z potocznym rozumieniem tego słowa. W opisywanym przypadku oznacza tę część popytu na pieniądz, która wynika z wygody posiadania pieniądza w formie najbardziej płynnej – gotówki i wkładów na żądanie. W zamian za to udogodnienie uczestnicy rynku gotowi są zrezygnować z dochodu, jaki mogliby uzyskać, lokując pieniądz w banku na oprocentowanej lokacie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pieniądza przy danej podaży pieniądza *M* i ustalonym poziomie cen *P*. Wyprowadzenie krzywej *LM* jest proste. Lewa strona rysunku 2.2 ukazuje sytuację na rynku pieniądza. Przy danym poziomie podaży pieniądza realnego M^*/P równowaga na rynku pieniądza ustali się w punkcie przecięcia prostej M^*/P z krzywą popytu na pieniądz realny $L/P = (hY - kr)$ – jest to punkt *A*. Krzywa popytu ma charakter opadający, ponieważ spadek stóp procentowych powoduje wzrost popytu na pieniądz przy danym poziomie dochodu. Załóżmy teraz, że z jakiegoś powodu dochód *Y* wzrasta – na przykład wskutek wzrostu PKB. Wówczas krzywa popytu na pieniądz przesunie się do góry do poziomu L'/P. Utrzymanie równowagi na rynku pieniądza przy jego niezmienionej podaży wymaga w tej sytuacji wzrostu stopy procentowej *r*, a nowym punktem równowagi staje się punkt *B*. Przenosząc na prawą stronę rysunku 2.2 odpowiednie wartości *Y* i *r*, dla których zachodzi równowaga na rynku pieniądza, otrzymujemy punkty *C* i *D*, które wyznaczają rosnącą krzywą *LM*.

Rysunek 2.2. Graficzne wyprowadzenie krzywej LM



Źródło: opracowanie własne.

Interpretacja ekonomiczna jest w tym przypadku następująca. Wzrost poziomu produkcji *Y* przekłada się na wzrost zapotrzebowania na pieniądz w celu obsłużenia większej liczby transakcji, a zatem krzywa popytu na pieniądz realny w lewej części rysunku 2.2 przesuwa się w prawo i w górę. Ponieważ ilość pieniądza *M* jest jednak dana i nie ulega zmianie, łączny popyt na pieniądz musi pozostać taki sam. Wymaga to wzrostu stóp procentowych z *r* do *r'*, aby skłonić uczestników rynku do

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

obniżenia spekulacyjnego popytu na pieniądz, czyli do trzymania mniejszej ilości pieniądza w postaci gotówki i wkładów na żądanie. Wzrost popytu transakcyjnego na pieniądz zostanie w ten sposób skompensowany spadkiem popytu spekulacyjnego, a łączny popyt nie ulegnie zmianie.

Ze względu na to, że wzrost produkcji (dochodu) Y wymaga wzrostu stopy procentowej r w celu utrzymania równowagi na rynku pieniądza, nachylenie krzywej LM pozostaje dodatnie. Wynika to bezpośrednio z zapisu równania 2.11b. Gdy realna podaż pieniądza M/P jest stała, prawa strona równania nie zmienia się, co to oznacza, że wzrostowi Y musi towarzyszyć wzrost r. Stopień nachylenia zależy zaś od wzajemnego położenia współczynników h i k, występujących w równaniu popytu na pieniądz (2.7). Przy danym poziomie k im bardziej wrażliwy jest popyt transakcyjny na pieniądz na zmiany poziomu Y, tym bardziej stroma staje się krzywa LM, ponieważ nawet niewielkie zmiany Y wymagają dużych modyfikacji r, aby zapewnić utrzymanie równowagi. Również odwrotnie, im bardziej wrażliwy staje się popyt spekulacyjny na pieniądz przy danym poziomie h, tym bardziej płaska będzie krzywa LM, ponieważ nawet niewielkie zmiany r wymagają dużych przekształceń na poziomie Y, aby zapewnić utrzymanie równowagi.

Nachylenie krzywej LM obliczamy, różniczkując równanie (2.11a), przy założeniu, że $dM = 0$ i $dP = 0$:

$$0 = hdY - kdr.$$

Daje to po przekształceniu:

$$dr/dY = h/k = \frac{\delta L}{\delta Y} \Big/ \frac{\delta L}{\delta r}. \tag{2.12}$$

Ponieważ współczynniki h i k pozostają zawsze dodatnie, prawa strona równania (2.12) jest większa od zera i krzywa LM ma stale nachylenie dodatnie. Krzywa LM będzie bardziej stroma dla wysokich wartości h, a zatem w przypadku dużej wrażliwości popytu transakcyjnego na pieniądz na zmiany poziomu Y, oraz wobec małych wartości k, czyli przy niewielkiej wrażliwości popytu spekulacyjnego na pieniądz na zmiany poziomu r.

Mimo że operujemy tu pojęciem „popytu na pieniądz", krzywa LM odzwierciedla w istocie zmiany w zasobach, a nie w strumieniach, jak miało to miejsce w przypadku krzywej IS. Zmiany w popycie na pieniądz oznaczają bowiem, że

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

następuje zmiana w strukturze portfela aktywów finansowych posiadanych przez uczestników rynku. Dany poziom popytu na pieniądz decyduje o tym, jaka część portfela utrzymywana jest w aktywach nieprzynoszących dochodu odsetkowego, a jaką część przechowuje się w pozostałych aktywach, takich jak obligacje czy lokaty terminowe.

Położenie krzywej *LM* określone jest także przez wartości pozostałych zmiennych egzogenicznych, typu *M* i *P*. Gdy rośnie podaż pieniądza *M*, krzywa *LM* przesuwa się w prawo i w dół, ponieważ przy tym samym poziomie produkcji *Y* zachowanie równowagi wymaga spadku stóp procentowych. Zachęca to uczestników rynku do trzymania większej ilości pieniądza. Gdy z kolei wzrasta poziom cen *P*, krzywa *LM* przesuwa się w lewo i w górę, ponieważ spada ilość pieniądza w kategoriach realnych, a przy tym samym poziomie produkcji *Y* stopy procentowe muszą być wyższe, aby zmniejszyć popyt spekulacyjny na pieniądz.

2.2.3. Równowaga ogólna i obszary nierównowagi

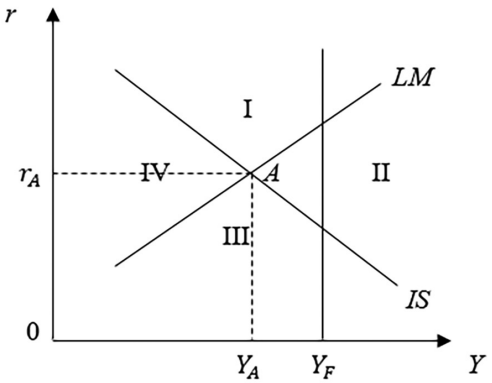
Wyprowadzenie krzywej *IS* oraz krzywej *LM* w ramach jednego układu współrzędnych pozwala wyznaczyć takie poziomy produkcji *Y* i stóp procentowych *r*, które, potraktowane łącznie, zapewniają jednocześnie równowagę na rynku produktów i na rynku pieniądza. Ponieważ krzywa *IS* jest nachylona ujemnie, a krzywa *LM* dodatnio, istnieje tylko jeden punkt przecięcia się obu krzywych. Oznacza to, że istnieje jedna i tylko jedna kombinacja wartości *Y* i *r*, zapewniająca jednocześnie równowagę na rynku produktów i na rynku pieniądza. Tę jedyną kombinację reprezentuje punkt *A* znajdujący się na rysunku 2.3.

Stan opisany punktem *A* niełatwo jest osiągnąć. Na ogół gospodarka znajduje się poza tym punktem i poszczególne rynki mają problem z utrzymaniem równowagi. Rysunek 2.3 pozwala zidentyfikować cztery różne typy nierównowagi w zależności od tego, w którym punkcie wykresu znajduje się w danym momencie gospodarka. Obszar I reprezentuje punkty położone powyżej krzywej *IS* i jednocześnie powyżej krzywej *LM*. Punkty powyżej krzywej *IS* charakteryzują się zbyt wysoką stopą procentową w stosunku do danego poziomu produkcji lub – alternatywnie – zbyt wysokim poziomem produkcji w stosunku do danego poziomu stopy procentowej. Oznacza to, że w punktach tych występuje przewaga podaży nad popytem na rynku produktów. Punkty położone powyżej krzywej *LM* charakteryzują się zbyt wysoką stopą procentową w stosunku do danego poziomu produkcji przy danej podaży pieniądza. Oznacza to, że podaż pieniądza jest za wysoka w odniesieniu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

do popytu na pieniądz. Nadwyżka podaży pieniądza spowoduje z kolei wzrost cen pozostałych aktywów (takich jak nieruchomości, obligacje, akcje), ponieważ uczestnicy rynku będą chcieli zrównoważyć swoje portfele majątkowe, pozbywając się nadwyżek pieniężnych i kupując inne aktywa. Gospodarka, która znajdzie się w punkcie położonym w obszarze I, będzie więc charakteryzować się z jednej strony tendencjami recesyjnymi w zakresie produkcji, spowodowanymi niedoborem popytu na produkty, z drugiej zaś strony wzrostem cen pozostałych aktywów wskutek pojawienia się nadwyżki podaży pieniądza nad popytem.

Rysunek 2.3. Równowaga i nierównowaga w gospodarce



Źródło: opracowanie własne.

Z kolei w obszarze II znajdują się punkty, które leżą powyżej krzywej *IS*, lecz jednocześnie poniżej krzywej *LM*. W punktach tych nadal utrzymuje się przewaga podaży nad popytem na produkty, ale tym razem stopy procentowe są za niskie przy danym poziomie produkcji i podaży pieniądza, a więc występuje przewaga popytu na pieniądz nad jego podażą. Gospodarka, która znajduje się w punkcie położonym w obszarze II, będzie więc charakteryzować się nadal tendencjami recesyjnymi wskutek niedoboru popytu na produkty, ale tym razem wobec niedoboru podaży pieniądza w stosunku do popytu nastąpi również spadek cen pozostałych aktywów.

W podobny sposób można przeanalizować stan gospodarki w obszarach III i IV. Podsumowanie tej analizy zostało przedstawione w tabeli 2.1.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 2.1. Typy nierównowagi gospodarczej w modelu IS-LM

Obszar nierównowagi	Rynek produktów	Rynek pieniądza	Stan gospodarki
I	podaż > popyt	podaż > popyt	recesja, wzrost cen aktywów
II	podaż > popyt	podaż < popyt	recesja, spadek cen aktywów
III	podaż < popyt	podaż < popyt	inflacja, spadek cen aktywów
IV	podaż < popyt	podaż > popyt	inflacja, wzrost cen aktywów

Źródło: opracowanie własne.

2.3. Model IS-LM z uwzględnieniem rynku pracy

W przedstawionej, najprostszej wersji model IS-LM ukazuje sytuację jedynie na dwóch rynkach – produktów i pieniądza krajowego. Nie jest to oczywiście kompletny obraz stanu gospodarki, ponieważ w wersji tej nie zostały uwzględnione dwa inne najważniejsze rynki – rynek pracy i rynek pieniądza zagranicznego (inaczej rynek walutowy). Ponadto, ponieważ model IS-LM jest typowo keynesowskim modelem popytowym, zakłada się w nim stałość cen i doskonale elastyczną podaż. W rzeczywistości ceny ulegają jednak zmianom, choć powoli, co ma wpływ na podaż i popyt. W dalszej części rozdziału założenia te będziemy stopniowo uchylać, uzupełniając model o kolejne warunki równowagi.

Przychodzi zacząć od omówienia sytuacji na rynku pracy. W standardowej analizie ekonomicznej przyjmuje się najczęściej, że podaż pracy oferowana przez pracowników jest rosnącą funkcją płac, a popyt na pracę zgłaszany przez przedsiębiorców jest malejącą funkcją płac (realnych). Jeśli płace są giętkie – czyli zmieniają się w krótkim okresie, reagując na zmiany podaży i popytu – równowaga na rynku pracy ustali się w punkcie przecięcia krzywych podaży i popytu na pracę, wyznaczając jednocześnie poziom płac oraz poziom zatrudnienia zapewniający „oczyszczenie" rynku pracy z siły roboczej dostępnej przy danym poziomie płac. Jest to więc stan pełnego zatrudnienia[10]. Z kolei na jego podstawie można wyznaczyć – za pośrednictwem funkcji produkcji – poziom produkcji odpowiadający pełnemu zatrudnieniu.

[10] Nie oznacza to oczywiście braku bezrobocia. Stan pełnego zatrudnienia należy rozumieć jako taki udział zatrudnienia, gdzie wszyscy pracownicy, chcący pracować przy *danym* poziomie wynagrodzeń, mogą znaleźć pracę. W gospodarce występuje więc jedynie bezrobocie dobrowolne i bezrobocie frykcyjne. Taka stopa bezrobocia jest równa stopie naturalnej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Zapiszmy równania popytu na pracę D_Z i podaży pracy S_Z jako dwie różne funkcje płacy realnej w:

$$D_Z = a - bw \tag{2.13}$$

$$S_Z = c + dw. \tag{2.14}$$

Warunkiem równowagi na rynku pracy jest, aby:

$$D_Z = S_Z. \tag{2.15}$$

Podstawiając równania (2.13) i (2.14) do (2.15) oraz przekształcając je, otrzymujemy wielkość zatrudnienia na poziomie zapewniającym równowagę podaży pracy i popytu na pracę:

$$Z_F = D_Z = S_Z = (ad + bc)/(b + d), \tag{2.16}$$

gdzie Z_F oznacza poziom pełnego zatrudnienia. Funkcja produkcji (2.17) pokazuje, że wielkość produkcji jest rosnącą funkcją wielkości zatrudnienia Z:

$$Y = f(Z), \quad \text{przy czym:} \quad f' > 0. \tag{2.17}$$

Wielkość produkcji, odpowiadająca pełnemu zatrudnieniu, wynosi natomiast:

$$Y_F = f(Z_F). \tag{2.18}$$

Poziom produkcji Y przy pełnym zatrudnieniu wyznaczony jest na rysunku 2.3 jako pionowa linia Y_F. Trudno jest *a priori* przewidzieć, czy prosta Y_F przejdzie przez punkt A, czy znajdzie się na lewo, czy też na prawo od tego punktu. Zauważmy, że poziom produkcji w równaniu (2.16) zależy wyłącznie od poziomu zatrudnienia, nie jest natomiast powiązany z innymi zmiennymi, takimi jak wydatki rządowe czy stopa procentowa. W tej sytuacji nie ma żadnego powodu, aby poziom produkcji, odpowiadający pełnemu zatrudnieniu Y_F, pozostawał koniecznie równy poziomowi produkcji Y_A, odpowiadającemu równowadze na rynku produktów i pieniądza. Pokazuje to rysunek 2.3, gdzie przyjęto, że poziom produkcji Y_F jest wyższy niż poziom Y_A. W tym przypadku w gospodarce wystąpi bezrobocie. Tylko

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

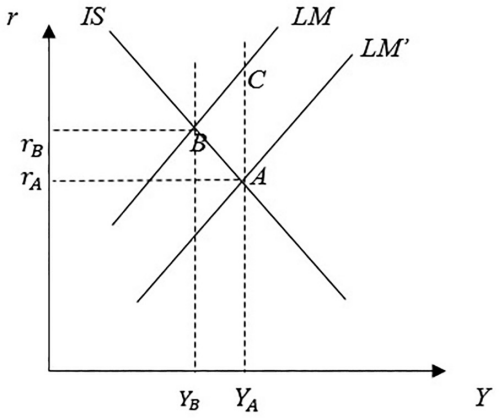
wówczas, gdy trzy linie – *IS, LM*, Y_F – przetną się w jednym punkcie, gospodarka osiągnie stan równowagi ogólnej.

2.4. Przywracanie równowagi i rola pieniądza

2.4.1. Przywracanie równowagi przy cenach giętkich i cenach sztywnych („lepkich")

Co dzieje się, gdy gospodarka znajduje się poza punktem równowagi ogólnej? Czy występuje jakiś samoczynny mechanizm, który odpowiada za automatyczny powrót do stanu równowagi? Czy raczej do akcji musi wkroczyć władza gospodarcza, która przy pomocy określonych instrumentów polityki gospodarczej wywoła pożądany proces dostosowawczy i w ten sposób przywróci równowagę? Odpowiedź zależy przede wszystkim od przyjętych założeń co do funkcjonowania mechanizmu cenowego. W ramach ekonomii klasycznej i neoklasycznej przyjmuje się, że ceny są doskonale giętkie (*flexible prices*). Oznacza to, że w przypadku wystąpienia nierównowagi pomiędzy podażą i popytem ceny natychmiast dostosowują się do zmian i doprowadzają do równowagi między podażą a popytem. Jeśli podaż przewyższa popyt, ceny spadają, a gdy popyt góruje nad podażą, ceny rosną.

Rysunek 2.4. Przywracanie równowagi w warunkach cen giętkich i sztywnych („lepkich")



Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Załóżmy, że poziom produkcji zapewniający pełne zatrudnienie w gospodarce jest równy Y_A, ale krzywe *IS* i *LM* przecinają się w punkcie *B*, co sytuuje stopy procentowe na poziomie r_B (rysunek 2.4). Oznacza to, że produkcja zapewniająca pełne zatrudnienie przewyższa popyt. Zgodnie z klasycznym założeniem o giętkości cen, przewaga podaży nad popytem powoduje spadek cen, a spadek cen, przy danej nominalnej podaży pieniądza, przekłada się na wzrost jego realnej podaży. Ten właśnie wzrost realnej podaży pieniądza wywołuje natychmiastowe przesunięcie krzywej *LM* w prawo do położenia *LM'*, co pociąga za sobą spadek stóp procentowych do poziomu r_A. Popyt na produkty rośnie i dostosowuje się do podaży, reagując na zmiany cen. W efekcie gospodarka stabilizuje się w punkcie równowagi ogólnej *A*. Przyjęcie założenia o giętkich cenach sprawia, że rynki praktycznie zawsze pozostają w równowadze, ponieważ dostosowanie cenowe jest samoczynne i natychmiastowe. Punkt równowagi będzie znajdował się zawsze na prostej odpowiadającej pełnemu zatrudnieniu Y_A. Oznacza to, że ingerencja władzy gospodarczej nie jest w takiej sytuacji potrzebna. Dlatego też w ramach założeń ekonomii klasycznej zagadnienie przywracania zachwianej równowagi nie jest w zasadzie analizowane.

Z kolei w ramach tradycyjnej ekonomii keynesowskiej zakłada się najczęściej, że ceny są sztywne (lub „lepkie" – *sticky prices*) w krótkim okresie, co oznacza, że przywracanie równowagi dokonuje się poprzez dostosowanie wielkości produkcji (podaży). Ponieważ ceny nie ulegają zmianom, również realna podaż pieniądza pozostaje stała, a krzywa *LM* nie zmienia swojego położenia. Nie ma również powodu, aby zmianie podlegał popyt (ponieważ ceny są stałe), a więc krzywa *IS* także nie zmienia danego położenia. W tych warunkach to podaż produktów musi zatem dostosować się do popytu. Przystosowanie produkcji w krótkim czasie nie jest jednak możliwe i gospodarka pozostaje nadal w punkcie *B* – z niższym poziomem produkcji i wyższymi stopami procentowymi. Ponieważ w ramach przyjętych założeń samoczynne przejście do punktu *A* jest niemożliwe, pojawia się uzasadnienie dla podjęcia odpowiednich działań w ramach polityki gospodarczej, pozwalających doprowadzić gospodarkę do punktu *A*. Przykładowo, bank centralny może zwiększyć podaż pieniądza, co wpłynie na przeniesienie krzywej *LM* w prawo do punktu *A*. Alternatywnie, rząd ma możliwość podniesienia wydatków publicznych, co pozwoli przesunąć krzywą *IS* w prawo do punktu *C*. W obu przypadkach poziom produkcji podniesie się z Y_B do Y_A, a to z kolei spowoduje pożądany wzrost zatrudnienia.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Czy keynesowskie założenie o sztywności cen jest uzasadnione? Badania potwierdzają, że przynajmniej w odniesieniu do towarów przemysłowych ceny rzeczywiście zmieniają się stosunkowo rzadko i generalnie w niewielkim stopniu[11]. Najczęściej wskazuje się na trzy przyczyny tej sztywności cen. Po pierwsze, producenci wiedzą, że zmiana cen wiąże się z przygotowaniem nowego „cenowego menu", co z kolei pociąga za sobą konieczność poniesienia dodatkowych kosztów administracyjnych i technicznych (nowe katalogi, nowe cenniki, nowe metki i nalepki, nowe opakowania). Po drugie, ważne są stosunki z klientami, zwłaszcza gdy chodzi o nabywców od dawna związanych z danym dostawcą i kupujących regularnie duże ilości produktów. Aby zachować konkurencyjną pozycję na rynku i nie tracić klientów, firma może powstrzymywać się przez dłuższy czas od wprowadzania podwyżek cen, godząc się na obniżenie marży zysku. Wreszcie jako trzecią z przyczyn wymienia się upowszechnienie długoterminowych kontraktów, w ramach których często ustala się stałą cenę dostawy. Dotyczy to na przykład kontraktów na regularne dostawy surowców i paliw (np. ceny dostaw węgla czy taryfy na dostawy gazu lub energii elektrycznej), a także określonych usług (np. usługi komunalne). Szczególnym rodzajem takich umów są kontrakty zbiorowe dotyczące płac oraz zwykłe indywidualne umowy o pracę, w których poziom wynagrodzeń nominalnych pozostaje ściśle określony.

Większość ekonomistów przyjmuje obecnie, że skrajne przypadki cen doskonale giętkich i doskonale sztywnych znajdują zastosowanie bardziej teoretyczne niż praktyczne i że w rzeczywistości ma się na ogół do czynienia z cenami, które w krótkim okresie pozostają sztywne (choć brakuje pełnej zgody w kwestii tego, co oznacza tzw. krótki okres). Rzeczywisty stopień sztywności cen nie tylko zależy też od struktury rynku (przykładowo, w warunkach monopolu ceny są sztywniejsze niż w przypadku wolnej konkurencji) i rodzaju produktu, którego dotyczą (na przykład, ceny produktów rolnych są bardziej giętkie od cen produktów przemysłowych), lecz ma także charakter asymetryczny, ponieważ różni się w zależności od tego, czy ma się do czynienia z ruchem cen w dół, czy w górę (przyjmuje się, że ceny są z reguły znacznie bardziej sztywne w dół, co oznacza, że łatwiej je podnosić niż obniżać).

Założenie o częściowej lub całkowitej sztywności cen sprawia, że w modelu IS-LM strona podażowa gospodarki przestaje funkcjonować jako ograniczenie. Podaż dostosowuje się do zmian popytu, podobnie jak podaż pracy przystosowuje

[11] Zob. Burda i Wypłosz (2000), s. 320.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

się do zmian produkcji w ramach istniejących zdolności wytwórczych. W takim modelu zasadniczo nie występuje też inflacja, co oczywiście jest wielkim uproszczeniem. W dalszych częściach rozdziału uwzględnione zostaną zmiany cen oraz ich wpływ na stronę podażową gospodarki. Kompletny opis funkcjonowania strony popytowej i podażowej przedstawiony będzie w ramach modelu AS-AD, czyli modelu zagregowanej podaży i zagregowanego popytu. Zanim to jednak nastąpi, należy opisać w kilku słowach rolę pieniądza w kształtowaniu zmiennych realnych, takich jak produkcja czy zatrudnienie.

2.4.2. Klasyczna dychotomia i neutralność pieniądza

Różnica pomiędzy podejściem (neo) klasycznym a podejściem keynesowskim nie polega tylko na przyjęciu odmiennego założenia, dotyczącego zachowania się cen. W obu opisanych doktrynach inaczej kształtuje się również rola, jaką pieniądz odgrywa w gospodarce. W przypadku cen giętkich wzrost nominalnej podaży pieniądza powoduje wzrost popytu na produkty, czemu towarzyszy natychmiastowy wzrost cen, a realna wielkość pieniądza w obiegu nie zmienia się. W efekcie krzywa *LM* zawsze przechodzi przez punkt równowagi ogólnej. Zmiany ilości pieniądza mają wpływ na ceny, ale nie przekładają się na zmiany poziomu produkcji i zatrudnienia. Innymi słowy, zmienne nominalne nie wpływają na zmienne realne – mamy więc do czynienia z tzw. klasyczną dychotomią (*classical dichotomy*), czyli rozdzieleniem sfery zmian nominalnych i sfery zmian realnych.

Ten brak zależności nazywany jest także hipotezą neutralności pieniądza. Pieniądz jest neutralny, gdy zmiany jego ilości nie wpływają na zmienne realne. Stwierdzenie to ma istotne konsekwencje praktyczne. Z hipotezy o neutralności pieniądza wynika, że próby pobudzania koniunktury, czyli stymulowania wzrostu produkcji i zatrudnienia, przy pomocy instrumentów polityki pieniężnej, polegających na zwiększaniu ilości pieniądza w obiegu, są nieskuteczne. W efekcie zastosowania takiej polityki nastąpi jedynie wzrost cen, a więc pojawi się inflacja. Zmienne nominalne oddziałują zatem na inne zmienne nominalne (ceny, kurs), lecz nie wpływają na poziom zmiennych realnych.

Hipoteza neutralności pieniądza, w swojej skrajnej postaci mówiąca o tym, że pieniądz nie ma w ogóle żadnego wpływu na zmienne realne, pozostaje jednak w pełni prawdziwa jedynie w odniesieniu do zmian gospodarczych zachodzących w długim okresie. W przypadku średniego i krótkiego okresu rezygnuje z niej bowiem większość ekonomistów. Należy jednak zauważyć, że odrzucana

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jest też hipoteza przeciwstawna, w myśl której przy pomocy zmian ilości pieniądza można w sposób trwały wpływać na zmienne realne. Większość ekonomistów skłania się obecnie do akceptowania hipotezy o ograniczonej neutralności pieniądza, która mówi o tym, że w krótkim i średnim okresie zmiany podaży pieniądza mogą wywierać pewien wpływ na zmienne realne, ale wpływ ten jest nietrwały i z upływem czasu zanika.

Przyjęcie założenia o sztywności cen (w krótkim okresie) oznacza, że w gospodarce nie występuje klasyczna dychotomia, a pieniądz nie ma charakteru neutralnego. Ponieważ ceny pozostają stałe, zmiany ilości pieniądza powodują przesunięcie krzywej *LM* i zmianę położenia punktu równowagi. Oznacza to, że w krótkim okresie zmienne realne – takie jak produkcja i zatrudnienie – zmieniają się pod wpływem zmian ilości pieniądza w obiegu.

2.5. Polityka makroekonomiczna w modelu IS-LM

Model IS-LM pozwala analizować wpływ zmiennych egzogenicznych – czyli zmiennych polityki gospodarczej, takich jak wydatki rządowe czy zmiany podaży pieniądza – na zmienne endogeniczne, takie jak produkcja, zatrudnienie bądź rynkowe stopy procentowe. Model IS-LM może więc dostarczać użytecznych ram analitycznych, służących ocenie skuteczności (*effectiveness*) polityki gospodarczej. Skuteczność polityki gospodarczej w kontekście działań stabilizacyjnych będziemy definiować jako zdolność polityki do przeciwdziałania cyklicznym załamaniom koniunktury, a zatem do podtrzymania określonego poziomu produkcji i zatrudnienia. Na tym etapie analizy zbadamy skuteczność podstawowych instrumentów polityki gospodarczej, tj. polityki fiskalnej i polityki pieniężnej, w podtrzymywaniu i zwiększaniu poziomu produkcji i zatrudnienia. Jest to ważne zagadnienie praktyczne, ponieważ, chcąc utrzymać koniunkturę, władza gospodarcza musi wiedzieć, które instrumenty – fiskalne czy pieniężne – są w danej sytuacji skuteczniejsze oraz na jaką skalę i w jakim zakresie powinny być one zastosowane.

W pierwszej kolejności omówione zostaną warunki skuteczności polityki fiskalnej. Jak wiadomo, polityka fiskalna polega m.in. na zwiększaniu lub ograniczaniu wydatków budżetowych (rządowych). Zmiana wydatków powoduje odpowiednie przesunięcie krzywej *IS*, co z kolei zmienia położenie punktu przecięcia się jej z krzywą *LM*, a zatem i usytuowanie punktu równowagi ogólnej. Zmiana położenia punktu równowagi oznacza, że zmieniają się także wartości produkcji *Y* i stopy

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

procentowej *r*. Na ogół w polityce gospodarczej chodzi o to, aby przy pomocy danego instrumentu polityki gospodarczej uzyskać jak największy wzrost produkcji i dochodu oraz związanego z nimi zatrudnienia. Należy więc stwierdzić, że polityka fiskalna jest skuteczna, gdy dany wzrost wydatków rządowych przekłada się na względnie wysoki wzrost produkcji i dochodu. W przeciwnym przypadku powiemy, że polityka fiskalna jest nieskuteczna.

W jaki sposób można ocenić skuteczność polityki fiskalnej? Najpierw należy zastanowić się nad tym, jak wzrost wydatków rządowych wpływa na poziom produkcji. Pierwszym czynnikiem, który trzeba wziąć pod uwagę, jest wielkość mnożnika autonomicznych wydatków rządowych. Całkowity wzrost popytu i produkcji wywołany wzrostem wydatków rządowych będzie z reguły większy od początkowego wzrostu, pod warunkiem wszakże istnienia w gospodarce wolnych mocy wytwórczych. Im większy będzie mnożnik, tym większe okażą się przyrosty produkcji wobec danego wzrostu wydatków rządowych[12]. Następnie należy zauważyć, że zwiększenie wydatków rządowych powoduje wzrost transakcyjnej części popytu na pieniądz (w celu obsłużenia zakupów rządowych), co przy niezmienionej podaży pieniądza musi przełożyć się na wzrost stóp procentowych. To z kolei wpłynie na ulokowanie większej ilości pieniądza gotówkowego na rachunkach bankowych w postaci lokat (lub wzrost zakupu obligacji ze względu na rosnące oprocentowanie). W efekcie część spekulacyjna popytu na pieniądz ulegnie obniżeniu. Dzięki temu wzrost popytu transakcyjnego może zostać skompensowany spadkiem popytu spekulacyjnego na pieniądz, skutkiem czego rynek pieniądza pozostanie w równowadze. Rosnące stopy procentowe wywołają jednak spadek tych kategorii wydatków, które zależą od poziomu stóp, czyli przede wszystkim inwestycji. Ostatecznie wpływ wzrostu wydatków rządowych na wzrost popytu i produkcji będzie więc pomniejszony o wielkość spadku inwestycji. Drugim czynnikiem decydującym o skuteczności polityki fiskalnej pozostaje zatem stopień wrażliwości inwestycji na wzrost stóp procentowych – im bardziej wrażliwe na wzrost stóp okażą się inwestycje, tym ich spadek będzie silniejszy i tym mniej skuteczna stanie się polityka fiskalna w zakresie pobudzania produkcji i zatrudnienia[13].

[12] Zob. rozdział 1. Opis mechanizmu mnożnika wydatków można także znaleźć w podstawowych podręcznikach makroekonomii. Zob. np. Hall i Taylor (2000), rozdział 6.

[13] W najprostszym modelu IS-LM stopa procentowa oddziałuje tylko na wydatki inwestycyjne. Trzeba jednak pamiętać, że w rzeczywistości ma ona także wpływ na inne kategorie wydatków, takie jak konsumpcja, w tym finansowana kredytem, oraz eksport netto.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

O ile wzrosną jednak rynkowe stopy procentowe w wyniku zwiększenia wydatków rządowych? To już z kolei zależy od postaci funkcji popytu na pieniądz, a w szczególności od wrażliwości spekulacyjnego popytu na pieniądz na zmiany stóp procentowych. Aby skompensować wzrost popytu transakcyjnego o określoną wielkość, wywołaną wzrostem wydatków rządowych, popyt spekulacyjny – przy danej podaży pieniądza – musi obniżyć się o tę samą wielkość, a to wymaga odpowiedniego wzrostu rynkowych stóp procentowych. Trzecim czynnikiem decydującym o skuteczności polityki fiskalnej będzie więc wrażliwość popytu na pieniądz na zmiany stóp procentowych – im mniej wrażliwy jest popyt na pieniądz na zmiany stóp, tym wyższy wzrost stóp jest konieczny w celu wywołania pożądanego spadku spekulacyjnego popytu na pieniądz i utrzymania równowagi na rynku pieniądza, i tym mniej skuteczna jest z kolei polityka fiskalna.

W podobny sposób można określić warunki skuteczności polityki pieniężnej. Mówi się, że polityka pieniężna jest skuteczna, jeśli dany wzrost podaży pieniądza powoduje względnie duże przyrosty produkcji, dochodu i zatrudnienia. W przeciwnym razie przyjmuje się, że polityka pieniężna będzie nieskuteczna. Podobnie jak w przypadku polityki fiskalnej pierwszym warunkiem skuteczności jest występowanie wolnych mocy produkcyjnych. Następnie należy zauważyć, że w ramach modelu IS-LM wzrost ilości pieniądza przekłada się na spadek rynkowych stóp procentowych. W dalszej kolejności wywołuje on wzrost tych kategorii wydatków, które pozostają zależne od stanu stopy procentowej – głównie inwestycji. Drugim czynnikiem wpływającym na skuteczność polityki pieniężnej jest więc wrażliwość inwestycji na spadek stóp procentowych – im silniejsza będzie reakcja inwestycji na dany spadek stóp procentowych, tym skuteczniejsza stanie się polityka pieniężna.

O ile jednak spadną rynkowe stopy procentowe na skutek wzrostu podaży pieniądza? To z kolei zależy od postaci funkcji popytu na pieniądz, a w szczególności od wrażliwości popytu spekulacyjnego na zmiany stóp procentowych. Ponieważ popyt transakcyjny na pieniądz w tym przypadku nie ulega zmianie, wzrostowi podaży pieniądza musi towarzyszyć wzrost popytu spekulacyjnego na pieniądz, aby rynek pieniądza pozostał w równowadze. Wymaga to spadku stóp procentowych. Trzecim czynnikiem określającym skuteczność polityki pieniężnej jest więc wrażliwość popytu na pieniądz na zmiany stóp procentowych – im mniej wrażliwy okaże się popyt na pieniądz, tym większy spadek stóp będzie wymagany w celu zwiększenia popytu niezbędnego dla zaabsorbowania wzrostu podaży pieniądza, i tym skuteczniejsza stanie się polityka pieniężna.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

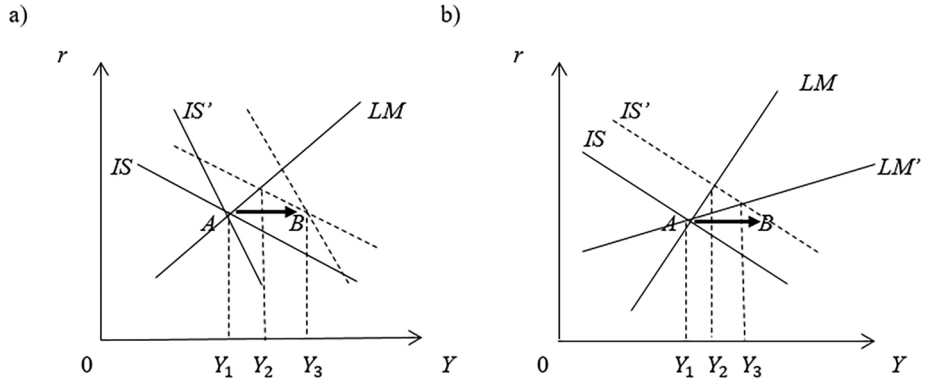
Powyższe wnioski można zilustrować za pomocą analizy położenia i nachylenia krzywych *IS* i *LM*. Należy zwrócić uwagę, że przesunięcia krzywej *IS*, wywołane zmianami poziomu wydatków rządowych, przekładają się na różne wartości przyrostu (spadku) poziomu produkcji (dochodu) w zależności od tego, czy krzywa *IS*, a także krzywa *LM* jest bardziej stroma czy bardziej płaska. Podobnie przesunięcia krzywej *LM*, związane ze zmianami podaży pieniądza, przekładają się na różne wartości przyrostu (spadku) poziomu produkcji (dochodu) w zależności od stopnia nachylenia obu krzywych. Łatwo zauważyć, że krzywa *IS* będzie względnie płaska, jeśli wrażliwość wydatków inwestycyjnych na zmiany stopy procentowej pozostanie wysoka, ponieważ małe zmiany stopy procentowej wywołują wówczas duże zmiany poziomu produkcji. Dzieje się również odwrotnie: krzywa *IS* będzie względnie stroma, gdy wrażliwość wydatków inwestycyjnych na zmiany stopy procentowej pozostanie niska. Podobnie krzywa *LM* stanie się względnie płaska, gdy wrażliwość popytu na pieniądz na zmiany stopy procentowej będzie wysoka, ponieważ małe zmiany stopy procentowej wywołują duże zmiany w zakresie popytu na pieniądz, co z kolei pozwala dostosować popyt spekulacyjny do stałej podaży pieniądza w przypadku znacznych zmian produkcji i dochodu.

Różne warianty nachylenia krzywych *IS* i *LM* zostały przedstawione na rysunku 2.5 i 2.6. Pierwszy z nich ilustruje zależność skuteczności polityki fiskalnej od nachylenia krzywej *IS* (część a) i od nachylenia krzywej *LM* (część b). Polityka fiskalna będzie skuteczna przy stromym położeniu krzywej *IS* (część a), ponieważ wobec danego wzrostu wydatków rządowych o odcinek *AB* uzyskany przyrost produkcji *Y* (mierzony na osi poziomej) w przypadku stromej krzywej *IS*' wyniesie ($Y_3 - Y_1$)i będzie większy niż przyrost produkcji uzyskany w przypadku płaskiej krzywej *IS* ($Y_2 - Y_1$), przy założeniu, że położenie krzywej *LM* pozostaje niezmienione. Dzieje się tak dlatego, że stroma krzywa *IS* oznacza niską wrażliwość popytu inwestycyjnego na wzrost stóp procentowych.

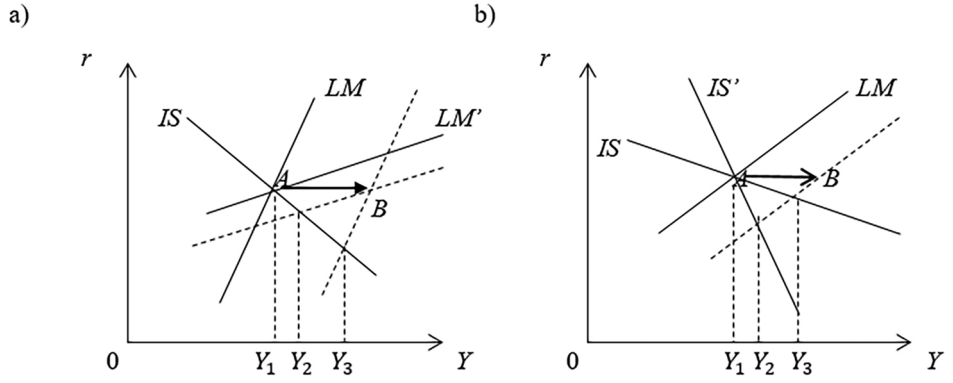
Polityka fiskalna będzie również bardziej skuteczna przy płaskim położeniu krzywej *LM* (część b), ponieważ wobec danego wzrostu wydatków rządowych o odcinek *AB*, a także przy danym nachyleniu krzywej *IS*, uzyskany przyrost produkcji *Y* w przypadku płaskiej krzywej *LM*' wyniesie ($Y_3 - Y_1$) i będzie większy niż przyrost produkcji ($Y_2 - Y_1$) uzyskany w przypadku stromej krzywej *LM*. Dzieje się tak dlatego, że płaska krzywa *LM* oznacza wysoką wrażliwość popytu na pieniądz na zmiany stóp procentowych. Polityka fiskalna jest zatem tym skuteczniejsza, im bardziej stroma staje się krzywa *IS* oraz im bardziej płaska jest krzywa *LM*.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 2.5. Skuteczność polityki fiskalnej w zakresie różnych wariantów położenia krzywych *IS* i *LM*



Rysunek 2.6. Skuteczność polityki pieniężnej w zakresie różnych wariantów położenia krzywych *IS* i *LM*



Źródło: opracowanie własne.

Rysunek 2.6 ilustruje zależność skuteczności polityki pieniężnej od nachylenia krzywej *LM* (część a) oraz od nachylenia krzywej *IS* (część b). Polityka pieniężna będzie skuteczna przy stromej krzywej *LM* (część a), ponieważ wobec danego wzrostu podaży pieniądza (przesunięcie krzywej *LM* o odcinek *AB*) i przy danej krzywej *IS* uzyskany przyrost produkcji (z Y_1 do Y_3) jest większy niż przyrost produkcji (z Y_1 do Y_2) uzyskany w przypadku płaskiej krzywej *LM'*. Dzieje się tak dlatego, że stroma krzywa *LM* oznacza niską wrażliwość popytu na pieniądz na zmiany

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

stóp procentowych. Podobnie polityka pieniężna jest bardziej skuteczna przy płaskiej krzywej *IS* (część b), ponieważ uzyskany przyrost produkcji jest większy niż w przypadku stromej krzywej *IS*. Dzieje się tak dlatego, że płaska krzywa *IS* oznacza wysoką wrażliwość popytu inwestycyjnego na spadek stóp procentowych. Polityka pieniężna okazuje się zatem tym skuteczniejsza, im bardziej stroma pozostaje krzywa *LM* oraz im bardziej płaska jest krzywa *IS*.

Tablica 2.2 zawiera podsumowanie powyższej analizy. Łatwo zauważyć, że warunki skuteczności polityki pieniężnej i polityki fiskalnej są względem siebie przeciwstawne, co oznacza, że w – w danych warunkach – jeśli jedna polityka pozostaje skuteczna, to druga z nich będzie nieefektywna. Skoro można więc określić ogólne parametry funkcji inwestycji i funkcji popytu na pieniądz (czyli stopień nachylenia krzywych *IS* i *LM*), możliwe jest w sposób jednoznaczny wybranie takiej polityki, która w danych warunkach okaże się skuteczniejsza. Należy również pamiętać o tym, że wnioski te uzyskano, przyjmując pewne istotne założenia, takie jak m.in. założenie niezmienności cen w krótkim okresie oraz założenie o braku wymiany z zagranicą (gospodarka zamknięta). Jak się jednak później przekonamy, uchylenie tych założeń nie zmienia zasadniczych konkluzji dotychczasowej analizy.

Tabela 2.2. Warunki skuteczności polityki gospodarczej w prostym modelu IS-LM

Parametry polityki	Polityka fiskalna		Polityka pieniężna	
	skuteczna	nieskuteczna	skuteczna	nieskuteczna
Wrażliwość inwestycji na zmiany stopy procentowej	mała	duża	duża	mała
Nachylenie krzywej *IS*	stroma	płaska	płaska	stroma
Wrażliwość popytu na pieniądz na zmiany stopy procentowej	duża	mała	mała	duża
Nachylenie krzywej *LM*	płaska	stroma	stroma	płaska

Źródło: opracowanie własne.

2.6. Uwzględnienie wymiany z zagranicą i bilansu płatniczego

2.6.1. Model IS-LM dla gospodarki otwartej

Gospodarka zamknięta, całkowicie oddzielona od otoczenia zagranicznego, jest pojęciem abstrakcyjnym. We współczesnym świecie, w warunkach globalizacji i integracji rynków oraz międzynarodowych przepływów czynników produkcji,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

technologii i informacji, nie można odizolować się od otoczenia zagranicznego – nawet gospodarki najbardziej autarkiczne podlegają w pewnym zakresie wpływom zagranicy. Poszczególne gospodarki narodowe są ze sobą wzajemnie powiązane i od siebie uzależnione. Związki te obejmują wymianę dóbr i usług, przepływy kapitałów i tytułów własności oraz powiązania prawne i instytucjonalne, wynikające z umów międzynarodowych, procesów tworzenia i rozwoju korporacji transnarodowych oraz tendencji do koordynowania polityki gospodarczej, zwłaszcza w obrębie ugrupowań integracyjnych. Chcąc więc bardziej urealnić model IS-LM, należy uzupełnić go o składniki odzwierciedlające wpływ zagranicy na równowagę na rynku produktów i pieniądza krajowego oraz wprowadzić do analizy zagadnienie równowagi bilansu płatniczego, tj. rynek pieniądza zagranicznego.

W warunkach gospodarki otwartej struktura i poziom produkcji wytworzonej w kraju różni się na ogół od struktury i poziomu wydatków końcowych. Część wytworzonego produktu jest bowiem eksportowana za granicę, a część popytu krajowego zaspokaja import z zagranicy. Uwzględniając handel zagraniczny w zakresie dóbr i usług, warunek równowagi dochodu i wydatków przyjmuje następującą postać:

$$Y = C + I + G + X, \tag{2.19}$$

gdzie wielkość eksportu netto X stanowi różnicę pomiędzy eksportem a importem dóbr i usług. Jeśli założymy, że saldo przepływu dochodów od czynników produkcji (płace, odsetki, zyski) oraz saldo transferów są równe zeru, eksport netto staje się równy saldu rachunku obrotów bieżących bilansu płatniczego:

$$X = EX - IM \Rightarrow X = EX - mY \Rightarrow X = X(Y), \tag{2.20}$$

gdzie:
$m = IM/Y$, $0 < m < 1$.

Zapis (2.20) pokazuje, że eksport EX jest zmienną egzogeniczną, określoną przez popyt zagranicy (nie uwzględniony w modelu), z kolei import IM jest zmienną endogeniczną, zależną od wielkości krajowego dochodu Y. Współczynnik m, nazywany skłonnością do importu (*propensity to import*) informuje o tym, jaka część z każdej jednostki dochodu krajowego Y przeznaczana jest na wydatki na import. Przy tych założeniach eksport netto staje się także funkcją dochodu Y[14].

[14] W rzeczywistości, ważnym czynnikiem określającym poziom eksportu i importu jest kurs walutowy. W ramach najprostszej wersji modelu IS-LM zakładamy, że kurs walutowy jest dany, podobnie jak

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Drugą częścią bilansu płatniczego jest rachunek obrotów finansowych (kapitałowych). Część transakcji ma tu charakter egzogeniczny, na przykład inwestycje bezpośrednie (fuzje, przejęcia, inwestycje typu *green-field*) lub inwestycje w instrumenty udziałowe (akcje, udziały). Pozostała część transakcji pozostaje zależna od poziomu krajowej stopy procentowej, która – przy danym poziomie stóp procentowych za granicą i przy danym oraz przewidywanym poziomie kursu walutowego – określa oczekiwaną opłacalność inwestycji w krajowe instrumenty finansowe (obligacje skarbowe, bony skarbowe, obligacje przedsiębiorstw, lokaty bankowe, a także tzw. instrumenty syntetyczne i instrumenty pochodne – *derivatives*)[15]. Im wyższa będzie krajowa stopa procentowa, tym większy – przy pozostałych warunkach niezmienionych – stanie się napływ kapitału zagranicznego. W świetle tych założeń, saldo bilansu obrotów kapitałowych będzie więc rosnącą funkcją krajowej stopy procentowej:

$$B = B(r). \tag{2.21}$$

Konkretny kształt funkcji (2.21) zależy od elastyczności przepływów kapitałowych wobec stopy procentowej. W gospodarce o walucie krajowej niewymienialnej na waluty zagraniczne elastyczność przepływów będzie zasadniczo równa zeru. W gospodarce w pełni otwartej, wszechstronnie powiązanej z zagranicą i bez jakichkolwiek ograniczeń w zakresie wymiany walut elastyczność przepływów kapitału pozostanie bardzo duża, niekiedy bliska nieskończoności.

Różne warianty elastyczności przepływów kapitału przedstawiono na rysunku 2.7. W gospodarce o walucie niewymienialnej zmiany stopy procentowej nie mają żadnego wpływu na ruch kapitału, a zmiany w zakresie salda bilansu płatniczego określone są jedynie za sprawą zmian na rachunku obrotów bieżących (czyli przez wielkość eksportu netto), a zatem przez poziom dochodu Y. Ten przypadek ilustruje pionowa krzywa BP – elastyczność przepływu jest równa zeru ($\gamma = 0$), ponieważ nie reaguje on na zmiany stóp procentowych. Z drugim krańcowym przypadkiem mamy do czynienia wówczas, gdy nawet najmniejsza zmiana stopy procentowej wywołuje nieskończenie duże przepływy kapitału. Saldo bilansu płatniczego jest określone w tym przypadku przez saldo obrotów kapitałowych, ponieważ przepływy kapitałowe są tak duże, że obroty bieżące nie mają znaczenia. Ów przypadek, kiedy

ceny krajowe P i ceny zagraniczne P^*. Kurs walutowy zostanie uwzględniony w analizie polityki gospodarczej w dalszych rozdziałach.

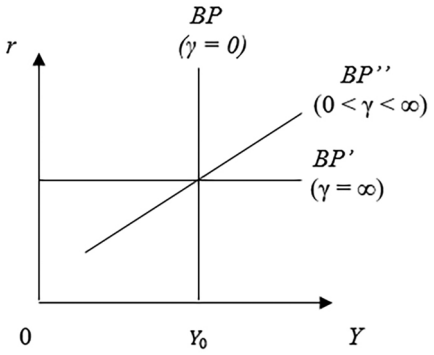
[15] Np. transakcje terminowe (*forward, futures*), opcje walutowe czy transakcje walutowe typu *swap*.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

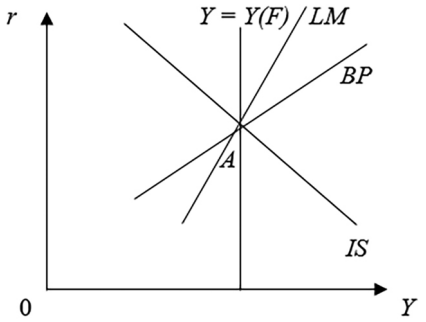
elastyczność pozostaje nieskończenie duża ($\gamma = \infty$), ilustruje pozioma krzywa *BP'*. Wreszcie przypadek pośredni, z uwzględnieniem skończonej, lecz różnej od zera elastyczności, ilustruje rosnąca krzywa *BP"*.

Ten ostatni wariant jest oczywiście najbliższy rzeczywistości. Wyższe krajowe stopy procentowe przyciągają kapitał zagraniczny, lecz wielkość przepływów staje się ograniczona, ponieważ krajowe i zagraniczne instrumenty finansowe nie są doskonałymi substytutami, a inwestorzy pozostają z reguły bardziej skłonni do kupowania krajowych aktywów (*home bias*). Inwestycje zagraniczne są ponadto obarczone dodatkowym ryzykiem, co wynika z możliwości wystąpienia zmian kursu waluty (deprecjacja waluty), niekorzystnej zmiany polityki gospodarczej (podniesienie podatków, wprowadzenie ograniczeń w wymienialności waluty krajowej) lub ryzyka niewypłacalności. Ponieważ różni inwestorzy w rozmaity sposób przejawiają awersję do ryzyka (*risk aversion*), tylko część z nich dokona inwestycji przy danym poziomie stóp procentowych[16]. Oznacza to, że aby przyciągnąć kolejne „transze" kapitału zagranicznego, krajowe stopy procentowe muszą być coraz wyższe.

Rysunek 2.7. Elastyczność przepływu kapitału



Rysunek 2.8. Równowaga ogólna



Źródło: opracowanie własne.

[16] Drugim czynnikiem ograniczającym napływ kapitału jest dążenie inwestorów do optymalizacji struktury portfela posiadanych aktywów. Optymalizacja polega na takim ich doborze, aby cały portfel zapewniał pożądaną równowagę pomiędzy dochodowością, ryzykiem i płynnością różnych rodzajów aktywów. Przy tych założeniach inwestorzy będą skłonni zwiększyć udział określonych aktywów w portfelu tylko wówczas, gdy rentowność tych aktywów wzrośnie lub ryzyko z nimi związane spadnie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Łącząc równania (2.20) i (2.21), otrzymujemy równanie salda całego bilansu płatniczego *BP*:

$$BP = X(Y) + B(r). \tag{2.22}$$

Bilans płatniczy będzie zrównoważony, gdy $BP = 0$, czyli kiedy deficyt (nadwyżka) rachunku bieżącego stanie się dokładnie równa nadwyżce (deficytowi) rachunku kapitałowego. Jest to równoznaczne z zachowaniem równowagi na krajowym rynku pieniądza zagranicznego, ponieważ dla $BP = 0$ podaż pieniądza zagranicznego równa się popytowi na ten pieniądz. Jeśli uzupełnimy prosty model IS-LM o równanie (2.22), uwzględniające przepływy kapitałowe i funkcjonowanie rynku pieniądza zagranicznego, otrzymamy rozszerzoną wersję modelu, która nazywana jest najczęściej modelem Mundella-Fleminga.

2.6.2. Model Mundella-Fleminga

Równanie (2.22) pokazuje, że saldo bilansu płatniczego składa się z dwóch elementów – salda obrotów bieżących i salda obrotów kapitałowych. Saldo obrotów bieżących pogarsza się, gdy rośnie produkcja (dochód). Saldo obrotów kapitałowych wykazuje poprawę, gdy rośnie stopa procentowa. Zachowanie się salda całego bilansu płatniczego zależy więc od tego, jaka jest wrażliwość obu składowych bilansu na zmiany obu zmiennych – dochodu Y i stopy procentowej r.

Krzywa *BP* reprezentuje takie kombinacje poziomu produkcji Y i poziomu stopy procentowej r, które zapewniają równowagę bilansu płatniczego. Z powodów opisanych wyżej, krzywa *BP* jest krzywą rosnącą. Interpretacja ekonomiczna jest w tym przypadku prosta. Załóżmy, że dla określonego poziomu dochodu (produkcji) Y_0 bilans płatniczy jest zrównoważony (rysunek 2.7). Wzrost dochodu oznacza wzrost wydatków, w tym wydatków na import, co spowoduje pojawienie się deficytu w obrotach bieżących. Aby gospodarka powróciła do równowagi, stopy procentowe muszą wzrosnąć, żeby wywołany w ten sposób napływ kapitału zagranicznego wyrównał deficyt w obrotach bieżących. Utrzymanie równowagi płatniczej wymaga więc, aby wzrostowi Y towarzyszył wzrost r – krzywa *BP* musi być więc krzywą rosnącą. Z kolei nachylenie krzywej *BP* zależy od elastyczności napływu kapitału względem stopy procentowej. Im bardziej wrażliwe są ruchy kapitału na zmiany stopy procentowej, tym bardziej płaska staje się krzywa *BP*.

Formalnie rzecz biorąc, nachylenie krzywej *BP* można obliczyć, różniczkując całkowicie funkcję (2.22), uwzględniając (2.19) i przyrównując wynik do zera:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$dB = -mdY + dr(\delta B/\delta r) = 0,$$

co daje po przekształceniu:

$$dr/dY = m/(\delta B/\delta r). \tag{2.23}$$

Wielkość dr/dY jest zawsze dodatnia, ponieważ zarówno licznik, jak i mianownik ułamka po prawej stronie równania (2.23) pozostają zawsze większe od zera. Nachylenie krzywej *BP* jest więc zawsze dodatnie. Można też zauważyć, że krzywa *BP* będzie tym bardziej stroma, im większa jest krańcowa skłonność do importu *m* oraz im mniejsza staje się wrażliwość bilansu kapitałowego na zmiany stopy procentowej ($\delta B/\delta r$).

Rysunek 2.8 przedstawia równowagę ogólną w gospodarce. Poza znanymi już krzywymi *IS* i *LM* zaznaczona została także prosta pełnego zatrudnienia (Y_F) oraz krzywa równowagi płatniczej *FB*. Jeśli wszystkie cztery krzywe przecinają się w jednym punkcie, zachowana jest równowaga na każdym z czterech rynków. Punkt *A* staje się więc punktem równowagi ogólnej. Jak wcześniej wspomniano, stan taki jest raczej wyjątkiem niż regułą. W rzeczywistości ma się najczęściej do czynienia z występowaniem nierównowagi cząstkowej na poszczególnych rynkach. Jest ona jednak w sposób ciągły eliminowana częściowo przez samoczynne procesy dostosowawcze w gospodarce, a częściowo przez określone działania podejmowane w ramach polityki gospodarczej. Ponieważ gospodarka ulega zaś nieprzerwanie zakłóceniom wynikającym z wpływu wahań cyklicznych i szoków zewnętrznych, nierównowagi nigdy nie da się definitywnie wyeliminować.

2.7. Model AS-AD

2.7.1. Wpływ zmian cen i krzywa łącznego (zagregowanego) popytu *AD*

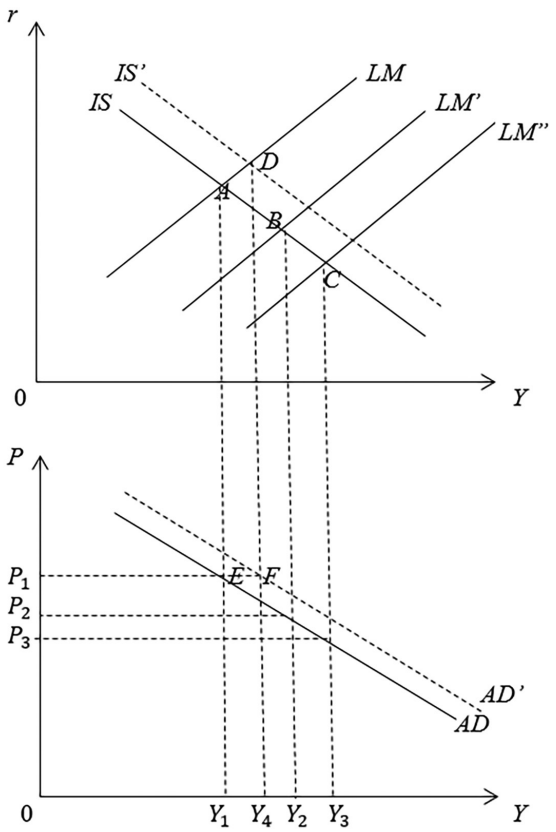
W dotychczasowej analizie przyjmowaliśmy założenie, że ceny nie ulegają zmianie. Założenie to można interpretować jako założenie doskonale elastycznej podaży. Oznacza ono wówczas, że producenci są gotowi dostarczać przy danej cenie dowolną ilość produktów. Taka sytuacja nie musi być odrealniona, zwłaszcza w odniesieniu do pojedynczego sprzedawcy, który ma niewielki wpływ na rynek.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wówczas model IS-LM może być traktowany jako model wyjaśniający kształtowanie się dochodu i produkcji pod wpływem zmian popytu.

W skali makroekonomicznej założenie o stałości cen jest jednak bardzo krępujące, ponieważ nie pozwala na wyjaśnienie tak powszechnego i ważnego zjawiska, jakim jest inflacja. Założenie to ogranicza również możliwości badania zmian na rynku pracy, wynikających z wahań płacy realnej, oraz czynników określających poziom i zmiany bezrobocia. Wprowadzimy więc obecnie do analizy zmiany cen – najpierw jako zmiany egzogeniczne, a następnie jako zmiany wywołane przez zmienne występujące w modelu (np. jako skutek zastosowania określonej polityki gospodarczej).

Rysunek 2.9. Zmiany cen i wyprowadzenie krzywej łącznego popytu *AD*



Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Zauważmy, że uwzględnienie zmian cen w równaniu wydatków (2.7) i w równaniu popytu na pieniądz (2.11) nie ma wpływu na warunki równowagi[17]. Lecz zmiany cen mają wpływ na realną podaż pieniądza. Przy danej nominalnej podaży pieniądza realna podaż pieniądza rośnie (spada), gdy ceny spadają (rosną). W efekcie położenie krzywej *LM* ulega zmianie, czego dowodzi rysunek 2.9. Gdy poziom cen obniża się z P_1 do P_2, a następnie do P_3, krzywa *LM* przesuwa się w prawo w dół do położenia *LM'* i *LM"*.

Dalsze punkty przecięcia krzywej *IS* z kolejnymi krzywymi *LM* wyznaczają nowe punkty równowagi, którym odpowiadają nowe, coraz wyższe poziomy popytu Y_1, Y_2, Y_3. Dzięki temu można do modelu wprowadzić ceny *P* jako zmienną egzogeniczną i wyrazić popyt *Y* jako funkcję poziomu cen. Zostało to zobrazowane w dolnej części rysunku. 2.9, gdzie coraz niższym poziomom cen P_1, P_2, P_3, odpowiadają coraz wyższe poziomy popytu Y_1, Y_2, Y_3. Łącząc kolejne punkty o współrzędnych (P_1, Y_1), (P_2, Y_2), (P_3, Y_3), otrzymujemy krzywą zagregowanego popytu *AD*.

Ujemne nachylenie krzywej *AD* jest intuicyjnie oczywiste i wynika z mikroekonomicznych założeń modelu równowagi cząstkowej (spadek cen na dane dobro powoduje wzrost popytu na to dobro). Lecz w ramach modelu IS-LM opadająca krzywa *AD* jest wynikiem działania tzw. efektu keynesowskiego. Im niższe ceny, tym wyższa realna podaż pieniądza, co wymaga obniżenia stopy procentowej, aby zwiększyć popyt na pieniądz i utrzymać równowagę na rynku pieniądza. Spadek stóp procentowych pobudza wydatki inwestycyjne i zwiększa wszystkie wydatki za pośrednictwem mechanizmu mnożnikowego. W rezultacie wielkość *Y* rośnie.

Zauważmy, że nachylenie krzywej *AD* zależy od nachylenia krzywych *IS* i *LM*. Im bardziej stroma krzywa *LM* i płaska krzywa *IS*, tym bardziej płaska jest krzywa *AD*. Taki układ krzywych oznacza, że realne zmiany podaży pieniądza (na skutek zmian cen) mają silny wpływ na stopy procentowe, a zmiany stóp mają z kolei silny wpływ na zmiany zagregowanego popytu. Z kolei im bardziej płaska krzywa *LM* i stroma krzywa *IS*, tym bardziej stroma staje się krzywa *AD*, co oznacza, że realne zmiany pieniądza mają niewielki wpływ na stopy procentowe i słaby wpływ na łączny (zagregowany) popyt. W pierwszym przypadku zmiany cen wywierają silny wpływ na zmiany popytu, co jest zgodne ze stanowiskiem ekonomistów (neo)klasycznych, z kolei w drugim przypadku zmiany cen mają niewielki wpływ na zagregowany popyt, co jest zgodne z poglądami ekonomistów keynesowskich.

[17] Wynika to z faktu, że oba równania są homogeniczne w stopniu zero pod względem zmian w zakresie cen. Funkcja stopnia α jest homogeniczna (jednorodna), gdy pomnożenie każdego z jej argumentów przez stałą λ spowoduje zmianę wartości funkcji w proporcji λ^{α}.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Położenie krzywej *AD* jest określone przez poziom autonomicznych składników wydatków (wydatki rządowe *G*) oraz przez egzogenicznie określony poziom podaży pieniądza *M*. Zmiany *G* lub *M* powodują więc przesunięcia krzywej *AD*. Na rysunku 2.9 przedstawiony jest wpływ wzrostu wydatków rządowych na położenie krzywej *AD*. Wzrost wydatków powoduje przesunięcie krzywej *IS* w prawo w górę do położenia *IS*', co przy danym *M* i *P* oznacza wzrost stopy procentowej, spadek spekulacyjnego popytu na pieniądz i wzrost produkcji z poziomu Y_1 do poziomu Y_4 (dolna część rysunku). Krzywa *AD* przesuwa się natomiast w prawo w górę o odcinek *EF*, do położenia *AD*'[18].

2.7.2. Krzywa łącznej (zagregowanej) podaży *AS*

Jak wspomniano wyżej, w swojej prostej postaci model IS-LM jest zasadniczo modelem popytu, ponieważ przyjmuje się, że podaż jest doskonale elastyczna i dostosowuje się do zmian popytu przy stałych cenach. Dla każdego, egzogenicznie danego poziomu cen, podaż (produkcja) jest określona wyłącznie przez popyt. Taka interpretacja zmian w gospodarce jest zgodna ze „skrajnie keynesowskim" poglądem, wedle którego, zwiększając zagregowany popyt (np. poprzez wzrost wydatków rządowych) można – bez żadnych dodatkowych kosztów i przezwyciężając prawo malejących przychodów – doprowadzić do pożądanego wzrostu produkcji i zatrudnienia. Na rysunku 2.10 doskonale elastyczną podaż reprezentuje linia pozioma *AS*. Oznacza ona, że przy danej cenie P_1 przedsiębiorcy są w stanie zapewnić dowolną ilość produkcji. Punkt przecięcia poziomej krzywej zagregowanej podaży *AS* z opadającą krzywą zagregowanego popytu *AD* wyznacza poziom produkcji przy danym poziomie cen. Przesunięcie krzywej popytu z *AD* do *AD*' powoduje wzrost produkcji z Y_1 do Y_2. Popyt jest w tym modelu jedynym czynnikiem określającym wielkość produkcji.

Taka „skrajnie keynesowska" ilustracja budzi jednak uzasadnione wątpliwości. Już intuicja podpowiada, że w praktyce trudno oczekiwać, aby produkcja rosła w dowolnym stopniu przy niezmienionych cenach. Codzienne doświadczenie potwierdza

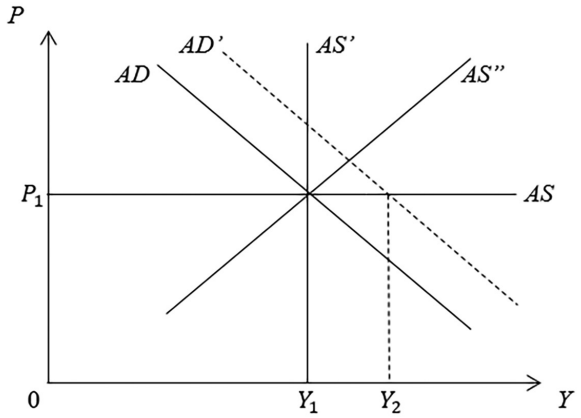
[18] Położenie krzywej *AD* zależy także od zasobu majątku. W standardowej analizie zasób majątku traktuje się jako dany, co jest uzasadnione tym, że strumienie oszczędności, powiększające majątek we wskazanym przedziale czasowym, są na ogół niewielkie w stosunku do zasobu majątku. Jednak gdy majątek jest wyrażony w kategoriach nominalnych, jego wartość może ulegać większym zmianom na skutek zmian cen. Dotyczy to takich składników majątku jak nieruchomości, aktywa realne (np. złoto i antyki), a także akcje i udziały. Wzrost cen składników majątku powoduje wzrost wartości majątku oraz wzrost konsumpcji. Zależności te nie są jednak uwzględnione w standardowym modelu IS-LM.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

to spostrzeżenie. Przeciwstawny – wobec „skrajnie keynesowskiego" – pogląd głosi, że produkcja jest stała przy danych zdolnościach wytwórczych i że zmiany łącznego popytu wywołują wyłącznie zmiany cen. Ten wariant na rysunku 2.10 ilustruje pionowa krzywa łącznej podaży *AS'*. Punkt przecięcia tej krzywej z krzywą łącznego popytu wyznacza poziom produkcji w gospodarce. Wzrost łącznego popytu przesuwa krzywą *AD* w prawo, ale nie zmienia to wielkości produkcji, a jedynie powoduje wzrost cen. Taka ilustracja jest zgodna z poglądem ekonomistów (neo) klasycznych, którzy uważają, że poziom produkcji określany jest wyłącznie przez czynniki leżące po stronie podaży.

Wariant pośredni ilustruje rosnąca krzywa podaży *AS"*. Pokazuje ona, że wzrost popytu powoduje częściowo wzrost produkcji i częściowo wzrost cen, przy czym wrażliwość produkcji na zmiany popytu zależy od stopnia nachylenia krzywej *AS"*. Ten wariant jest obecnie powszechnie akceptowany w głównym nurcie makroekonomii.

Rysunek 2.10. Warianty położenia krzywej łącznej podaży *AS*



Źródło: opracowanie własne.

Połączenie krzywych *AD* i *AS* pozwala rozpatrywać zmiany w gospodarce w kategoriach zmian zagregowanego popytu i zagregowanej podaży. Co więcej, umożliwia to również bezpośrednie uwzględnienie zmian cen i ich wpływu na podaż i popyt, co eliminuje jedno z istotnych ograniczeń występujących w prostym modelu IS-LM. Dlatego też model AS-AD jest często wykorzystywany do wyjaśniania mechanizmu inflacji oraz wpływu inflacji na produkcję, podaż, zatrudnienie i popyt.

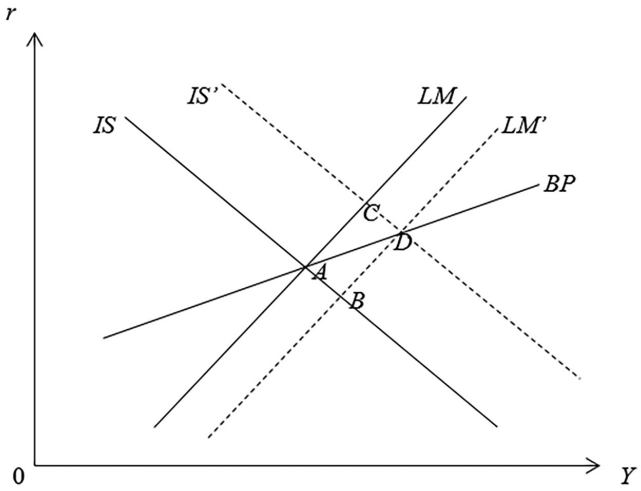
Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

2.8. Polityka pieniężna i polityka fiskalna w modelu Mundella-Fleminga w warunkach różnych systemów kursu walutowego

2.8.1. W systemie stałego kursu walutowego

W ramach modelu Mundella-Fleminga tradycyjny model IS-LM został uzupełniony o analizę równowagi w bilansie płatniczym, czyli równowagi na rynku pieniądza zagranicznego. Przyjęliśmy jednocześnie założenie o stałości kursu walutowego, czyli ceny jednostki pieniądza zagranicznego wyrażonej w walucie krajowej. Dlatego też nie zastanawialiśmy się nad wpływem zmian w równowadze zewnętrznej i wewnętrznej kraju na poziom kursu walutowego, mimo że taki wpływ niewątpliwie występuje. W kolejnych punktach tego rozdziału zbadamy więc skuteczność polityki gospodarczej oraz proces przywracania równowagi makroekonomicznej w warunkach stałego, a następnie płynnego kursu walutowego[19].

Rysunek 2.11. Ekspansja pieniężna i ekspansja fiskalna przy stałym kursie walutowym



Źródło: opracowanie własne.

[19] Szerzej na temat skuteczności polityki fiskalnej i polityki pieniężnej w ramach funkcjonowania różnych systemów kursowych w rozdziale 6.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Załóżmy, że dążąc do zwiększenia produkcji i zatrudnienia władza gospodarcza chce pobudzić koniunkturę przy pomocy polityki makroekonomicznej. Można w tym celu posłużyć się polityką pieniężną lub polityką fiskalną. Przyjmijmy, że kurs walutowy jest stały (sztywny). Zaczniemy od polityki pieniężnej przy założeniu, że bank centralny zwiększa podaż pieniądza (rysunek 2.11). Wzrost podaży pieniądza spowoduje przesunięcie krzywej *LM* do *LM'* i gospodarka przenosi się z punktu równowagi ogólnej *A* wzdłuż krzywej *IS* do punktu *B*. To z kolei wywoła spadek stóp procentowych, odpływ kapitału za granicę i presję na deprecjację waluty krajowej. Pojawia się również deficyt w bilansie płatniczym (ponieważ punkt *B* leży poniżej krzywej równowagi płatniczej *BP*). Broniąc ustalonego kursu waluty, bank centralny musi interweniować na rynku walutowym, sprzedając walutę zagraniczną w zamian za walutę krajową. Przekłada się to na spadek rezerw walutowych, co w konsekwencji powoduje spadek podaży pieniądza krajowego. Krzywa *LM* przesuwa się z powrotem do położenia początkowego, rosną stopy procentowe i gospodarka powraca do punktu równowagi *A*. Okazuje się więc, że w warunkach swobody przepływów kapitału i przy stałym kursie walutowym polityka pieniężna okazuje się nieskuteczna w pobudzaniu aktywności gospodarczej.

Najprostszym sposobem zwiększenia skuteczności polityki pieniężnej jest oczywiście rezygnacja z kursu stałego i upłynnienie kursu walutowego. Wówczas krajowa polityka pieniężna odzyskuje swoją autonomię[20]. Jeśli jednak upłynnienie kursu jest niemożliwe (na przykład w ramach unii walutowej) lub niewskazane, bank centralny może próbować zwiększyć skuteczność polityki pieniężnej, stosując interwencję sterylizowaną. Sprzedając waluty obce w celu obrony danego poziomu kursu walutowego, bank może jednocześnie prowadzić zasilające operacje „otwartego rynku" (operacje *repurchase agreement* – w skrócie „*repo*"), to znaczy kupować papiery wartościowe od banków komercyjnych, zwiększając w ten sposób podaż pieniądza krajowego. Operacji takich nie da się jednak prowadzić w nieskończoność, ponieważ są one kosztowne i obciążają wynik finansowy banku centralnego. Co ważniejsze, nie można też w nieskończoność sprzedawać obcych walut, ponieważ rezerwy walutowe banku centralnego są zawsze ograniczone. Interwencja sterylizowana może być zatem stosowana tylko w ograniczonym zakresie[21].

Innym sposobem zwiększenia skuteczności polityki pieniężnej jest ograniczenie swobody przepływów kapitału, na przykład w postaci opodatkowania tych

[20] Więcej o autonomii polityki pieniężnej w rozdziale 6.

[21] Zagadnienie interwencji walutowych i operacji sterylizacyjnych jest omówione szerzej w rozdziale 6.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przepływów (tzw. podatek Tobina) lub nałożenia obowiązku utrzymywania nieoprocentowanych rezerw w banku centralnym. Doświadczenie pokazuje jednak, że tego rodzaju restrykcje są na ogół mało skuteczne, ponieważ po pewnym czasie uczestnicy rynku znajdują sposoby, aby te ograniczenia omijać[22]. Wreszcie skrajnym rozwiązaniem problemu nieskuteczności polityki pieniężnej jest zawieszenie wymienialności waluty krajowej. Pozwala to wprawdzie odzyskać autonomię w polityce pieniężnej, lecz nie jest to rozwiązanie zalecane z uwagi na poważne koszty związane z utratą wiarygodności międzynarodowej kraju oraz nieuchronne zniekształcenia rynkowe, jakie rodzi brak wymienialności.

Przejdźmy teraz do polityki fiskalnej. Załóżmy, że nadal obowiązuje system stałego kursu i że, dążąc do zwiększenia aktywności gospodarczej, władza gospodarcza zwiększa wydatki rządowe (rysunek 2.11). Krzywa *IS* przesuwa się do położenia *IS'*, gospodarka przesuwa się z punktu *A* do punktu *C* i rosną stopy procentowe (przy danym poziomie podaży pieniądza). Napływa kapitał zagraniczny, pojawia się nadwyżka w bilansie płatniczym i rośnie presja na aprecjację waluty krajowej. Broniąc ustalonego kursu waluty, bank centralny skupuje waluty obce na rynku walutowym, płacąc również walutą krajową. Rosną rezerwy walutowe, zwiększa się podaż pieniądza krajowego, i w rezultacie krzywa *LM* przesuwa się w prawo i w dół. Wzrost podaży pieniądza będzie trwał do momentu, aż krzywa *LM* przetnie nową krzywą *IS'* w punkcie *D*, czyli na linii równowagi płatniczej *BP*. Wówczas kapitał przestaje bowiem napływać (ponieważ nowy poziom stóp procentowych zapewnia wyrównanie dochodowości aktywów finansowych w kraju i za granicą po uwzględnieniu ryzyka kursowego) i gospodarka ustabilizuje się w nowym punkcie równowagi ogólnej *D*. Punkt *D* odznacza się wyższym poziomem produkcji (i stóp procentowych) niż punkt *A*. Okazuje się zatem, że – w przeciwieństwie do polityki pieniężnej – polityka fiskalna jest skuteczna w pobudzaniu aktywności gospodarczej w warunkach swobody przepływów kapitału i przy stałym kursie walutowym.

Posługiwanie się wyłącznie polityką fiskalną w celu pobudzenia aktywności gospodarczej jest jednak kosztowne, ponieważ z reguły powoduje wzrost stóp procentowych, a to może wywołać spadek wydatków inwestycji prywatnych (głównie inwestycji) i ograniczyć, a nawet całkowicie zniwelować efekt wzrostu wydatków rządowych. Z kolei polityka pieniężna – jak wykazano – jest w ogóle nieskuteczna.

[22] Świadczą o tym doświadczenia krajów, które stosowały restrykcje w przepływie kapitału, takich jak Chile czy Islandia – zob. rozdział 6.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W tych warunkach pojawia się uzasadnienie dla jednoczesnego zastosowania instrumentów polityki fiskalnej i polityki pieniężnej – a więc zastosowania określonej kombinacji polityki fiskalnej i polityki pieniężnej (*policy-mix*). Jednoczesne zwiększenie wydatków rządowych i podaży pieniądza przy zastosowaniu odpowiedniej skali powoduje przesuwanie się krzywych *IS* i *LM* w prawo, przy czym punkt ich przecięcia wędruje wzdłuż linii *BP*. W efekcie gospodarka przesuwa się z punktu *A* do punktu *D* wzdłuż linii *BP*, bez konieczności przechodzenia przez punkt *C*. Efekt końcowy jest więc ten sam, co przy zastosowaniu jedynie polityki fiskalnej, lecz koszty przejścia są niższe. Ponadto inny jest mechanizm kreacji pieniądza. W przypadku polityki fiskalnej źródłem wzrostu podaży pieniądza jest napływ kapitału zagranicznego i wzrost rezerw walutowych w banku centralnym, podczas gdy przy zastosowaniu kombinacji polityki pieniężnej i polityki fiskalnej źródłem wzrostu podaży pieniądza staje się wzrost aktywów krajowych banku centralnego w wyniku zakupu papierów wartościowych od banków komercyjnych.

2.8.2. W systemie płynnego kursu walutowego

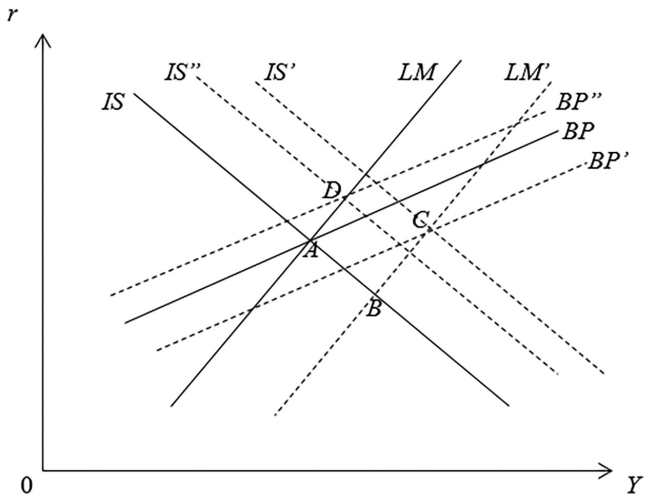
Załóżmy, że władza gospodarcza dąży do pobudzenia gospodarki przy pomocy ekspansywnej polityki pieniężnej w warunkach płynnego kuru walutowego. Bank centralny zwiększa podaż pieniądza i w efekcie krzywa *LM* przesuwa się w prawo do *LM'* (rysunek 2.12). To z kolei wywołuje spadek stóp procentowych, co przekłada się na wzrost inwestycji oraz wzrost produkcji z punktu *A* do punktu *B*. Jednocześnie spadek stóp procentowych powoduje jednak odpływ kapitału za granicę. Pojawia się deficyt w bilansie płatniczym i kurs walutowy ulega deprecjacji. W efekcie rośnie eksport, maleje import i rośnie produkcja krajowa. Deprecjacja oraz spowodowany nią wzrost popytu powoduje przesunięcie krzywych *IS* i *BP* w prawo, a gospodarka przesuwa się do nowego punktu równowagi *C*. Punkt ten cechuje się wyższym poziomem produkcji niż punkt *A*. Polityka pieniężna okazuje się w tych warunkach skuteczna przy pobudzaniu produkcji.

Przejdźmy teraz do polityki fiskalnej. Załóżmy, że nadal obowiązuje system płynnego kursu i że, dążąc do zwiększenia aktywności gospodarczej, władza gospodarcza zwiększa wydatki rządowe (rysunek 2.12). Krzywa *IS* przesuwa się w prawo do położenia *IS'*, rośnie produkcja oraz stopy procentowe (przy danym poziomie podaży pieniądza). Napływa kapitał zagraniczny, pojawia się nadwyżka w bilansie płatniczym i waluta krajowa ulega aprecjacji. Powoduje to spadek eksportu netto i spadek produkcji krajowej, a dodatkowo maleją też inwestycje krajowe na skutek

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wzrostu stóp procentowych. Krzywa *IS* przesuwa się z powrotem w lewo, na przykład do położenia *IS*". W tych warunkach zachowanie równowagi płatniczej wymaga wyższego poziomu stóp procentowych dla każdego danego poziomu dochodu, tak aby dodatkowy napływ kapitału zrównoważył pogorszenie salda rachunku bieżącego na skutek aprecjacji. W efekcie krzywa *BP* przesuwa się w lewo, do położenia *BP*". Przecięcie nowej krzywej *BP*" i krzywej *IS*" w punkcie *D* wyznacza nowy punkt równowagi, w którym nadwyżka płatnicza znika. Polityka fiskalna okazuje się w tych warunkach umiarkowanie skuteczna: poziom produkcji i dochodu w punkcie *D* jest wyższy niż w początkowym punkcie *A*, chociaż niższy niż w punkcie przecięcia krzywej *IS*' i krzywej *LM*'.

Rysunek 2.12. Ekspansja pieniężna i ekspansja fiskalna przy płynnym kursie walutowym



Źródło: opracowanie własne.

Tak jak w przypadku stałego kursu walutowego, także i w systemie kursu płynnego pojawia się uzasadnienie dla jednoczesnego zastosowania instrumentów polityki fiskalnej i polityki pieniężnej (*policy-mix*). Zwiększanie podaży pieniądza równolegle ze wzrostem wydatków rządowych zapobiega wzrostowi stóp procentowych, co pozwala osiągać większe przyrosty produkcji. Jednoczesne zastosowanie polityki fiskalnej i pieniężnej przesuwa krzywe *IS* i *LM* w prawo (rysunek 2.12), przy czym punkt ich przecięcia może przesuwać się wzdłuż linii *BP* lub – gdy polityka pieniężna jest bardziej ekspansywna – poniżej linii *BP*. W efekcie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gospodarka przesuwa się z punktu *A* do punktu *D* bez konieczności przechodzenia przez punkt *C* lub *B*.

2.9. Reguła efektywnego przyporządkowania

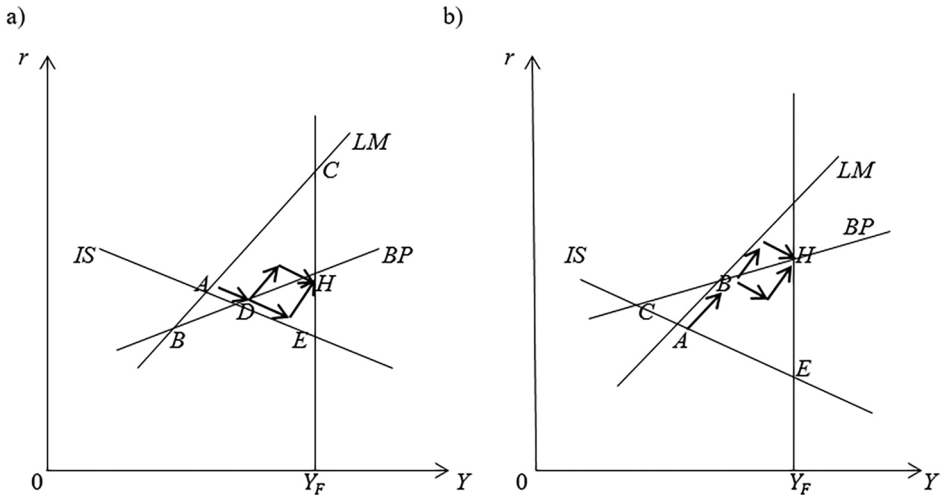
W poprzednim akapicie wprowadzono pojęcie *policy-mix*, czyli jednoczesnego stosowania polityki fiskalnej i polityki pieniężnej, służącego przywracaniu równowagi makroekonomicznej, ale nie powiedziano nic na temat wzajemnych proporcji i skali oddziaływania obu rodzajów polityki. Mundell (1960) pierwszy zwrócił uwagę, że nie jest to bynajmniej obojętne i sformułował prostą zasadę, w myśl której do realizacji każdego celu polityki powinien być zastosowany instrument wywierający na ten cel relatywnie największy wpływ. Zasada ta, którą można nazwać regułą efektywnego przyporządkowania, została zilustrowana na rysunku 2.13.

Załóżmy, że mamy system stałego kursu walutowego. Rysunek 2.13 (część a) prezentuje sytuację w gospodarce z bezrobociem i z nadwyżką płatniczą (punkt *A*). Ograniczenie wydatków rządowych (przesunięcie krzywej *IS* w dół wzdłuż krzywej *LM*) pozwala przywrócić równowagę zewnętrzną (płatniczą) w punkcie *B*, ale kosztem wyraźnego pogłębienia nierównowagi wewnętrznej (spadek produkcji i dalszy wzrost bezrobocia). Wzrost wydatków rządowych (przesunięcie krzywej *IS* w górę wzdłuż krzywej *LM*) pozwala przywrócić równowagę wewnętrzną w punkcie *C*, ale kosztem wyraźnego pogłębienia nierównowagi zewnętrznej (dalszy wzrost nadwyżki płatniczej). Z kolei wzrost podaży pieniądza (przesunięcie krzywej *LM* w prawo wzdłuż krzywej *IS*) pozwala albo przywrócić równowagę zewnętrzną w punkcie *D* przy utrzymującej się nadal nierównowadze wewnętrznej (mniejszy wzrost podaży pieniądza), albo przywrócić równowagę wewnętrzną w punkcie *E* kosztem pojawienia się deficytu w bilansie płatniczym (większy wzrost podaży pieniądza). Z kolei ograniczenie podaży pieniądza (przesunięcie krzywej *LM* w lewo) jednoznacznie pogłębia obie nierównowagi. Wynika z tego, że przy pomocy tylko jednego instrumentu polityki gospodarczej (polityka fiskalna albo polityka pieniężna) nie uda się osiągnąć jednocześnie dwóch różnych celów (równowaga zewnętrzna i równowaga wewnętrzna). Wniosek ten jest oczywiście spójny z omówioną już wcześniej zasadą Tinbergena, zgodnie z którą realizacja dwóch niezależnych celów wymaga zastosowania co najmniej dwóch niezależnych instrumentów[23].

[23] Zob. rozdział 1.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 2.13. Przywracanie równowagi ogólnej przy pomocy polityki fiskalnej i polityki pieniężnej. Reguła efektywnego przyporządkowania.



Źródło: opracowanie własne.

Ustaliwszy, że istnieje potrzeba jednoczesnego zastosowania polityki fiskalnej i polityki pieniężnej, musimy następnie rozstrzygnąć kwestię „przyporządkowania" (*assignment*) poszczególnych instrumentów do danych celów. Z rysunku 2.13 (część a) wynika, że równowaga zewnętrzna może być osiągnięta albo przy pomocy restrykcji fiskalnych (przesunięcie krzywej *IS* w lewo, z *A* do *B*), albo przy pomocy ekspansji pieniężnej (przesunięcie krzywej *LM* w prawo, z *A* do *D*). Drugie rozwiązanie jest oczywiście lepsze, ponieważ równowaga zewnętrzna jest osiągana mniejszym kosztem, gdyż skala nierównowagi wewnętrznej w tym wariancie jest mniejsza. Polityka pieniężna daje więc bezwzględnie lepsze rezultaty w przywracaniu równowagi zewnętrznej niż polityka fiskalna. Z kolei równowaga wewnętrzna (czyli pełne zatrudnienie) może być osiągnięta albo przy pomocy ekspansji fiskalnej (przesunięcie krzywej *IS* w prawo, z *A* do *C*), albo przy pomocy ekspansji pieniężnej (przesunięcie krzywej *LM* w prawo, z *A* do *E*). W pierwszym przypadku pojawia się nadwyżka płatnicza, w drugim zaś deficyt. Trudno jest z góry określić, który wariant polityki okaże się korzystniejszy – zależy to od tego, jaka jest skala nierównowagi płatniczej w obu wariantach.

Rysunek 2.13 (część a) prezentuje jednak tylko przykładowy układ poszczególnych krzywych. Przy innej konfiguracji preferencje w przypadku poszczególnych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

polityk mogą być różne[24]. Generalnie, nie można ograniczyć się do zastosowania tylko jednego instrumentu (ponieważ nie zostanie wówczas osiągnięta ogólna równowaga), a zatem przychodzi zdecydować, która polityka daje względnie lepsze rezultaty w odniesieniu do poszczególnych celów. Jeśli chcemy osiągnąć najpierw równowagę *zewnętrzną*, to w punkcie wyjścia *A* musimy oczywiście zastosować politykę pieniężną, która pozwala osiągnąć równowagę zewnętrzną i jednocześnie przybliża nas do stanu równowagi wewnętrznej (podczas gdy zastosowanie polityki fiskalnej oddala nas od tego celu). Znalazłszy się w punkcie *D*, i chcąc dojść do punktu równowagi ogólnej *H*, musimy następnie zastosować jednocześnie (lub naprzemiennie) politykę fiskalną i politykę pieniężną, ponieważ obie przybliżają nas stopniowo do równowagi wewnętrznej i obie pozwalają utrzymać się na linii równowagi zewnętrznej *BP*. Wynika stąd, że jeśli polityka pieniężna jest jednoznacznie skuteczniejsza w osiąganiu równowagi zewnętrznej, polityka fiskalna będzie względnie bardziej skuteczna w osiąganiu równowagi wewnętrznej.

W podobny sposób można analizować politykę przywracania równowagi ogólnej w gospodarce z bezrobociem i z deficytem płatniczym. Sytuację tę ilustruje punkt *A* na rysunku 2.13 (część b). Ekspansja fiskalna (przesunięcie krzywej *IS* w prawo, z *A* do *B*) pozwala odzyskać równowagę zewnętrzną, lecz nadal występuje bezrobocie. Restrykcje fiskalne oczywiście powiększają obie nierównowagi. Z kolei restrykcje pieniężne (krzywa *LM* przesuwa się w lewo, z *A* do *C*) pozwalają odzyskać równowagę zewnętrzną, ale za cenę dalszego pogłębienia się nierównowagi wewnętrznej. Wreszcie ekspansja pieniężna (krzywa *LM* przesuwa się w prawo, z *A* do *E*) pozwala osiągnąć równowagę wewnętrzną, za cenę jednak pogłębienia nierównowagi zewnętrznej.

Jak w tych warunkach rozwiązać kwestię „przyporządkowania" instrumentów polityki poszczególnym celom? Z rysunku 2.13 (część b) wynika, że równowaga zewnętrzna może być osiągnięta albo przy pomocy ekspansji fiskalnej (przesunięcie krzywej *IS* w prawo), albo przy pomocy restrykcji pieniężnych (przesunięcie krzywej *LM* w lewo). Pierwsze rozwiązanie jest oczywiście lepsze, ponieważ równowaga zewnętrzna jest osiągana mniejszym kosztem (mniejsza skala nierównowagi wewnętrznej). W tych warunkach polityka fiskalna okazuje się bezwzględnie skuteczniejsza niż polityka pieniężna. Z kolei równowaga wewnętrzna

[24] Przykładowo, gdy krzywa *BP* jest bardziej stroma niż krzywa *LM*, przywracanie równowagi wewnętrznej przy pomocy polityki fiskalnej okazuje się korzystniejsze, ponieważ wiąże się z mniejszą skalą nierównowagi zewnętrznej. Z kolei w celu przywrócenia równowagi zewnętrznej lepiej jest zastosować politykę pieniężną, która w mniejszym stopniu przekłada się na stan nierównowagi wewnętrznej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

może być osiągnięta albo przy pomocy ekspansji fiskalnej (przesunięcie krzywej *IS* w prawo), albo za sprawą ekspansji pieniężnej (przesunięcie krzywej *LM* w prawo). W pierwszym przypadku deficyt płatniczy zamienia się w nadwyżkę, a w drugim ulega pogłębieniu. Trudno jest z góry określić, który wariant polityki jest korzystniejszy – zależy to od tego, jaka będzie skala nierównowagi płatniczej w obu wariantach.

Tak jak poprzednio, nie można ograniczyć się do zastosowania tylko jednego instrumentu (ponieważ nie zostanie osiągnięta ogólna równowaga). Należy więc zdecydować, która polityka daje względnie lepsze rezultaty w odniesieniu do poszczególnych celów. Jeśli chcemy najpierw osiągnąć równowagę zewnętrzną, to w punkcie wyjścia *A* musimy oczywiście zastosować politykę fiskalną, która pozwala osiągnąć równowagę zewnętrzną i jednocześnie przybliża nas do równowagi wewnętrznej. Znalazłszy się w punkcie *D*, i chcąc dojść do punktu równowagi ogólnej *H*, trzeba następnie zastosować naprzemiennie politykę fiskalną i politykę pieniężną, ponieważ obie przybliżają nas do równowagi wewnętrznej i obie pozwalają utrzymać równowagę zewnętrzną. Wynika stąd, że jeśli polityka fiskalna jest jednoznacznie skuteczniejsza w osiąganiu równowagi zewnętrznej, polityka pieniężna będzie względnie bardziej skuteczna w osiąganiu równowagi wewnętrznej. Tak więc w przypadku gospodarki z deficytem płatniczym efektywne przyporządkowanie jest odwrotne niż w przypadku gospodarki z nadwyżką płatniczą.

2.10. Przywracanie równowagi ogólnej przy pomocy polityki fiskalnej i polityki pieniężnej w warunkach stałego kursu walutowego*

Standardowa forma modelu IS-LM prezentuje wpływ instrumentów polityki gospodarczej na cele polityki – równowagę wewnętrzną i równowagę zewnętrzną – w sposób pośredni, poprzez wpływ na produkt Y i stopę procentową r. Wygodniej jest przedstawić cele polityki gospodarczej bezpośrednio jako funkcje dwóch instrumentów polityki gospodarczej – wydatków rządowych i podaży pieniądza. Rozszerzymy zarazem definicję równowagi wewnętrznej, łącząc wymóg pełnego zatrudnienia z zachowaniem równowagi na rynku produktów i rynku pieniądza krajowego. Jej warunkiem będzie więc teraz równość zagregowanego popytu i wydatków przy poziomie dochodu, odpowiadającym pełnemu zatrudnieniu, Y_F. Otrzymujemy więc:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$Y_F = C(Y_F) + I(r) - IM(Y_F) + EX + G \tag{2.24}$$

oraz

$$M = L = L(Y_F, r),$$

czyli:

$$r = q(Y_F, M). \tag{2.25}$$

Podstawiając równanie (2.25) do (2.24), otrzymujemy warunek równowagi wewnętrznej:

$$Y_F = C(Y_F) + I\left[q(Y_F, M)\right] - IM(Y_F) + EX + G. \tag{2.26}$$

Przy danym poziomie Y_F i EX równowaga zależy od zmiennych polityki gospodarczej – wydatków rządowych G i podaży pieniądza M. Różniczkując (2.26) dla $dY_F = 0$, otrzymujemy formułę ukazującą zależność pomiędzy zmianami wydatków rządowych a zmianami podaży pieniądza w warunkach równowagi wewnętrznej[25]:

$$\frac{dM}{dG} = -\frac{1}{\left(\frac{\delta I}{\delta r}\right)\left(\frac{\delta r}{\delta M}\right)}. \tag{2.27}$$

Znak wyrażenia (2.27) jest określony przez znak mianownika prawej strony równania. Lecz $\delta I/\delta r < 0$ oraz $\delta r/\delta M < 0$, a zatem (2.27) jest zawsze ujemne. Interpretacja ekonomiczna tego wyniku jest prosta. Wzrost wydatków rządowych przy danym poziomie produkcji Y_F wymaga ograniczenia podaży pieniądza, jeśli równowaga wewnętrzna ma być nadal zachowana.

Z kolei warunek równowagi zewnętrznej możemy zapisać, podstawiając $BP = 0$ do równania (2.22):

$$0 = EX - IM(Y) + B(r). \tag{2.28}$$

[25] Różniczkując wyrażenie $I[q(Y_F, M)] = I(r)$, wykorzystujemy wzór na pochodną funkcji złożonej: $I'(r) = \frac{\delta I}{\delta r}\frac{\delta r}{\delta Y_F}dY_F + \frac{\delta I}{\delta r}\frac{\delta r}{\delta M}dM$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Zarówno dochód Y, jak i stopę procentową r można wyrazić jako funkcję zmiennych polityki gospodarczej – wydatków rządowych i podaży pieniądza. Wówczas równanie (2.28) można zapisać następująco:

$$0 = EX - IM\left[f(G,M)\right] + B\left[q(G,M)\right]. \tag{2.29}$$

Różniczkując (2.29) dla $EX = const$, otrzymujemy formułę ukazująca zależność równowagi zewnętrznej od polityki fiskalnej i polityki pieniężnej:

$$0 = (mg)dG + m\mu)dM + (bg)dG + (b\mu)dM,$$

czyli:

$$dM\mu(m+b) = -dGg(m+b)$$

i ostatecznie

$$dM/dG = -\left(g/\mu\right), \tag{2.30}$$

gdzie $m = \delta IM/\delta Y$ oznacza krańcową skłonność do importu, $g = \delta Y/\delta G$ – krańcowy mnożnik wydatków rządowych, $\mu = \delta Y/\delta M$ – krańcowy wzrost dochodu w wyniku wzrostu podaży pieniądza oraz $b = \delta B/\delta Y$ – krańcową zmianę salda płatniczego na skutek wzrostu dochodu.

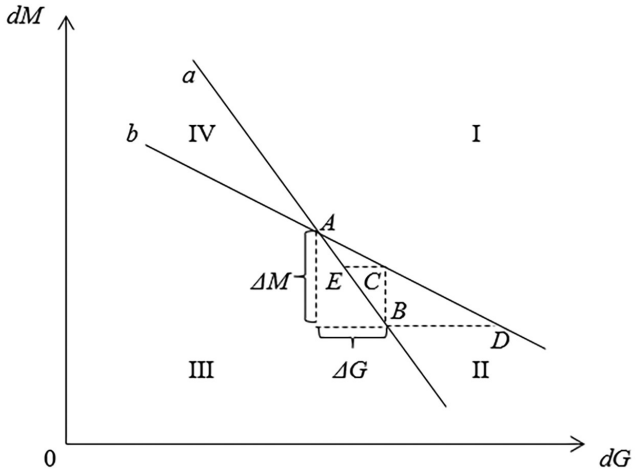
Ponieważ $g > 0$ oraz $\mu > 0$, prawa strona równania (2.30) jest zawsze ujemna. Oznacza to, że gdy rośnie podaż pieniądza, zachowanie równowagi zewnętrznej wymaga spadku wydatków rządowych. Wyjaśnienie jest proste. Wzrost podaży pieniądza powoduje spadek stóp procentowych (dla danego Y), odpływ kapitału za granicę i pojawienie się deficytu płatniczego. Przywrócenie równowagi wymaga spadku importu, co można osiągnąć ograniczając wydatki rządowe. Również odwrotnie, gdy podaż pieniądza maleje, wydatki rządowe muszą wzrosnąć, aby przywrócić równowagę zewnętrzną.

Równania (2.27) oraz (2.30) zostały przedstawione na rysunku 2.14, gdzie na osiach odłożono tym razem wydatki rządowe G i podaż pieniądza M (zob. Mundell 1962). Prosta a reprezentuje równowagę wewnętrzną dla danego poziomu dochodu zapewniającego pełne zatrudnienie (Y_F). Jej nachylenie jest ujemne na mocy równania (2.27), ponieważ wzrost wydatków rządowych zwiększa popyt, a więc utrzymanie równowagi wewnętrznej wymaga ograniczenia podaży pieniądza i wzrostu stóp

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

procentowych. Prosta *b* reprezentuje równowagę zewnętrzną (także dla poziomu dochodu zapewniającego pełne zatrudnienie). Jej nachylenie jest również ujemne na mocy równania (2.30), ponieważ wzrost wydatków rządowych zwiększa import, a więc utrzymanie równowagi zewnętrznej wymaga spadku podaży pieniądza i wzrostu stóp procentowych, aby przyciągnąć kapitał zagraniczny. Zauważmy następnie, że przesuwanie się wzdłuż prostej *a* nie zapewnia utrzymania równowagi zewnętrznej. Choć saldo handlowe pozostaje stałe (bo Y_F nie zmienia się), wzrost (lub spadek) ilości pieniądza przekłada się na spadek (lub wzrost) stopy procentowej, a w konsekwencji odpływ (lub napływ) kapitałów. Gdy przesuwamy się w górę po prostej *a*, nasila się tendencja do powstania deficytu płatniczego, gdy zaś przesuwamy się w dół, pojawia się tendencja do nadwyżki. Mamy więc prawo przypuszczać, że na prostej *a* znajduje się taki punkt, dla którego bilans płatniczy znajduje się w równowadze. Jest to oczywiście punkt przecięcia prostej równowagi wewnętrznej *a* i prostej równowagi zewnętrznej *b*.

Rysunek 2.14. Współzależność polityki pieniężnej i fiskalnej w przywracaniu równowagi ogólnej



Źródło: na podstawie Mundell (1962).

Powyższe rozumowanie pozwala zarazem wykazać, że przy przyjętych założeniach linia *b* musi mieć mniejsze nachylenie niż linia *a*. Załóżmy, że gospodarka znajduje się w punkcie równowagi ogólnej *A* i że władza gospodarcza zwiększa wydatki rządowe o $\Delta G > 0$. Narusza to równowagę wewnętrzną (popyt przewyższa

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

podaż) i, dążąc do jej przywrócenia, władza gospodarcza zmniejsza podaż pieniądza o ΔM, tak aby utrzymać produkcję i dochód na poziomie Y_F. Lecz zmniejszenie podaży pieniądza wywołuje pojawienie się nadwyżki w bilansie płatniczym (ponieważ rosną stopy procentowe) i gospodarka znajduje się w punkcie *B*. Przywrócenie równowagi wewnętrznej idzie więc w parze z powstaniem nadwyżki w bilansie płatniczym (punkt *B* leży na lewo od linii *b*). Zlikwidowanie tej nadwyżki może dokonać się wyłącznie poprzez odpływ kapitału, co wymagałoby zwiększenia podaży pieniądza (i spadku stóp procentowych) ponad poziom określony dla punktu B, na przykład do poziomu określonego punktem *C*. Ale w tym punkcie naruszona jest znowu równowaga wewnętrzna, ponieważ podaż pieniądza jest zbyt duża. Wynika z tego, że prosta równowagi zewnętrznej *b* poniżej punktu równowagi *A* musi leżeć powyżej prostej równowagi wewnętrznej *a*, a zatem nachylenie linii *b* będzie zawsze mniejsze niż nachylenie linii *a*[26]. Oznacza to, że przy tym samym wzroście wydatków rządowych zachowanie równowagi wewnętrznej wymaga większego ograniczenia podaży pieniądza niż zachowanie równowagi zewnętrznej ze względu na reakcję bilansu kapitałowego[27].

Rysunek 2.14 pozwala zidentyfikować cztery odrębne obszary trudności gospodarczych, związanych z różnymi układami nierównowagi wewnętrznej i zewnętrznej. W obszarze I występuje deficyt płatniczy i napięcia inflacyjne. W ramach tego obszaru sytuacja jest jednak na tyle zróżnicowana, że nie można

[26] Przy założeniu, że występuje swoboda przepływu kapitałów skali międzynarodowej. Gdy kapitały są niemobilne, prosta *b* jest linią poziomą.

[27] Alternatywnie, można wykazać, że krzywa *b* ma niższe nachylenie niż krzywa *a*, wykorzystując zależność przepływu kapitałów od stopy procentowej. Załóżmy chwilowo, że kapitał nie reaguje na zmiany stopy procentowej, czyli pozostaje niemobilny w skali międzynarodowej. Wówczas saldo bilansu płatniczego jest wyłącznie określone przez saldo obrotów bieżących, czyli przez zmiany importu dóbr i usług będących funkcją dochodu (ponieważ eksport z założenia jest stały). Zmiany importu są określone przez skłonność do importu, czyli $IM = mY$. Zakładając liniowość tej relacji, wzrost wydatków rządowych o ΔG spowoduje wzrost importu równy $m\Delta G$, i pojawi się tendencja do deficytu w bilansie płatniczym. Aby przywrócić równowagę zewnętrzną, krajowa podaż pieniądza musi spaść o ΔM, co spowoduje wzrost stóp procentowych i spadek pozostałych wydatków. W efekcie spadną także wydatki na import, a ponieważ ogólny poziom dochodu nie zmienia się ($Y = Y_F$), deficyt znika i bilans płatniczy powraca do równowagi. Wynika z tego, że nachylenie krzywej *b* będzie dokładnie równe nachyleniu krzywej *a* i równe stosunkowi $\Delta M/\Delta G$. Załóżmy teraz, że przepływy kapitału są wrażliwe na zmiany stopy procentowej. Wówczas wzrost wydatków rządowych o ΔG, wywołujący wzrost stóp procentowych, spowoduje napływ kapitału z zagranicy, co przełoży się na powstanie nadwyżki w bilansie płatniczym (punkt *B*). W takiej sytuacji ograniczenie podaży pieniądza krajowego może być mniejsze, na przykład $\Delta M^* < \Delta M$ (lub może w ogóle nie być potrzebne), co oznacza, że nachylenie krzywej równowagi płatniczej *b* jest mniejsze niż krzywej równowagi wewnętrznej *a*, ponieważ $\left(\frac{\Delta M^*}{\Delta G}\right) < \left(\frac{\Delta M}{\Delta G}\right)$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jednoznacznie przesądzić, jakie konkretne działania z zakresu polityki gospodarczej są konieczne do podjęcia w celu przywrócenia równowagi. Przykładowo, w lewej górnej części obszaru I konieczne działania obejmują spadek podaży pieniądza i wzrost wydatków rządowych, podczas gdy w prawej dolnej części tego obszaru konieczne okazują się działania odwrotne: ograniczenie podaży pieniądza i spadek wydatków rządowych. Podobnie niejednoznaczna sytuacja panuje w obszarze III, gdzie mamy do czynienia z nadwyżką płatniczą oraz niedoborem popytu i bezrobociem. Z kolei w obszarze II (nadwyżka i inflacja) oraz w obszarze IV (bezrobocie i deficyt) pożądane kierunki działań polityki fiskalnej i pieniężnej są precyzyjnie określone. Nie wystarczy zatem wiedzieć, jakiego rodzaju trudności występują w gospodarce, lecz trzeba również w miarę dokładnie znać aktualne położenie punktu nierównowagi, aby zastosować właściwą kombinację instrumentów polityki fiskalnej i pieniężnej.

Rysunek 2.14 pozwala też zilustrować działanie zasady efektywnego przyporządkowania. Załóżmy, że gospodarka znajduje się na linii równowagi wewnętrznej w punkcie *B* i władza gospodarcza pragnie przywrócić równowagę zewnętrzną za cenę naruszenia równowagi wewnętrznej. Jest oczywiste, że lepszym rozwiązaniem okaże się wykorzystanie polityki pieniężnej i wzrostu podaży pieniądza, co pozwoli przesunąć się do punktu *C*, niż zastosowanie polityki fiskalnej w postaci wzrostu wydatków rządowych i przesunięcie się do punktu *D*. Punkt *C* jest korzystniejszy od punktu *D*, ponieważ równowaga zewnętrzna osiągana jest mniejszym kosztem w postaci naruszenia równowagi wewnętrznej, a ponadto punkt *C* znajduje się bliżej punktu równowagi ogólnej *A*. Podobnie przedstawia się wykorzystanie zasady efektywnego przyporządkowania, gdy gospodarka znajduje się na linii równowagi zewnętrznej, na przykład w punkcie *C*, i władza gospodarcza chce przywrócić równowagę wewnętrzną za cenę naruszenia równowagi zewnętrznej. Lepiej jest w takiej sytuacji zastosować ograniczenia fiskalne i osiągnąć równowagę wewnętrzną w punkcie *E*, niż stosować ograniczenia pieniężne i poszukiwać równowagi wewnętrznej w punkcie *B*.

Zasada Mundella – czyli reguła efektywnego przyporządkowania – pozostaje atrakcyjna dzięki swej prostocie, ale posiada też pewne niedoskonałości. Jest mało prawdopodobne, aby władza gospodarcza zawsze posługiwała się tylko jednym instrumentem, a nie dwoma jednocześnie. Nawet jeśli za politykę pieniężną odpowiada niezależny bank centralny, a za politykę fiskalną rząd (wspólnie z parlamentem), na ogół ma się do czynienia w pewnym zakresie z koordynacją działań obu ramion władzy gospodarczej. Z drugiej jednak strony polityka pieniężna

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jest bardziej elastyczna i może szybciej reagować na pojawiające się zagrożenia. Decyzje o zmianach podaży pieniądza są podejmowane autonomicznie przez władze pieniężne (zarząd i prezes banku centralnego, rada polityki pieniężnej), podczas gdy decyzje z zakresu polityki fiskalnej wymagają na ogół dłuższego procesu politycznego, a często legislacyjnego z udziałem parlamentu. Ta asymetria elastyczności obu rodzajów polityki może powodować zakłócenia w koordynacji i zmuszać do posługiwania się polityką pieniężną w sytuacjach, w których lepiej byłoby zastosować politykę fiskalną, ale ze względu na ograniczenia polityczne nie jest to możliwe.

Ponadto, wrażliwość przepływów kapitału na zmiany stóp procentowych jest uzależniona od wielu czynników i może ulegać gwałtownym zmianom. W pewnych warunkach – zwiększone ryzyko kursowe, wysoki poziom zadłużenia zagranicznego, ograniczenia w wymienialności waluty, niestabilność polityczna – mechanizm wyrównywania deficytu płatniczego poprzez napływ kapitału może się załamać. W takiej sytuacji narzędzia polityki pieniężnej zawodzą i potrzebny jest dodatkowy instrument polityki. Takim instrumentem może być kurs walutowy, który, zmieniając relacje opłacalności między produkcją i konsumpcją towarów krajowych oraz będących przedmiotem wymiany zagranicznej, może wpływać na zmiany poziomu eksportu i importu, nawet przy danym poziomie dochodu. Angielski ekonomista James Meade wykazał w 1955 roku, że wszędzie tam, gdzie zachodzi konflikt między dążeniem do równowagi wewnętrznej a dążeniem do równowagi zewnętrznej, zmiana kursu waluty pozwala na przezwyciężenie tego konfliktu i odzyskanie równowagi ogólnej (Meade 1955).

Rekomendacja ta odnosi się jednak do sytuacji, gdy kraj kontroluje kurs walutowy w ramach stałego lub okresowo zmienianego kursu. W latach 50. i 60. XX wieku ogromna większość krajów na świecie stosowała takie właśnie systemy kursowe. W okresie ostatnich 20–30 lat w gospodarce światowej zaszły jednak zmiany, które skłoniły wiele państw do rezygnacji z systemu kontrolowanego kursu na rzecz kursu płynnego. W warunkach globalizacji, postępującej integracji rynków finansowych i szybko rosnących transgranicznych przepływów kapitału płynne kursy pozwalają utrzymać pożądany zakres autonomii w polityce pieniężnej i skuteczniej chronią przez ryzykiem kryzysów walutowych. W rezultacie polityka kursowa straciła na znaczeniu jako instrument przywracania równowagi ogólnej[28].

[28] Szerzej na ten temat w rozdziale 5.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

2.11. Alternatywne modele makroekonomiczne

2.11.1. Wprowadzenie

Dotychczas poruszaliśmy się w obrębie podstawowego modelu IS-LM. Jak wspomniano, mimo upływu wielu lat model ten nadal może dostarczać pożytecznych informacji w zakresie polityki gospodarczej i cieszy się niezmiennie dużą popularnością w nauczaniu makroekonomii, a także wśród większości ekonomistów podejmujących badania empiryczne i stosowane. Ekonomiści zainteresowani głównie teoretycznymi podstawami funkcjonowania gospodarki od wielu lat zwracali jednak uwagę na liczne ograniczenia modelu IS-LM, takie jak założenie o sztywności cen i płac, ignorowanie oczekiwań wśród uczestników rynku oraz brak przekonującego mikroekonomicznego uzasadnienia dla przyjętych założeń o kształtowaniu się konsumpcji i produkcji[29]. Dążąc do uzupełnienia modelu o brakujące związki i wyeliminowania jego empirycznych oraz teoretycznych defektów, ekonomiści wypracowali w okresie minionego półwiecza szereg alternatywnych modeli makroekonomicznych. Poniżej zostaną omówione najważniejsze z nich, tj. modele syntezy neoklasycznej, modele monetarystyczne i modele racjonalnych oczekiwań, modele realnego cyklu koniunkturalnego, nowe modele klasyczne oraz nowe modele keynesowskie (więcej na ten temat zob. Romer 2000).

2.11.2. Synteza neoklasyczna

Synteza neoklasyczna (*Neoclassical Synthesis* – NS) była nurtem w makroekonomii, w ramach którego podjęto próbę połączenia głównych składników opozycyjnych wobec siebie koncepcji ekonomicznych – keynesizmu i ekonomii neoklasycznej – w obrębie jednolitych, spójnych ram analitycznych. Synteza neoklasyczna, rozwinięta głównie przez wybitnego amerykańskiego ekonomistę Paula Samuelsona w latach 50. i 60., przejęła z keynesizmu przekonanie, że w gospodarce występuje systematycznie przymusowe bezrobocie spowodowane sztywnościami w krótkookresowym dostosowaniu płac i cen oraz że zmiany zagregowanego popytu są głównym czynnikiem wpływającym na produkcję i zatrudnienie w ramach

[29] Chodzi tu m.in. o brak uwzględnienia tendencji do optymalizacji konsumpcji w dłuższym okresie i tendencji do maksymalizacji zysku przez przedsiębiorstwa.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

cyklu koniunkturalnego. Oznaczało to, że klasyczna dychotomia, postulująca brak związku między zmiennymi nominalnymi i zmiennymi realnymi, nie obowiązuje, przynajmniej w krótkim okresie. Z drugiej jednak strony synteza neoklasyczna utrzymywała tradycyjne założenia o istnieniu konkurencyjnych rynków, braku efektów pozarynkowych i doskonałej informacji. Na poziomie mikroekonomicznym gospodarka miała więc funkcjonować zgodnie z walrasowskim modelem zrównoważonych rynków i doskonałej konkurencji[30] (Samuelson 1967; Okun i in. 1969; Okun 1970; Tobin 1972, 1974). Podstawowym narzędziem analitycznym wykorzystywanym w ramach syntezy neoklasycznej był omówiony wcześniej model IS-LM, uzupełniony kluczowym składnikiem w postaci krzywej Phillipsa, która zakładała krótkookresową wymienność pomiędzy inflacją i bezrobociem.

Synteza neoklasyczna zdominowała makroekonomię w latach 50. i 60. ubiegłego wieku. Jej schludna logika załamała się jednak we wczesnych latach 70., kiedy okazało się, że można ją fundamentalnie zakwestionować zarówno na gruncie teoretycznym, jak i empirycznym. Od strony empirycznej synteza neoklasyczna nie była w stanie wyjaśnić zjawiska stagflacji, czyli jednocześnie rosnących stóp inflacji i bezrobocia, co pozostawało sprzeczne z funkcjonowaniem mechanizmu krzywej Phillipsa. Z kolei na gruncie teoretycznym nie dało się usunąć niezgodności zachodzącej miedzy mikroekonomicznymi zasadami konkurencji doskonałej i uwarunkowaniami elastycznych rynków a występowaniem systematycznego bezrobocia, świadczącego o trwałej i nieusuwalnej nierównowadze na rynku pracy. Zwolennicy syntezy neoklasycznej nie byli więc w stanie w sposób przekonujący odpowiedzieć na dość fundamentalne pytanie: dlaczego mianowicie – jeśli rynki rzeczywiście są konkurencyjne (czyli „walrasowskie") – na rynku pracy dochodzi do trwałej nierównowagi?

Krytyka syntezy neoklasycznej obrała generalnie dwa kierunki. Ekonomiści żywiący większy sentyment do myśli neoklasycznej próbowali usuwać niedostatki modeli NS, powracając do głównych paradygmatów ekonomii neoklasycznej, odnoszących się do „walrasowskiej" równowagi na poszczególnych rynkach i klasycznej dychotomii (czyli neutralności pieniądza), ale uwzględniali oni także niektóre elementy konkurencji niedoskonałej. W ramach tego nurtu myślenia, nazwanego „nową syntezą neoklasyczną" (*New Neoclassical Synthesis* – NNS), rozwinęły się takie kierunki jak monetaryzm i teoria realnego cyklu koniunkturalnego. U podstaw

[30] Równowaga „walrasowska" zakłada, że rynki znajdują się w stanie nieustannej równowagi, którą zapewnia doskonała konkurencja i giętkie ceny.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

modeli budowanych w ramach „nowej syntezy neoklasycznej" leżało przekonanie, że w gospodarce generalnie nie występuje przymusowe bezrobocie, a stopa bezrobocia zbliżona jest do stopy naturalnej[31] (poza bezrobociem frykcyjnym). Oznaczało to, że klasyczna dychotomia nadal obowiązuje, a więc zmienne nominalne, takie jak podaż pieniądza, nie mają wpływu na zmienne realne. Ważną nowością modeli NNS było uwzględnienie koncepcji racjonalnych oczekiwań.

Z kolei ekonomiści mający większe przekonanie do koncepcji keynesowskich próbowali doskonalić modele syntezy neoklasycznej, zakładając, że w rzeczywistości gospodarka nie ma cech „walrasowskich" i że mamy w tym przypadku do czynienia z trwałymi odstępstwami od doskonałej konkurencji. Ponadto, ponieważ klasyczna dychotomia nie obowiązuje w krótkim okresie, przy pomocy polityki pieniężnej i polityki fiskalnej można wpływać na poziom produkcji i zatrudnienia. Modele powstałe w ramach tego nurtu nosiły nazwę nowych modeli keynesowskich (*New Keynesian Models* – NKM). Podobnie jak modele NNS, także modele NMK w znacznym stopniu uwzględniały koncepcję racjonalnych oczekiwań (Mankiw 1990, Romer 1993).

2.11.3. Monetaryzm i Nowa Synteza Neoklasyczna

Modele makroekonomiczne w ramach Nowej Syntezy Neoklasycznej (*New Neoclassical Synthesis* – NNS) wykorzystują neokeynesowski mechanizm kształtowania produkcji i dochodu w warunkach sztywnych cen oraz racjonalne oczekiwania i wielookresową optymalizację zachowań uczestników rynku. W ten sposób modele makroekonomiczne są osadzone na mikroekonomicznych podstawach określających decyzje dotyczące cen i produkcji (zaczerpnięte z modeli neokeynesowskich) oraz decyzje odnoszące się do konsumpcji i inwestycji w ujęciu dynamicznym. Nowym elementem w przypadku modeli NNS było uwzględnienie niedoskonałej informacji dotyczącej cen, zaczerpnięte z modeli neoklasycznych oraz modeli realnego cyklu koniunkturalnego (*real business cycle models* – RBC, zob. Lucas 1973). Podobnie jak synteza neoklasyczna modele NNS uwzględniają istnienie krótkookresowej dodatniej zależności między poziomem produkcji i poziomem inflacji, ale wyjaśniają ją, odwołując się do niedoskonałej informacji na temat oczekiwanej inflacji po stronie podaży, co powoduje inercję w zachowaniu się racjonalnych uczestników rynku. W konsekwencji tylko nieoczekiwane skoki inflacji mogą wywołać

[31] Więcej na temat naturalnej stopy bezrobocia w rozdziale 5.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zmiany produkcji. Gdy zmiany inflacji są jednak w pełni przewidywane z punktu widzenia uczestników rynku, polityka staje się nieskuteczna. W rezultacie jednym z zasadniczych wniosków płynących z zastosowania modeli NNS jest sformułowana przez amerykańskich ekonomistów Thomasa Sargenta i Neila Wallace'a hipoteza o nieistotności polityki (*policy irrelevance* – Sargent i Wallace 1975). Zgodnie z jej założeniami, tylko nieantycypowane zmiany w polityce pieniężnej wywierają wpływ na zmienne realne. Co więcej, ponieważ uczestnicy rynku pozostają racjonalni, są oni w stanie w pełni przewidzieć skutki jakiejkolwiek systematycznej (powtarzalnej) polityki. W efekcie regularnie prowadzona polityka nie może wpływać na zmienne realne, takie jak produkcja i zatrudnienie (więcej na temat teoretycznych podstaw modeli NNS zob. Bludnik 2010).

Najgłośniejszym kierunkiem krytyki modeli syntezy neoklasycznej stał się w latach 70. monetaryzm połączony z koncepcją racjonalnych oczekiwań. Nurt ten stanowił wyzwanie w przypadku zarówno oryginalnych koncepcji keynesowskich, jak i syntezy neoklasycznej. Rozbieżność poglądów wywodziła się przede wszystkim z głównych założeń dotyczących nie tylko funkcjonowania gospodarki rynkowej w skali makroekonomicznej, lecz także mechanizmu transmisji impulsów pieniężnych i rekomendacji dla polityki gospodarczej. Związany głównie z nazwiskiem Miltona Friedmana, ale rozwijany również przez innych ekonomistów z tzw. szkoły chicagowskiej (Philip Cagan, Anna Schwartz, Karl Brunner, Alan Metzler, Harry Johnson), monetaryzm z jednej strony kwestionował rolę polityki stabilizacyjnej (zarówno fiskalnej, jak i pieniężnej) w wygładzaniu cyklu koniunkturalnego, a z drugiej przypisywał polityce pieniężnej decydujący wpływ na inflację. Centralny punkt monetaryzmu stanowiła ilościowa teoria pieniądza, zgodnie z którą kluczowe znaczenie w określaniu poziomu dochodu miała nominalna podaż pieniądza i popyt na realne rezerwy gotówkowe, zgodnie z równaniem:

$$Y = MV, \tag{2.31}$$

czyli:

$$log\, Y = log\, M + log\, V, \tag{2.32}$$

gdzie Y to dochód nominalny (czyli Py), M oznacza podaż pieniądza, a V to szybkość obiegu pieniężnego. Zarazem monetaryści nie byli jednak w stanie określić, w jakiej proporcji zmiany dochodu nominalnego Y po lewej stronie równania (2.32) przekładają się na zmiany dochodu realnego y i zmiany cen P. W istocie bowiem

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

monetaryzm w ogóle nie wyjaśniał kształtowania się cen w krótkim okresie, ograniczając się jedynie do podkreślania średnio- i długookresowej zależności zmian cen od ilości pieniądza w obiegu.

Monetaryści jako pierwsi wykorzystali koncepcję oczekiwań inflacyjnych w kontekście polityki stabilizacyjnej, wskazując, że uczestnicy rynku nie pozostają bierni w obliczu zmian polityki, lecz są w stanie racjonalnie przewidywać jej skutki i odpowiednio dostosowywać do nich swoje postępowanie. Badania dotyczące roli i wpływu racjonalnych oczekiwań na skuteczność polityki gospodarczej, zapoczątkowane pracami Johna Mutha, Edmunda Phelpsa i Miltona Friedmana, były następnie kontynuowane i rozwijane m.in. przez Roberta Lucasa, Thomasa Sargenta i Neila Wallace'a. John Muth (1961) wprowadził do teorii makroekonomicznej koncepcję racjonalnych oczekiwań. Była to innowacja przełomowa, choć całkowicie oczywista w świecie gospodarki „walrasowskiej". Rzeczywiście, trudno było zakładać, by uczestnicy rynku z jednej strony zachowywali się racjonalnie, dążąc do maksymalizacji zysku i użyteczności, a z drugiej strony postępowali nieracjonalnie, nie uwzględniając własnych oczekiwań co do przyszłego stanu gospodarki. Z kolei Friedman (1968) jako pierwszy wykazał, w jaki sposób częściowe dostosowanie oczekiwań przyczynia się do krótkookresowej sztywności cen i płac, dzięki czemu polityka pieniężna wpływa w krótkim okresie na sektor realny. Jednocześnie też udowodnił, że w dłuższym okresie oczekiwania dostosowują się w pełni do prowadzonej polityki, co powoduje, że wpływ polityki pieniężnej na produkcję i zatrudnienie szybko zanika, a jedynym trwałym efektem ekspansji pieniężnej staje się rosnąca inflacja.

Uwzględnienie racjonalnych oczekiwań w analizie polityki pieniężnej było prawdziwym przełomem w modelowaniu makroekonomicznym. Wykorzystując tę innowację, Rober Lucas (1972, 1976) w sposób przekonujący wykazał, że w warunkach istnienia racjonalnych oczekiwań nie można z góry przesądzić o wynikach określonej polityki, ponieważ racjonalni uczestnicy rynku dostosowują swoje zachowania do prognozowanych skutków tej polityki (o tzw. krytyce Lucasa wspomniano już w rozdziale 1, więcej na ten temat można przeczytać w rozdziale 5). Z kolei Thomas Sargent (1986) zwrócił w tym kontekście uwagę na znaczenie wiarygodności polityki, argumentując, że koszty polityki dezinflacji w postaci utraconego dochodu i produkcji stają się niższe, gdy zobowiązanie banku centralnego do obniżenia inflacji jest wiarygodne w oczach uczestników rynku. Racjonalne oczekiwania są obecnie standardowym elementem większości modeli makroekonomicznych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

2.11.4. Modele realnego cyklu koniunkturalnego

Drugim ważnym nurtem w ramach nowej syntezy neoklasycznej był rozwój teorii realnego cyklu koniunkturalnego (*real business cycle theory* – RBC). Teoria RBC wyjaśnia zmiany produkcji i dochodu przy pomocy zmian w wydajności czynników produkcji, będących skutkiem przemian dokonujących w zakresie technologii produkcyjnych. Zmiany technologiczne mają w przeważającej mierze charakter losowy. Uczestnicy rynku są racjonalni i dostosowują do tych zmian poziom konsumpcji i podaż pracy. Podstawowe elementy składowe teorii RBC obejmują międzyokresową optymalizację konsumpcji i podaży pracy, określającą zachowanie gospodarstw domowych, oraz międzyokresową optymalizację inwestycji przez firmy maksymalizujące zysk. Decyzje gospodarstw domowych i firm są następnie analizowane w ramach modeli równowagi ogólnej (*general equilibrium models* – GE) w celu określenia wielkości produkcji i cen. Takie ujęcie zostało po raz pierwszy zaproponowane przez amerykańskich ekonomistów Kydlanda i Prescotta (1982).

Zgodnie z teorią RBC, wstrząsy w postaci zmian wydajności czynników produkcji, spowodowane przekształcenami technologicznymi, wpływają na przebieg cyklu koniunkturalnego (Prescott 1986; Plosser 1989). Zapiszmy funkcję produkcji o stałych przychodach względem skali produkcji w następującej postaci:

$$y_t = a_t f\left(n_t, k_t\right), \tag{2.33}$$

gdzie n_t, k_t oznaczają nakłady pracy i kapitału w czasie t, a_t jest zaś całkowitą wydajnością czynników produkcji. Jest to tzw. reszta Solowa, występująca w kanonicznym modelu wzrostu gospodarczego zbudowanym przez amerykańskiego ekonomistę Roberta Solowa (*Solow residual* – zob. Solow, 1957). Różniczkując, otrzymujemy równanie zmian produktu w reakcji na zmiany czynników produkcji i ich wydajności:

$$\frac{dy_t}{y_t} = \left[\beta_n \frac{dn_t}{n_t} + \beta_k \frac{dk_t}{k_t}\right] + \frac{da_t}{a_t}, \tag{2.34}$$

gdzie β_n, β_k oznaczają odpowiednio udziały zasobu pracy i zasobu kapitału w całości produkcji. Wstrząsy w postaci zmian wydajności czynników produkcji wpływają w równaniu (2.34) na zmiany produkcji (dy_t) *bezpośrednio* – poprzez zmiany

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

czynnika $\frac{da_t}{a_t}$ – oraz *pośrednio* – poprzez wpływ na popyt na wykorzystanie zasobów produkcyjnych *n* i *k*, a w efekcie na położenie krzywej produktu krańcowego. Przykładowo, przejściowy wzrost wydajności pracy zwiększa skłonność gospodarstw domowych do pracy (ponieważ obniża relatywną użyteczność czasu wolnego), podczas gdy przejściowy wzrost produktywności kapitału zwiększa skłonność do inwestowania. Wyjaśnia to procykliczność produkcji w ramach cyklu koniunkturalnego – wzrost produkcji spowodowany wzrostem wydajności jest dodatkowo spotęgowany wzrostem skali wykorzystania zasobów (Goodfriend i King 1997).

Badania empiryczne dowodzą, że wyniki zastosowania modeli realnego cyklu koniunkturalnego, w przypadku których jedynym źródłem zaburzeń losowych są wahania produktywności, okazują się w dużym stopniu zgodne z rzeczywistym przebiegiem cyklu (więcej na ten temat zob. Kolasa 2012).

2.11.5. Nowe modele keynesowskie

Krytyka pod adresem tradycyjnych modeli keynesowskich (takich jak na przykład model IS-LM) ze strony monetarystów, przedstawicieli koncepcji racjonalnych oczekiwań i zwolenników teorii realnego cyklu koniunkturalnego skłoniła ekonomistów reprezentujących ten nurt do gruntownej rewizji dotychczasowych koncepcji. Nowe modele keynesowskie (*New Keynesian Models* – NKM), rozwinięte w latach 90., uwzględniają już dynamiczny mechanizm kształtowania się cen (Gordon 1982), wpływ racjonalnych oczekiwań na mechanizm kształtowania się płac (Taylor 1980; Ball i in. 1988) oraz niedoskonałą konkurencję (Blanchard i Kiyotaki 1987; Rotemberg 1987), przy utrzymaniu założenia o sztywności cen i płac nominalnych (Mankiw i Romer 1991; Romer 1993).

Modele NKM miały przede wszystkim odpowiedzieć na pytanie, dlaczego gospodarka rynkowa nie zachowuje się w sposób „walrasowski", to znaczy dlaczego rynek pracy znajduje się w stanie permanentnej nierównowagi. Punktem wyjścia było spostrzeżenie, że bezrobocie może mieć dwojaki charakter. Bezrobocie klasyczne występuje wówczas, gdy nie ma ograniczeń w popycie na produkty, a więc firmy są w stanie sprzedać swoje produkty po danych cenach, a przyczyną bezrobocia jest zbyt wysoki poziom płac realnych. Bezrobocie keynesowskie ma miejsce wówczas, gdy istnieją ograniczenia w popycie przy danych cenach, co zmusza firmy do zwalniania pracowników. Przyczyną bezrobocia może być zatem albo

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zawodność na rynku pracy (bezrobocie klasyczne), albo zawodność na rynku produktów (bezrobocie keynesowskie). W pierwszym przypadku mamy do czynienia ze sztywnością płac, co teoria NKM wyjaśnia istnieniem kontraktów płacowych i rolą związków zawodowych, podczas gdy w drugim przypadku spotykamy się ze sztywnością cen, co teoria NKM wyjaśnia istnieniem konkurencji monopolistycznej i kosztami dostosowań cenowych (Mankiw 1990).

Kontrakty płacowe, zarówno w postaci indywidualnych umów o pracę, jak i w postaci kontraktów zbiorowych, będących rezultatem negocjacji płacowych, ustalają poziom płac nominalnych w określonym przedziale czasu. Oznacza to, że płace nominalne (a więc i realne przy danych cenach) nie reagują – przynajmniej w krótkim okresie – na wstrząsy popytowe i podażowe. Niesie to ważne konsekwencje dla skuteczności polityki gospodarczej. Gdy płace *nominalne* nie reagują na zmiany koniunktury, polityka pieniężna staje się skutecznym narzędziem stabilizowania gospodarki, zmiany podaży pieniądza wpływają bowiem na zmiany płacy *realnej* i w ten sposób oddziałują na poziom zatrudnienia[32].

Z kolei sztywne ceny lub ceny zmieniające się z opóźnieniem (*staggered prices*), wynikające z istnienia niedoskonałej (monopolistycznej) konkurencji, wyjaśniają silny wpływ polityki pieniężnej na zmienne realne w krótkim okresie, ponieważ wzrost podaży pieniądza zwiększa zagregowany popyt realny. Gdy cena przewyższa koszt krańcowy, wzrost popytu skłania monopolistyczną firmę do zwiększania produkcji przy zachowaniu stałej marży (*markup*). Ponadto w warunkach konkurencji monopolistycznej firmy poziom produkcji i zatrudnienia kształtuje się systematycznie poniżej poziomu optymalnego, a poziom bezrobocia – powyżej stopy naturalnej. Tłumaczy to dążenie „krótkowzrocznej" władzy do obniżenia bezrobocia przy pomocy ekspansywnej polityki pieniężnej (Barro i Gordon 1983)[33]. Reprezentatywny dla nurtu NKM model makroekonomiczny rozwinięty przez Claridę, Galì i Gertnera (1999) omówiony zostanie w rozdziale piątym.

[32] W okresie recesji wzrost podaży pieniądza wywołuje wzrost cen i obniża płacę realną.
[33] Zob. punkt 5.6.1.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Literatura cytowana

Acocella N., *Zasady polityki gospodarczej*, Wydawnictwo Naukowe PWN, Warszawa 2002.

Acocella N., *Economic Policy in the Era of Globalisation*, Cambridge University Press, Cambridge 2005.

Ball L., Mankiw N.G., Romer D., *The New Keynesian Economics and the Output-Inflation Trade-off*, "Brookings Papers on Economic Activity" 1988, no. 1.

Barr, R.J., Gordon D., *Rules, Discretion and Reputation in a Model of Monetary Policy*, "Journal of Monetary Economics" 1983, no. 12.

Barro R.J., *Makroekonomia*, PWE, Warszawa 1997.

Begg D., Fischer S., Dornbusch R., *Makroekonomia*, wydanie 3 zmienione, PWE, Warszawa 2003.

Blanchard O., *Makroekonomia*, Oficyna a Wolters Kluwer Business, Warszawa 2011.

Blanchard O.J., Kiyotaki N., *Monopolistic Competition and the Effects of Aggregate Demand*, "American Economic Review" 1987, no. 77(4).

Bludnik I., *Nowa synteza neoklasyczna w makroekonomii*, „Bank i Kredyt" 2010, t. 41, nr 2.

Burda M., Wyplosz Ch., *Makroekonomia. Podręcznik europejski*, wydanie 2 zmienione, PWE, Warszawa 2000.

Clarida R., Galí J., Gertler M., *The Science of Monetary Policy: A New Keynesian Perspective*, "Journal of Economic Literature" 1999, vol. XXXVII, December.

Goodfriend M., King R.G., *The New Neoclassical Synthesis and the Role of Monetary Policy*, "NBER Macroeconomics Annual" 1997, vol. 12.

Gordon R.J., *Price inertia and policy ineffectiveness in the United States, 18901980*, "Journal of Political Economy" 1982, no. 90(6).

Hall R.E., Taylor J.B., *Makroekonomia*, wydanie nowe zmienione, PWE, Warszawa 2000.

Hicks J.R., *Mr. Keynes and the "Classics": A Suggested Interpretation*, "Econometrica" 1937, vol. 5, issue 2.

James T., *The cruel dilemma*, w: *The Battle Against Unemployment*, red. A.M. Okun, W.W. Norton and Company, New York 1972.

Kolasa M., *Teoria realnego cyklu koniunkturalnego*, „Studia i Prace Kolegium Zarządzania i Finansów SGH" 2012, z. 119.

Kydland F.E., Prescott E.C., *Time to Build and Aggregate Fluctuations*, "Econometrica" 1982, vol. 50, no. 6.

Lucas R.E., *Expectations and the Neutrality of Money*, "Journal of Economic Theory" 1972, vol. 4, no. 2.

Lucas R.E., *International Evidence on Output-Inflation Trade-Offs*, "American Economic Review", vol. 63, no. 3.

Lucas R.E. Jr., *Econometric policy evaluation: A critique*, "Carnegie-Rochester Conference Series on Public Policy" 1976, vol. 1, issue 1.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Mankiw G.N., Romer D., *New Keynesian Macroeconomics*, vol. 1–2, The MIT Press, Cambridge 1991.

Meade J., *The Theory of International Economic Policy*, Oxford University Press, Oxford 1955.

Mundell R.A., *The Monetary Dynamics of International Adjustment under Fixed and Flexible Exchange Rates*, "The Quarterly Journal of Economics" 1960, no. 4.

Mundell R.A., *The Appropriate Use of Monetary and Fiscal Policy under Fixed Exchange Rates*, "IMF Staff Papers", March 1962.

Okun A.M., *Inflation: The Problems and Prospects before us,* w: *Inflation: The Problems it Creates and the Policies It Requires*, red. H. Fowler, M. Gilbert, New York University Press, New York 1970.

Okun A.M., Peck M., Smith W., *The 1969 Annual Report of the Council of Economic Advisors*, U.S. Government Printing Office, Washington 1969.

Plosser Ch.I., *Understanding real business cycles*, "Journal of Economic Perspectives" 1989, no. 3(3).

Prescott E.C., *Theory ahead of business cycle measurement,* "Federal Reserve Bank of Minneapolis Quarterly Review" 1986, vol. 10, no. 4.

Romer D., *The New Keynesian Synthesis*, "Journal of Economic Perspectives" 1993, vol. 7, no. 1.

Romer D., *Makroekonomia dla zaawansowanych*, Wydawnictwo Naukowe PWN, Warszawa 2000.

Rosati D., *Polityka proeksportowa*, PWE, Warszawa 1990.

Rose A.K., *A Review of Some of the Economic Contributions of Robert A. Mundell, Winner of the 1999 Nobel Memorial Prize in Economic*, Haas School of Business, Berkeley 2000, http://faculty.haas.berkeley.edu/arose/mundell.pdf (17.01.2015).

Rotemberg J.J., *The New Keynesian Microfoundations*, "NBER Macroeconomics Annual" 1987, vol. 2.

Samuelson P.A., *Economics: An Introductory Analysis*, wydanie 7, McGraw-Hill, New York 1967.

Sargent T.J., *Rational Expectations and Inflation*, Harper and Row, New York 1986.

Sargent T.J., Wallace N., *Rational Expectations. The Optimal Monetary Instrument and the Optimal Money Supply Rule*, "Journal of Political Economy" 1975, vol. 83, no. 2.

Solow R., *Technical change and the aggregate production function*, "Review of Economics and Statistics" 1957, vol. 39(3).

Summers L., *Some Sceptical Observations on Real Business Cycle Theory*, "Federal Reserve Bank of Minneapolis Review" 1986, vol. 10, no. 4.

Taylor J.B., *Aggregate dynamics and staggered contracts*, "Journal of Political Economy" 1980, vol. 88, no. 1.

Tobin J., *The New Economics One Decade Older*, Princeton University Press, Princeton–New York 1974.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozdział 3

POLITYKA FISKALNO-BUDŻETOWA

3.1. Zakres, cele i funkcje polityki fiskalno-budżetowej

3.1.1. Wprowadzenie

Polityka fiskalno-budżetowa, nazywana też w skrócie polityką fiskalną[1], polega na gromadzeniu i wydatkowaniu środków finansowych przez władzę publiczną. Termin „fiskalny" wywodzi się z łacińskiego słowa *fiscus,* które w dosłownym tłumaczeniu oznaczało „koszyk", a służyło na określenie skarbca władzy cesarskiej w starożytnym Rzymie. Odwołując się więc do etymologii tego pojęcia, można uznać, że polityka fiskalna oznacza gospodarowanie pieniędzmi publicznymi. U jej podstaw leży przekonanie, że zorganizowane społeczeństwa dążą w obszarze działalności gospodarczej i społecznej do osiągnięcia pewnych istotnych celów, których podmioty prywatne nie są w stanie zrealizować. Zbiorowości te powołują zatem władzę publiczną, wyposażoną w określone prawem kompetencje, w tym prawo do gromadzenia środków finansowych. Jej zadanie polega na przeznaczaniu tych środków na sprawowanie określonych funkcji społecznych i realizowanie wyznaczonych celów na rzecz całego społeczeństwa.

Jak pamiętamy z rozdziału 1, zgodnie z „triadą musgrawiańską", władza publiczna gromadzi i wydatkuje środki publiczne w celu dostarczania dóbr publicznych (takich jak edukacja, budowa dróg i mostów czy utrzymanie sądów i policji), dokonywania pożądanej społecznie redystrybucji dochodów do najbardziej potrzebujących członków społeczeństwa oraz stabilizowania makroekonomicznego gospodarki

[1] W dalszej części rozdziału będziemy posługiwać się terminem „polityka fiskalna".

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w reakcji na wahania cykliczne i wstrząsy jednorazowe. Zwłaszcza w przypadku tej pierwszej funkcji, wydatki publiczne stanowią ważne uzupełnienie, a niekiedy nawet niezbędny warunek podjęcia przez sektor prywatny działalności inwestycyjnej i produkcyjnej. W tym sensie, jak wskazuje amerykański ekonomista Robert Barro, wydatki publiczne, które mają służyć zapewnianiu ważnych dóbr publicznych, przyczyniają się do zwiększenia krańcowej produktywności kapitału prywatnego, w związku z czym podnoszą efektywność gospodarowania (Barro 1989). Chodzi tu przede wszystkim o takie dobra publiczne jak bezpieczeństwo obywateli i ochrona praw własności, zapewnienie przestrzegania kontraktów i umów oraz dostarczanie dóbr publicznych, których rynek żadną miarą nie byłby w stanie zapewnić (Buchanan 1975; North 1990). Pozytywne skutki w przypadku efektywności ogólnospołecznej mogą też przynieść wydatki na cele stabilizacyjne, dzięki którym gospodarka unika przewlekłych okresów spadku produkcji i dużego bezrobocia (choć wśród ekonomistów od dawna toczy się spór na temat skuteczności takich działań).

Z drugiej strony, gromadzenie środków w celu ich wydatkowania przez sektor publiczny pociąga za sobą koszty ekonomiczne i społeczne. Zgodnie z klasyczną teorią finansów publicznych (zob. np. Samuelson 1954; Barro 1979), wydatki finansowane są podatkami, które stają się źródłem zniekształceń rynkowych (*market distortions*). Każdy podatek, poza dostarczaniem dochodów do budżetu, powoduje ograniczenie skali działalności objętej danym podatkiem. Przykładowo, podatek od zysku ogranicza skalę działalności gospodarczej, podatek obrotowy – sprzedaż, a podatek od płac – podaż pracy (więcej na ten temat w rozdziale 4). Ponadto, sam proces redystrybucji dochodów poprzez podatki wymaga ponoszenia dodatkowych nakładów finansowych na utworzenie i funkcjonowanie administracji publicznej, rejestrację dochodów i identyfikację adresatów redystrybucji, co ogranicza wielkość zasobów, które można wykorzystać na cele produkcyjne. Wszystko to może obniżać efektywność gospodarowania. W związku z tym w ramach polityki fiskalnej zachodzi potrzeba porównywania korzyści płynących z wydatków publicznych z kosztami ich finansowania. Władza publiczna powinna realizować tylko takie wydatki, które przynoszą korzyść społeczną większą niż suma kosztów społecznych wynikających ze zniekształceń spowodowanych podatkami oraz związanych z funkcjonowaniem systemu skarbowego. Dlatego też określenie skali zaangażowania władzy publicznej w funkcjonowanie gospodarki, a zwłaszcza określenie wielkości i przeznaczenia wydatków publicznych, jest jednym z najważniejszych zadań polityki fiskalnej w każdym kraju.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W celu uzyskania lepszej przejrzystości wykładu zagadnienia polityki fiskalno-budżetowej zostały w niniejszej książce podzielone na dwie grupy. Przedmiotem uwagi w tym rozdziale będą przede wszystkim kwestie makroekonomiczne związane z wydatkowaniem środków publicznych, długiem, deficytem oraz zagadnieniami polityki stabilizacyjnej. Z kolei problematyka podatków oraz zagadnienia polityki podatkowej zostaną omówione w rozdziale następnym.

3.1.2. Udział sektora publicznego w gospodarce

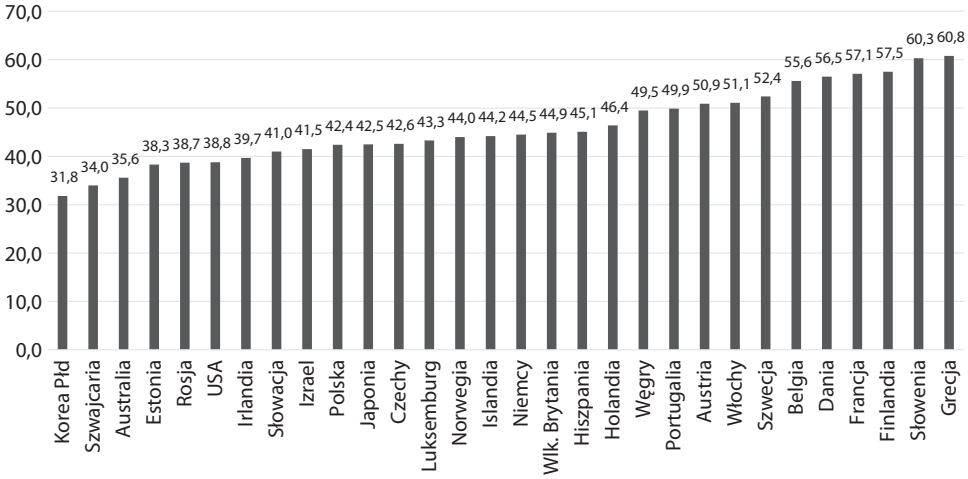
Cele, funkcje i zakres polityki fiskalnej zmieniały się w miarę upływu czasu, wraz z postępującym rozwojem społecznym i gospodarczym, a także w zależności od zmieniających się poglądów na rolę państwa w społeczeństwie. Jeśli uznać, że miarą zakresu polityki fiskalnej jest wielkość środków finansowych gromadzonych i wydatkowanych przez sektor publiczny, to ostatnie stulecie było okresem niespotykanego wcześniej wzrostu znaczenia i zakresu polityki fiskalnej. Jak pokazuje wykres 1.1 (z rozdziału 1), w grupie krajów rozwiniętych udział wydatków (dochodów) sektora publicznego w PKB zwiększył się z średnio 10–12% w końcu XIX wieku i w latach poprzedzających I wojnę światową do około 28% w latach 60. oraz aż do około 45% w połowie lat 90. XX wieku. Tylko w okresie między 1960 a 1990 rokiem udział sektora rządowego w PKB w krajach członkowskich OECD zwiększył się średnio o niemal jedną piątą (o 19,3% PKB – Gwartney i in. 1998). W następnych latach wskaźnik ten zaczął się obniżać, lecz kryzys finansowy i gospodarczy lat 2008–2011 doprowadził do ponownego wzrostu udziału wydatków publicznych w PKB w związku z rosnącym zaangażowaniem rządów w pobudzanie wzrostu gospodarczego i pomoc publiczną dla sektora finansowego, w tym zwłaszcza dla banków.

Za tymi wielkościami średnimi kryją się jednak duże różnice występujące pomiędzy poszczególnymi państwami. O ile w krajach skandynawskich oraz w wybranych krajach Europy kontynentalnej (Francja, Belgia, Austria, Włochy) udział sektora publicznego w PKB przekracza 50%, o tyle średnia w przypadku krajów OECD wynosi niecałe 43%, a w Szwajcarii i w niektórych krajach Europy Środkowo-Wschodniej kształtuje się na poziomie poniżej 40% (wykres 3.1). Jeszcze niższy jest udział sektora publicznego w nowo uprzemysłowionych krajach azjatyckich oraz w krajach słabiej rozwiniętych. Gdy wziąć pod uwagę dwa skrajne przypadki spośród krajów OECD, przychodzi zauważyć, że sektor publiczny w Grecji czy w Słowenii jest dwa razy większy niż w Korei Południowej. Różnice

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

te odzwierciedlają zarówno różny poziom rozwoju gospodarczego, jak i – przede wszystkim – rozbieżne koncepcje udziału państwa w gospodarce oraz odmienne podejście do polityki publicznej w poszczególnych krajach.

Wykres 3.1. Udział wydatków publicznych w PKB w krajach OECD w 2013 roku (dane w % PKB)



Źródło: opracowanie własne na podstawie danych OECD ("Economic Outlook" 2014, no. 97, issue 2).

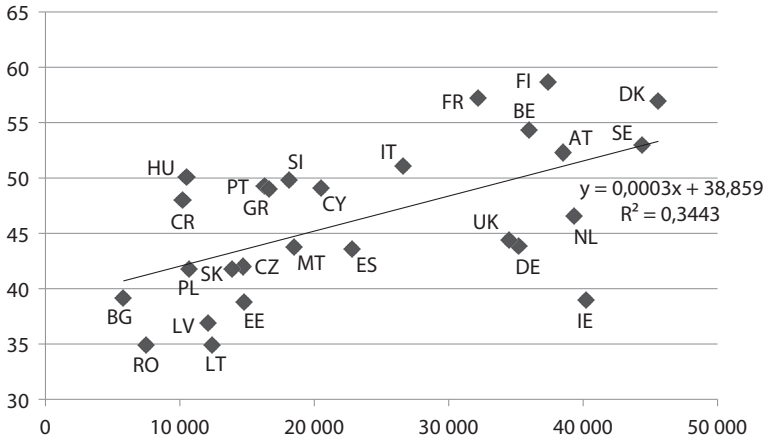
Długotrwała tendencja do wzrostu udziału sektora publicznego w PKB w miarę podnoszenia się poziomu rozwoju gospodarczego bywa nazywana prawem Wagnera, od nazwiska niemieckiego ekonomisty Adolfa Wagnera (1835–1917), który zauważył, że rosnący poziom rozwoju gospodarczego przekłada się na wzrost wydatków publicznych, zarówno bezwzględnie, jak i w relacji do PKB – ze względu na szybciej rosnące potrzeby zbiorowe niż indywidualne. Wykres 3.2 ilustruje sytuację w tym zakresie w krajach Unii Europejskiej w 2014 roku. Jak widać, występuje wyraźna dodatnia korelacja między wielkością PKB na mieszkańca a udziałem wydatków publicznych w PKB (współczynnik korelacji 0,59), co oznacza, że wyższemu poziomowi PKB na mieszkańca towarzyszy przeciętnie wyższy udział wydatków publicznych w PKB. Oszacowane równanie regresji wskazuje, że na każde dodatkowe 1000 EUR wartości PKB na mieszkańca udział wydatków publicznych w PKB wzrasta o około 0,3 punktu procentowego[2]. Jeśli uznać tę relację za punkt odniesienia,

[2] Barro (1989) pokazuje, że zależność ta jest szczególnie silna w przypadku wydatków transferowych, oraz – do pewnego stopnia – w ramach wydatków na edukację. Pozostałe kategorie wydatków

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wykres 3.3 pozwala także zauważyć, że niektóre kraje UE mają zbyt wysoki udział wydatków publicznych w stosunku do osiągniętego poziomu rozwoju gospodarczego (np. Węgry, Chorwacja, Francja czy Finlandia), a z kolei w innych krajach udział ten jest zbyt niski (np. w Irlandii, Łotwie i Rumunii).

Wykres 3.2. Prawo Wagnera w krajach UE[3]: poziom rozwoju gospodarczego (PKB na mieszkańca, EUR '000) a udział wydatków publicznych w PKB w 2014 roku (dane w %)



Źródło: opracowanie własne na podstawie danych Eurostatu.

Opisana tendencja do wzrostu roli sektora publicznego w miarę podnoszenia się poziomu rozwoju gospodarczego państwa była obserwowana niemal przez cały XX wiek. W latach 90. doszło jednak do odwrócenia tego trendu. W okresie obejmującym ostatnią dekadę XX wieku i pierwszą dekadę XXI wieku niemal we wszystkich krajach rozwiniętych udział sektora publicznego w gospodarce wyraźnie się obniżył, przeciętnie spadając o około 4% PKB, w tym w strefie euro średnio o 7% PKB, a w niektórych krajach skandynawskich nawet o 10% PKB. Wykres 3.3 pokazuje, że – poza Irlandią, Słowenią i Kanadą – we wszystkich pozostałych krajach OECD udział wydatków publicznych w PKB w 2007 roku był wyraźnie niższy niż w 1993 roku (niemal wszystkie wartości zmiennej leżą po prawej stronie przekątnej). Największe spadki odnotowano w Słowacji (o 14,3% PKB), Finlandii

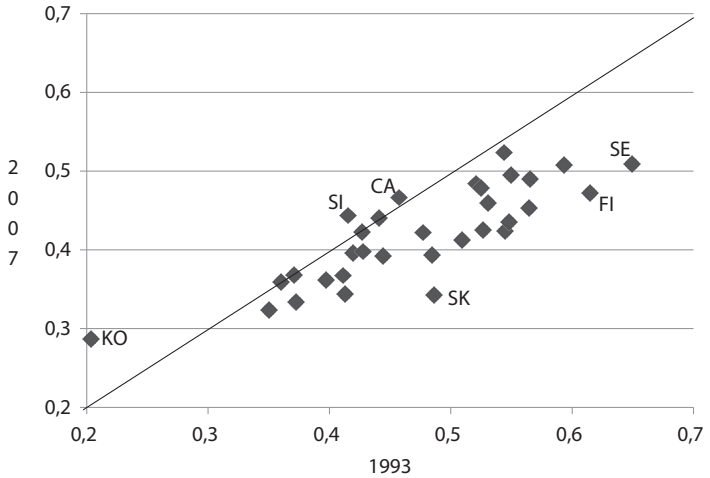
– na inwestycje i obronę narodową – nie wykazują statystycznie istotnego związku z poziomem rozwoju gospodarczego.

[3] UE 27 bez Luksemburga.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(o 14,1% PKB) i Szwecji (o 13,9% PKB), a także w Czechach (o 12,0% PKB), Niemczech (o 11,3% PKB) oraz Holandii (o 11,2% PKB). W Polsce udział wydatków publicznych w PKB spadł o 5,5% PKB (z 47,7% w 1993 roku do 42,2% w 2007 roku).

Wykres 3.3. Zmiany udziału wydatków publicznych w PKB (32 kraje OECD oraz strefa euro i OECD jako całość) w latach 1993–2007 – czy to już koniec prawa Wagnera? (dane w % PKB)



Źródło: opracowanie własne na podstawie danych OECD ("Economic Outlook" 2011, no. 89, issue 1).

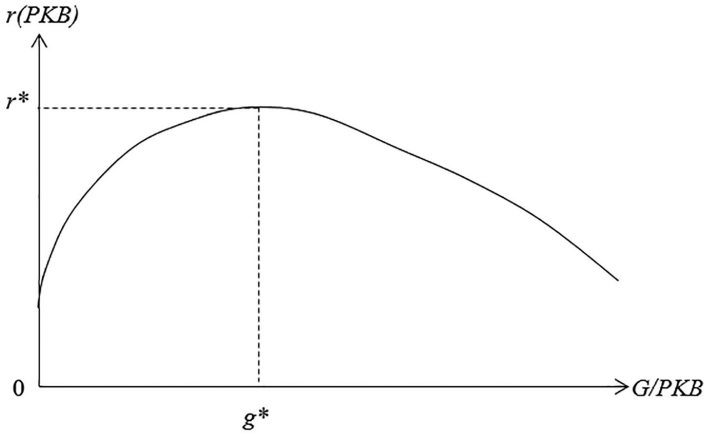
Bezpośrednią przyczyną tej zmiany był wzrost zadłużenia publicznego w krajach rozwiniętych oraz konieczność dostosowania wielkości bieżących wydatków do aktualnych dochodów budżetu. Należy bowiem podkreślić, że systematyczny wzrost wydatków rządowych tylko częściowo był opłacany ze przyrostu dochodów podatkowych – w coraz większym stopniu dodatkowe wydatki były finansowane pożyczkami zaciąganymi przez rządy na rynkach finansowych w ramach rosnącego deficytu budżetowego. Przełożyło się to na zwiększenie zadłużenia, wzrost kosztów obsługi długu publicznego i rosnące obawy o stabilność finansów publicznych w długim okresie.

Jednakże zahamowanie wzrostu, a następnie obniżenie udziału wydatków publicznych w PKB, było spowodowane także coraz powszechniejszym przekonaniem – znajdującym odzwierciedlenie w wynikach wielu badań empirycznych – co do tego, że wzrost udziału sektora publicznego w krajowej gospodarce przynosi maleące korzyści, a zarazem generuje rosnące koszty ekonomiczne. Należą do nich w szczególności wysokie i rosnące podatki, a także marnotrawstwo środków

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i rozrost biurokracji, co powoduje osłabienie tempa wzrostu gospodarczego. Proces zmniejszania się udziału wydatków publicznych w PKB został wprawdzie zahamowany w latach 2008–2011 na skutek kryzysu finansowego i gospodarczego, lecz wiele wskazuje na to, że było to tylko przejściowe zakłócenie nowego trendu i że po przezwyciężeniu kryzysu udział wydatków publicznych w PKB w krajach rozwiniętych będzie się nadal obniżał spadając do poziomu co najmniej przedkryzysowego. Równolegle jednak w krajach o tradycyjnie niskim udziale wydatków publicznych w PKB (rozwinięte kraje pozaeuropejskie), a także wśród państw nowouprzemysłowionych udział ten nie maleje, a wykazuje nadal nawet tendencję wzrostową. W rezultacie rozpiętość wartości pomiędzy poszczególnymi krajami OECD pod względem udziału sektora publicznego w gospodarce uległa wyraźnemu zmniejszeniu. Być może oznacza to, że wraz z upływem czasu udział wydatków publicznych w PKB we wszystkich krajach rozwiniętych będzie konwergował do pewnego wspólnego poziomu, zapewniajacego nabardziej pożądaną relację korzyści i kosztów redystrybucji.

Rysunek 3.1. Krzywa Barro: wpływ zmian udziału wydatków rządowych w PKB (*G/PKB*) na tempo wzrostu gospodarczego (*r*(*PKB*))



Źródło: Barro (1989).

Ekonomiści od dawna spierali się o zakres roli państwa w gospodarce, zadając sobie pytanie, czy istnieje jakiś optymalny poziom udziału sektora publicznego w gospodarce? W latach 80. i 90. powstała spora literatura na ten temat. Wyniki większości badań potwierdzają wcześniejsze obserwacje, że wzrost wydatków

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

publicznych do pewnego poziomu sprzyja wzrostowi gospodarczemu, ale po przekroczeniu tego poziomu przyczynia się do spowolnienia tempa wzrostu gospodarczego, choć nadal występują różnice poglądów co do tego, jaki poziom wydatków należy uznać za optymalny. Przy pewnych założeniach wpływ zmian poziomu wydatków rządowych na tempo wzrostu gospodarczego można przedstawić w postaci funkcji w kształcie odwróconej litery „U": wzrost udziału wydatków w PKB do pewnego stopnia przyspiesza wzrost gospodarczy, a po przekroczeniu tego poziomu hamuje ten trend. W literaturze ekonomicznej funkcja ta nosi nazwę krzywej Barro[4] (zob. rysunek 3.1 oraz ramka 3.1). Jej przebieg pokazuje, że optymalny udział wydatków publicznych w PKB wynosi g^*, ponieważ wówczas tempo wzrostu PKB jest najwyższe (r^*).

Ramka 3.1. Optymalne rozmiary sektora publicznego*

Jak wspomniano wyżej, dostarczanie przez państwo pewnych podstawowych dóbr publicznych zwiększa krańcową produktywność kapitału prywatnego i podnosi efektywność gospodarowania (Barro 1989, Ghali 1998, Ostry i in. 2010). Lecz jednocześnie, gdy wydatki publiczne szybko rosną i wykraczają poza ramy podstawowych funkcji państwa, mogą one obniżać efektywność gospodarowania i hamować rozwój gospodarczy. Powodów takiego stanu rzeczy jest kilka. Po pierwsze, w parze z wysokim poziomem wydatków publicznych idą także z reguły wysokie podatki, a te osłabiają bodźce do podejmowania pracy i działalności gospodarczej. Po drugie, gdy działalność rządu rozszerza się na obszar dostarczania dóbr niepublicznych, inwestycje rządowe mogą w części lub w całości „wypierać" inwestycje prywatne. Efekt „netto" jest w tym wypadku zerowy lub wręcz ujemny. Po trzecie, efektywność wydatków publicznych maleje w miarę ich wzrostu, gdy rząd wychodzi poza zakres swoich podstawowych funkcji i wkracza na płaszczyzny, w których sprawnie działa sektor prywatny. Wreszcie po czwarte, koszt zarządzania działalnością gospodarczą przez rząd jest z reguły wyższy niż koszt zarządzania przez sektor prywatny, i wiąże się z rozrostem administracji publicznej oraz nadmierną regulacją (Gwartney i in. 1998, Bajo-Rubio 2000).

[4] Niekiedy nazywa się ją krzywą BARS, od pierwszych liter nazwisk ekonomistów, którzy wnieśli największy wkład w badanie kwestii optymalnych rozmiarów sektora publicznego (Barro 1989, 1990; Armey 1995; Rahn, Fox i Fox 1996; Scully 1994, 1998). Przegląd tej literatury można znaleźć w pracy Magazzino i Forte (2010).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

To, że z jednej strony pewne wydatki publiczne są niezbędne dla efektywnego funkcjonowania gospodarki, a z drugiej strony ich nadmierny wzrost może prowadzić do obniżenia efektywności, sprawia, że warto postawić pytanie o optymalną wielkość sektora publicznego w gospodarce. Jedną z pierwszych prób wyjaśnienia tej kwestii podjął Robert Barro (1989, 1990). Posłużył się on modelem wzrostu endogenicznego, w którym reprezentatywne gospodarstwo domowe maksymalizuje zdyskontowaną sumę użyteczności konsumpcji w nieskończonym horyzoncie czasu, zgodnie z następującą formułą:

$$U = \int_0^\infty u(c)e^{-\rho t}dt, \tag{3.1}$$

gdzie ρ oznacza (stałą) stopę dyskonta. Z kolei funkcja chwilowej (*momentary*) użyteczności konsumpcji ma postać:

$$u(c) = \frac{c^{1-\theta} - 1}{1 - \theta}, \tag{3.2}$$

gdzie $\theta > 0$ jest (stałą) elastycznością krańcowej użyteczności konsumpcji. Funkcja produkcji jest typu Cobb–Douglas i ma stałe przychody względem skali produkcji. Zakładając, że liczba ludności nie ulega zmianie, otrzymujemy:

$$y = Ak^{1-\alpha}g^{\alpha} = Ak\left(\frac{g}{k}\right)^{\alpha}, \tag{3.3}$$

gdzie y oznacza produkcję na osobę, k – kapitał (fizyczny i ludzki) na osobę, g – realne wydatki rządowe na osobę, A jest stałą reprezentującą postęp technologiczny, zaś $0 < \alpha < 1$. Równanie (3.3) pokazuje, że wydatki rządowe stanowią „czynnik" produkcji w sensie opisanym wyżej jako dostarczanie przez rząd sektorowi prywatnemu ważnych dóbr publicznych. Produkcja y jest sumą konsumpcji, inwestycji (równych przyrostowi kapitału) i wydatków rządowych:

$$y = c + dk + g. \tag{3.4}$$

Przy czym wydatki rządowe są w pełni finansowane podatkami:

$$g = \tau y = \tau Ak^{1-\alpha}g^{\alpha}, \tag{3.5}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gdzie τ jest stopą opodatkowania dochodu. Różniczkując równanie (3.3) względem k otrzymujemy wielkość krańcowej produktywności kapitału:

$$\frac{\delta y}{\delta k} = Ak(\frac{g}{k})^{\alpha}(1-\alpha). \tag{3.6}$$

Równania (3.1) i (3.2) implikują dowód na to, że tempo wzrostu konsumpcji w każdym momencie jest równe:

$$\frac{dc}{c} = \gamma = \frac{1}{\theta}[(1-\tau)\left(\frac{\delta y}{\delta k}\right) - \rho\Big] = \frac{1}{\theta}\Big[(1-\tau)Ak\left(\frac{g}{k}\right)^{\alpha}(1-\alpha) - \rho\Big]. \tag{3.7}$$

Równanie (3.7) pokazuje, że tempo wzrostu konsumpcji jest tym większe, im bardziej wzrasta krańcowa produktywność kapitału pomniejszona o płacony podatek τ, im niższa jest stopa dyskonta ρ oraz im mniejsza staje się elastyczność krańcowej użyteczności konsumpcji θ. Gdy stopa opodatkowania nie zmienia się, wówczas relacja g/y także jest stała, a tempo wzrostu konsumpcji γ nie ulega zmianie. Wówczas również produkcja y i kapitał k rosną w tym samym tempie co konsumpcja – gospodarka jest w stanie wzrostu ustalonego (*steady-state*).

Przejdźmy teraz do zbadania wpływu zmian w poziomie wydatków publicznych na wzrost y. Formuła (3.7) pokazuje, że zmiany wielkości sektora publicznego – g/y – wywierają dwojaki wpływ na tempo wzrostu. Wzrost stawki podatku – τ – obniża tempo wzrostu, a wzrost g w wyrażeniu $\left(\frac{g}{k}\right)^{\alpha}$ przyspiesza tempo wzrostu (na podstawie równania 3.6). Różniczkując formułę (3.7) względem g/y, otrzymujemy:

$$\frac{d\gamma}{d\left(\frac{g}{y}\right)} = \frac{1}{\theta}\cdot Ak\left(\frac{g}{k}\right)^{\alpha}\cdot\left[\left(\frac{\delta g}{\delta k}\right) - 1\right]. \tag{3.8}$$

Znak prawej strony równania (3.8) zależy od znaku wyrażenia w nawiasie kwadratowym. Tempo wzrostu γ rośnie wraz ze wzrostem relacji g/y w przypadku małych wartości g/k, ponieważ wówczas $\frac{\delta g}{\delta k} > 1$, i maleje w miarę wzrostu relacji g/y w odniesieniu do dużych wartości g/k, gdyż wtedy $\frac{\delta g}{\delta k} < 1$. Reasumując, gdy

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

udział sektora publicznego w gospodarce jest niewielki, jego przyrost przyspiesza wzrost gospodarczy, z kolei gdy udział sektora publicznego w gospodarce jest znaczny, dalszy jego rozrost obniża tempo wzrostu. Zależność ta ukazana jest na rysunku 3.1 w postaci tzw. krzywej Barro.

Barro (1989) zastosował swój model do danych przekrojowych dotyczących 72 krajów świata w okresie 1960–1985. Oszacował, że optymalny udział wydatków publicznych w PKB wynosi 25%, przy czym wzrost wydatków powyżej tego poziomu przekłada się na spadek tempa wzrostu gospodarczego. Trzeba jednak zaznaczyć, że zastosowana przez niego definicja wydatków obejmowała tylko te z nich, które mogą bezpośrednio wspomagać aktywność sektora prywatnego. W efekcie z obliczeń wyłączone zostały między innymi transfery społeczne oraz takie usługi publiczne jak edukacja. Jeśli jednak uznamy, że wydatki na edukację i – w pewnym przynajmniej zakresie – na transfery są we współczesnych społeczeństwach niezbędne, próg 25% PKB można uznać za zdecydowanie zaniżony. Otrzymane przez Barro wyniki dowodzą też, że wzrost wszystkich kategorii wydatków publicznych, z wyjątkiem wydatków inwestycyjnych, jest ujemnie skorelowany z tempem wzrostu gospodarczego.

Afonso, Schuknecht i Tanzi (2003), posługując się dokładniejszą miarą wydatków publicznych niż Barro (1989), szacują na próbie 23 uprzemysłowionych krajów OECD, że udział wydatków sektora publicznego w PKB powyżej 30% obniża tempo wzrostu gospodarczego i nie przynosi dodatkowych korzyści w zakresie dobrobytu społecznego. Wniosek ten potwierdzają obserwacje empiryczne. Kraje o „małym" rządzie wykazywały się generalnie lepszymi wynikami w zakresie efektywności gospodarowania niż kraje o „dużym" rządzie.

Chobanov i Mladenova (2009) oszacowali optymalny udział sektora publicznego na podstawie danych przekrojowych dotyczących 28 krajów OECD w okresie 1970–2007 na poziomie 25% PKB, mimo że wzięli pod uwagę całość wydatków sektora publicznego, czyli po uwzględnieniu znacznie szerszej definicji niż Barro (1989).

Z kolei Magazzino i Forte (2010) badali wpływ wydatków rządowych na wzrost gospodarczy wśród 27 krajów Unii Europejskiej w latach 1970–2009. Otrzymane wyniki potwierdzają wcześniejsze rezultaty uzyskane przez innych ekonomistów, choć oszacowany przez nich optymalny poziom wydatków publicznych jest wyższy. Ogólnie rzecz biorąc, większym udziałom sektora rządowego w PKB odpowiadają niższe wartości tempa wzrostu: wzrostowi udziału wydatków średnio o 1% PKB towarzyszy spadek tempa wzrostu o 0,21 punktu procentowego. Autorzy szacują, że optymalny poziom wydatków wśród krajów UE wynosi między 35,6% i 37,3%

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

PKB (w zależności od specyfikacji modelu). Zauważmy, że wszystkie te wyniki sugerują zbyt wysoki poziom wydatków publicznych w krajach UE.

3.1.3. Cele i funkcje polityki fiskalnej

Niezależnie od różnic dotyczących poziomu i zakresu wydatków publicznych w poszczególnych krajach, panuje powszechna zgoda co do konieczności realizowania przez politykę fiskalną pewnych określonych celów (choć brak jest już jednomyślności co do ich szczegółowego zakresu i sposobu realizacji). Jak już wspomniano, w klasycznej nauce o finansach publicznych do celów tych zalicza się najczęściej:

a) dostarczanie tzw. dóbr i usług publicznych (*public goods*), czyli takich produktów, których sektor prywatny z różnych powodów nie jest w stanie zapewnić;
b) redystrybucja dochodów w celu zagwarantowania sprawiedliwego społecznie podziału dochodów;
c) stabilizowanie produkcji i zatrudnienia na wysokim poziomie poprzez regulowanie wielkości zagregowanego popytu.

Równolegle mówi się o też podstawowych funkcjach polityki fiskalnej (Musgrave i Musgrave 1989). Realizując pierwszy z wymienionych celów, polityka fiskalna spełnia funkcję alokacyjną, ponieważ dostarczanie dóbr publicznych w postaci wytwarzania danych produktów i usług wymaga przeprowadzenia określonych inwestycji i poniesienia pewnych kosztów bieżących, związanych z ich funkcjonowaniem. W ten sposób polityka fiskalna wpływa na strukturę alokacji zasobów produkcyjnych w gospodarce. Realizując drugi cel, polityka fiskalna spełnia funkcję dystrybucyjną, kiedy przy pomocy podatków i subsydiów koryguje pierwotny rozkład dochodów, kształtujący się w wyniku działania mechanizmu rynkowego. We współczesnych gospodarkach mechanizm dystrybucji jest wielokierunkowy, co oznacza, że przyjmuje postać transferu dochodów pomiędzy poszczególnymi grupami społecznymi. W szczególności polityka fiskalna polega na nakładaniu podatków na dochody bezpośrednich uczestników procesu produkcji i wymiany, a uzyskane w ten sposób środki zostają rozdysponowane między pozostałych członków społeczeństwa, w postaci płatności transferowych. Wreszcie, realizacja trzeciego z celów, pozwala polityce fiskalnej spełniać funkcję stabilizacyjną w odniesieniu do całej gospodarki. Chodzi tu o stabilizację makroekonomiczną, a więc o zapewnienie makroekonomicznej równowagi i unikanie cyklicznych załamań popytu oraz produkcji. Cele i funkcje polityki fiskalnej przedstawione są w tabeli 3.1.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 3.1. Cele i funkcje polityki fiskalnej

Cele	Funkcje
Dostarczanie dóbr publicznych	Funkcja alokacyjna
Redystrybucja dochodów	Funkcja dystrybucyjna
Stabilizacja makroekonomiczna	Funkcja stabilizacyjna

Źródło: opracowanie własne.

Należy pamiętać, że wymienione funkcje nie pozostają względem siebie w pełni niezależne i częściowo się pokrywają (Owsiak 2006, s. 66–76). Przykładowo, wydatkowanie środków publicznych w celu sfinansowania dóbr publicznych, takich jak edukacja, wymiar sprawiedliwości czy obrona narodowa, oznacza zarazem redystrybucję dochodów na rzecz określonych grup zawodowych, takich jak nauczyciele, sędziowie i prokuratorzy czy zawodowi wojskowi. Z kolei redystrybucja dochodów poprzez opodatkowanie zamożniejszych grup społecznych, odznaczających się wysoką skłonnością do oszczędzania, na rzecz grup uboższych, cechujących się niską skłonnością do oszczędzania, oznacza obniżenie średniej skłonności do oszczędzania w całym społeczeństwie, co w efekcie przekłada się (w warunkach *ceteris paribus*) na wzrost zagregowanego popytu i wydatków. To wzajemne nakładanie się efektów spełniania określonych funkcji oznacza, że poszczególnych gałęzi polityki nie można rozpatrywać w izolacji i trzeba mieć świadomość, że zastosowanie każdej z nich wpływa równocześnie na różne aspekty życia gospodarczego.

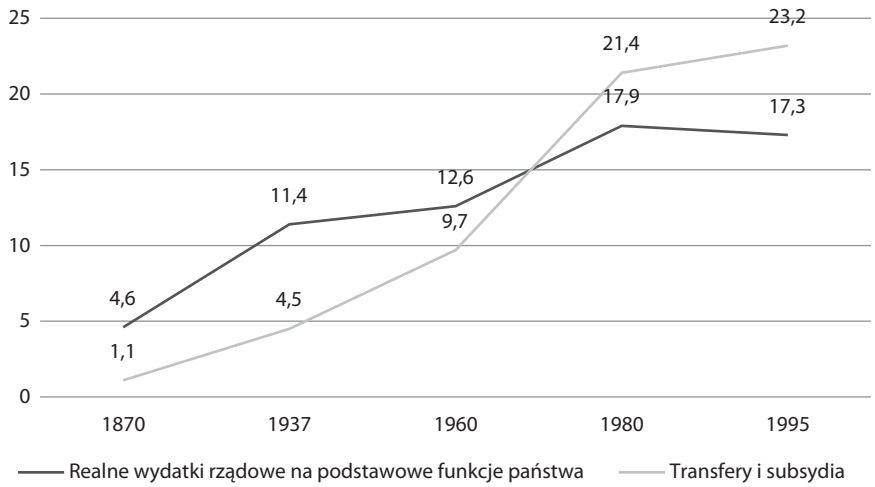
W wąskim znaczeniu politykę fiskalną często utożsamia się z polityką stabilizacyjną lub inaczej – makroekonomiczną. Jej zadaniem jest utrzymanie produkcji, zatrudnienia i dochodów na możliwie wysokim poziomie, tak aby nie występowała presja inflacyjna ani deflacyjna, a bezrobocie było bliskie stopie naturalnej. Inaczej rzecz biorąc, polityka stabilizacyjna powinna dążyć do utrzymania rzeczywistego PKB na poziomie bliskim potencjalnemu PKB. O ile dwie pozostałe funkcje polityki fiskalnej – funkcja dystrybucyjna i funkcja alokacyjna – są realizowane przez wszystkie szczeble władzy publicznej, o tyle funkcja stabilizacyjna jest z reguły założeniem realizowanym przez szczebel centralny władzy publicznej.

Wraz z upływem czasu różnie kształtowało się znaczenie poszczególnych funkcji w ramach całej polityki fiskalnej, a w ślad za tym zmieniały się również proporcje wydatków. W końcu XIX wieku państwo spełniało tylko bardzo ograniczone funkcje, które sprowadzały się zasadniczo do utrzymywania armii i policji, instytucji wymiaru sprawiedliwości oraz administracji państwowej, a więc koncentrowały

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

się na dostarczaniu pewnej, ściśle ograniczonej liczby dóbr publicznych. Było to odzwierciedleniem liberalnej i leseferystycznej koncepcji państwa „minimalnego", pełniącego w gospodarce wyłącznie rolę „stróża nocnego" (*night watchman state*), którego celem jest stanowienie prawa, zapewnianie porządku publicznego oraz bezpieczeństwa wewnętrznego i zewnętrznego. Wydatki publiczne kształtowały się w tym czasie na poziomie około 8–10% PKB. W okresie po I wojnie światowej państwo zaczęło jednak przejmować coraz więcej obowiązków, a rola polityki fiskalnej rosła. Przy czym należy zauważyć, że tradycyjne funkcje państwa (administracja, obrona narodowa, prawo i porządek publiczny) odgrywały relatywnie coraz mniejszą rolę, podczas gdy wzrastało znaczenie „nowych" funkcji, takich jak redystrybucja i wydatki stabilizacyjne.

Wykres 3.4. Zmiany wydatków publicznych w XX wieku (dane w % PKB)



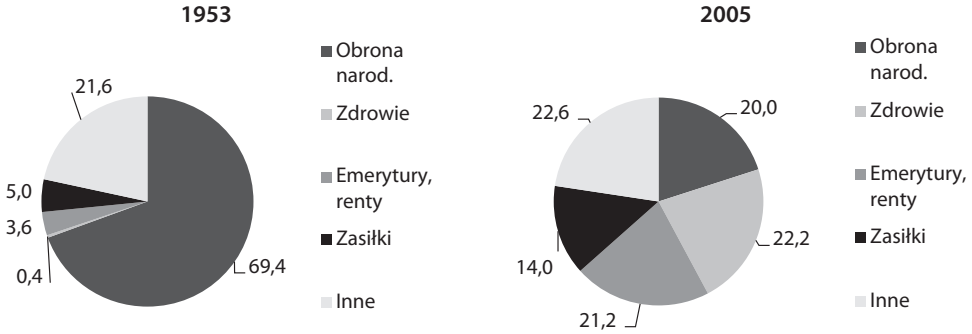
Źródło: na podstawie Tanzi i Schuknecht (2000), s. 25, 33.

Wykres 3.4, zaczerpnięty z pracy Tanzi i Schuknecht (2000), pokazuje, jak rosły wydatki publiczne w XX wieku i zarazem – zwłaszcza w drugiej połowie stulecia – jak zmieniała się ich struktura na korzyść wydatków transferowych, takich jak świadczenia społeczne, subsydia i dotacje. Wydatki te, w reakcji na nasilenie się lewicowych ruchów społecznych i pracowniczych, a także w odpowiedzi na politykę ówczesnego Związku Radzieckiego, wzrosły z 4,5% PKB w latach bezpośrednio poprzedzających II wojnę światową do ponad 23% PKB w drugiej połowie lat 90. XX wieku.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Dobrą ilustracją zmieniających się priorytetów w zakresie wydatków publicznych w krajach rozwiniętych w okresie po II wojnie światowej są zmiany struktury wydatków publicznych rządu federalnego w Stanach Zjednoczonych między 1953 a 2005 rokiem, przedstawione na wykresie 3.5. Jak widać, na przestrzeni pięćdziesięciu lat udział wydatków na tzw. podstawowe funkcje państwa (obrona narodowa, prawo i porządek publiczny, wymiar sprawiedliwości, administracja) w USA spadł z 91,0% do 42,6% ogólnej puli wydatków publicznych, podczas gdy udział środków przeznaczanych na opiekę zdrowotną wzrósł z 0,4% do 22,2%, a udział wydatków transferowych (emerytury, renty zasiłki) z 8,6% do 35,2% całości.

Wykres 3.5. Zmiany w strukturze wydatków rządu federalnego USA między 1953 a 2005 rokiem (dane w %)



Źródło: opracowanie własne na podstawie Gruber (2007), rozdział 1.2.

3.2. Budżet

3.2.1. Podstawowa klasyfikacja pozycji budżetowych

Głównym instrumentem polityki fiskalnej jest budżet państwa, rozumiany jako zestawienie dochodów i wydatków publicznych w określonym czasie, najczęściej w danym roku. Dochody publiczne to głównie dochody z podatków i innych danin publicznych, a także rozmaite dochody niepodatkowe, takie jak dywidendy z zysków realizowanych przez przedsiębiorstwa państwowe czy dochody odsetkowe z lokat utrzymywanych w bankach przez instytucje rządowe. Z kolei wydatki obejmują zakup dóbr i usług na potrzeby sektora publicznego, koszty utrzymania administracji rządowej i samorządowej, wydatki na inwestycje publiczne oraz

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wydatki na obsługę długu publicznego. Różnica między sumą dochodów a sumą wydatków w danym roku stanowi saldo budżetu – gdy jest ono ujemne, mamy do czynienia z deficytem budżetowym lub inaczej z deficytem fiskalnym (*budgetary deficit, fiscal deficit*), kiedy zaś saldo jest dodatnie, powstaje nadwyżka budżetowa (*budgetary surplus*).

Rozróżnia się dwie główne metody tworzenia budżetu. Pierwsza z nich bierze za punkt wyjścia możliwe do uzyskania dochody i do ich wielkości dostosowuje poziom wydatków. Jeśli dochody są niewystarczające do sfinansowania planowanych wydatków, konieczne staje się ich ograniczenie. Ta metoda nazywana jest metodą dochodową. Druga z metod, określana mianem metody wydatkowej, przyjmuje za podstawę konieczne do sfinansowania wydatki publiczne, i to poziom tych wydatków wyznacza niezbędny do ich sfinansowania poziom dochodów (Owsiak 2006, s. 53–54). Jeśli dochody okazują się niewystarczające, należy je dostosować do wymaganego poziomu wydatków (na przykład, podnosząc istniejące podatki lub nakładając nowe). Większy sens ekonomiczny posiada drugie z podejść, pod warunkiem, że uzna się jednak, iż wydatki nie mogą być nieograniczone oraz że realistycznie określi się niezbędne i społecznie uzasadnione potrzeby wydatkowe państwa. Wartość społeczną realizowanych zadań publicznych należy porównać z kosztem ekonomicznym ponoszonym w związku z koniecznością opodatkowania działalności gospodarczej. Zwiększanie zadań publicznych ma bowiem sens tylko wówczas, gdy społeczna korzyść z ich realizacji przewyższa społeczną wartość utraconej produkcji i dochodu w wyniku nałożenia podatków. Więcej na ten temat piszemy w rozdziale 4.

Całokształt środków publicznych składa się z budżetu państwa (budżetu centralnego), budżetów samorządów lokalnych, budżetów funduszy celowych, w tym funduszy ubezpieczeń społecznych, oraz budżetów agencji rządowych, a także innych organizacyjnie wyodrębnionych jednostek budżetowych. Całkowity zasób środków publicznych jest w Polsce nazywany budżetem sektora finansów publicznych – prezentuje on w sposób skonsolidowany dochody i wydatki wszystkich części sektora publicznego. Dane dotyczące głównych pozycji budżetu sektora finansów publicznych w ostatnich latach w Polsce zostały ukazane w tabeli 3.2.

Jak widać, największą pozycją po stronie dochodów publicznych pozostają podatki, których udział w całości dochodów wyniósł w 2014 roku 51%, a w całym PKB 19,7%. Na tle innych krajów rozwiniętych Polska ma stosunkowo niski poziom fiskalizmu – średni udział podatków w PKB w krajach OECD w 2012 roku wyniósł 24,6%. Różnica ta maleje jednak, gdy uwzględni się również składki

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

na ubezpieczenia społeczne, które stanowią swoisty *parapodatek* – wówczas łączny udział danin publicznych w Polsce wzrasta do 32,9% i wynosi niewiele mniej niż średnia w krajach OECD – 33,6%. W całym okresie 2007–2014 udział podatków w PKB w Polsce wykazywał umiarkowaną tendencję spadkową (z 22,9% do 19,7%), z kolei udział składek na ubezpieczenia społeczne nieznacznie wzrósł (z 12,8% do 13,2). Relacja dochodów publicznych do PKB obniżyła się wyraźnie w latach 2009–2010 w wyniku kryzysu finansowego i gospodarczego. Kolejny spadek tej wartości nastąpił w 2013 roku.

Tabela 3.2. Dochody, wydatki i saldo sektora finansów publicznych w Polsce w latach 2007–2014 (dane w mld PLN wg ESA 2010)

Pozycja/rok	2007	2008	2009	2010	2011	2012	2013	2014
1. Dochody	**489,1**	**521,0**	**516,0**	**549,7**	**606,5**	**632,6**	**634,9**	**667,6**
1.1. Podatki	269,8	293,0	274,8	292,8	319,9	323,3	324,5	341,0
1.2. Składki na ubezpieczenia społeczne	151,2	155,9	164,9	169,7	188,9	210,6	219,6	227,5
1.3. Pozostałe dochody	68,1	72,1	76,4	87,2	97,7	98,8	90,8	99,1
Dochody jako % PKB	41,6	41,0	38,4	38,8	39,8	39,1	38,2	38,6
2. Wydatki	**511,1**	**567,4**	**615,7**	**659,4**	**682,6**	**693,1**	**701,6**	**722,9**
2.1. Wynagrodzenia (brutto)	123,8	138,8	150,3	158,6	164,5	167,5	171,3	176,3
2.2. Zużycia pośrednie	71,1	80,1	82,1	93,3	90,1	95,4	97,8	104,0
2.3. Świadczenia społeczne	189,1	203,9	225,3	237,8	243,5	256,9	270,3	281,0
2.4. Odsetki	25,8	27,2	33,6	36,0	39,6	43,2	41,3	33,8
2.5. Inwestycje	53,5	61,3	68,8	80,7	91,3	76,4	68,1	76,3
2.6. Pozostałe wydatki	36,6	43,0	43,3	40,0	41,8	42,1	41,5	41,1
Wydatki jako % PKB	43,4	44,6	45,8	46,6	44,8	42,9	42,2	41,8
3. Wynik sektora	**–22,0**	**–46,4**	**–99,7**	**–109,7**	**–76,2**	**–60,4**	**–66,7**	**–55,2**
Wynik sektora jako % PKB	–1,9	–3,6	–7,3	–7,6	–4,9	–3,7	–4,0	–3,2
4. Memo: wielkość PKB w mld PLN	1176,7	1271,7	1343,4	1415,4	1523,2	1615,9	1662,7	1728,7

Źródło: Ministerstwo Finansów, *Informacja kwartalna o sytuacji makroekonomicznej i stanie finansów publicznych*, nr 2, 2015 oraz dane GUS i obliczenia własne.

Wśród wydatków najwięcej środków publicznych pochłaniają świadczenia społeczne, głównie emerytury, renty i rozmaite zasiłki, których udział w całości wydatków w 2014 roku wyniósł prawie 39%, a w przypadku PKB nieco ponad 16%. Drugą co do wielkości pozycję stanowiły wynagrodzenia w sektorze publicznym (udział w wydatkach i w PKB wyniósł odpowiednio 24,3% i 10,2%). Udział obu tych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kategorii wydatków w PKB nie uległ większym zmianom od 2007 roku. Stabilny pozostał także udział wydatków na bieżące zakupy dóbr i usług, z kolei wydatki inwestycyjne najpierw silnie wzrosły w latach 2007–2011 (z 4,5% do 6,0% PKB), by następnie spaść (do poziomu 4,4% PKB w 2014 roku). Ogólnie rzecz biorąc, wydatki sektora publicznego stanowiły w 2014 roku 41,8% PKB i były niższe niż zarówno te w 2007 roku (43,4% PKB), jak i w kryzysowym 2010 roku (46,6% PKB).

Saldo sektora finansów publicznych było przez cały okres ujemne i zmieniało się zgodnie z przebiegiem cyklu koniunkturalnego – najpierw silnie wzrosło z poziomu –1,9% PKB w ostatnim przedkryzysowym 2007 roku do –7,6% PKB w 2010 roku, a następnie stopniowo spadało do poziomu –3,2% w 2014 roku. Znaczny wzrost deficytu w okresie kryzysu finansowego i gospodarczego był spowodowany zarówno procyklicznym spadkiem dochodów w 2009 roku, jak i antycyklicznym wzrostem wydatków w latach 2008–2010. Zacieśnienie fiskalne (wyraźne ograniczenie tempa wzrostu wydatków), jakie miało miejsce po 2010 roku, pozwoliło na stopniowe odzyskiwanie równowagi budżetowej.

W Unii Europejskiej budżet środków publicznych nazywany jest budżetem instytucji rządowych i samorządowych (*general government*), a jego zakres podmiotowy i przedmiotowy różni się nieco od zakresu stosowanego w Polsce w odniesieniu do sektora finansów publicznych[5].

Proporcje, w jakich środki publiczne dzielone są pomiędzy budżet państwa a budżety pozostałych części sektora publicznego, zależą od rozwiązań ustrojowych

[5] Według Europejskiego Systemu Rachunków Narodowych i Regionalnych ESA 2010 (do 2013 roku stosowano system ESA 1995, od 2013 roku obowiązuje system ESA 2010), sektor instytucji rządowych i samorządowych dzieli się na trzy podsektory (GUS, Warszawa 2000):
a) instytucje rządowe na szczeblu centralnym (*central government*),
b) instytucje samorządowe na szczeblu lokalnym (*local government*),
c) fundusze ubezpieczeń społecznych (*social insurance funds*).

Z kolei według ustawy o finansach publicznych z 2009 roku sektor finansów publicznych w Polsce składa się z następujących części:
a) Skarb Państwa oraz budżet państwa wraz z państwowymi jednostkami budżetowymi,
b) państwowe fundusze celowe,
c) Narodowy Fundusz Zdrowia,
d) agencje wykonawcze oraz instytucje gospodarki budżetowej,
e) pozostałe państwowe osoby prawne,
f) jednostki samorządu terytorialnego oraz ich budżety wraz z samorządowymi jednostkami i zakładami budżetowymi,
g) zakłady i fundusze ubezpieczeń społecznych.

Te różnice zakresowe i terminologiczne zachodzące między klasyfikacją krajową a klasyfikacją ESA 2010 nie są uzasadnione, ponieważ zmniejszają przejrzystość finansów publicznych i utrudniają analizę porównawczą (zob. na ten temat Wernik 2011, s. 18–22).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przyjętych w poszczególnych państwach. W krajach o strukturze federalnej udział budżetu centralnego jest z reguły niższy niż w państwach o strukturze unitarnej, ponieważ większą rolę odgrywają tu budżety władz regionalnych. Pozostaje on również zależny od sposobu finansowania wydatków na zabezpieczenia społeczne, w tym głównie emerytur i rent. W Polsce udział budżetu centralnego stanowi około połowy wszystkich wydatków publicznych.

Budżet sektora finansów publicznych może być rozpatrywany w węższym, technicznym ujęciu jako rachunek (fundusz), na który wpływają wszystkie dochody publiczne i z którego są dokonywane wydatki (Wernik 2011). Jednakże budżet pozostaje także, a nawet przede wszystkim, instrumentem podziału dochodu wytworzonego w danym kraju i w tym sensie pełni bardzo ważną funkcję ekonomiczną i polityczną (Owsiak 2006, rozdział 5). Gromadząc dochody w formie podatków i finansując określone zadania społeczne i ekonomiczne, budżet oddziałuje na skalę i strukturę działalności gospodarczej, rozkład dochodów w społeczeństwie, i ma decydujący wpływ na realizację ważnych celów społecznych. Wielkość oraz struktura dochodów i wydatków w ramach budżetu nie tylko zależą od decyzji politycznych, lecz wiążą się także z daleko idącymi skutkami ekonomicznymi i społecznymi. Z tych właśnie powodów często mówi się, że budżet państwa jest najważniejszym instrumentem polityki gospodarczej.

3.2.2. Równowaga i nierównowaga budżetowa

Gdy w danym okresie dochody budżetowe pozostają równe wydatkom, budżet jest zrównoważony. Wśród współczesnych państw, z różnych powodów, taka sytuacja występuje stosunkowo rzadko. Z reguły ma się do czynienia z sytuacją, w której dochody nie są równe wydatkom, a zatem występuje nierównowaga budżetowa. Saldo budżetu stanowi wówczas różnicę między dochodami i wydatkami, którą można zapisać w postaci następującej tożsamości budżetowej, uwzględniającej podstawowe kategorie dochodów i wydatków:

$$B \equiv (T + NT) - (Cg + TR + Int + Ig), \tag{3.9}$$

gdzie:
B = saldo budżetu,
T = dochody podatkowe,
NT = dochody niepodatkowe,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Cg = konsumpcja publiczna (wydatki bieżące),
TR = transfery (bieżące i kapitałowe),
Ig = inwestycje publiczne (wydatki inwestycyjne),
Int = odsetki od długu publicznego.

Pierwsze wyrażenie w nawiasie po lewej stronie równania (3.9) obejmuje wszystkie dochody budżetu, drugie zaś – wszystkie wydatki. Saldo budżetu *B* może być dodatnie (nadwyżka) lub ujemne (deficyt). Ekonomiczne znaczenie obu tych pojęć nie jest jednak takie same i nie zachodzi między nimi symetria. Wynika to między innymi z różnic w trybie decydowania o wielkości wydatków i wielkości dochodów oraz z politycznych i społecznych konsekwencji zmian w wydatkach i zmian w dochodach. Nadwyżka – choć formalnie jest przejawem braku równowagi budżetowej – nie ogranicza mianowicie swobody polityki fiskalnej, ponieważ w warunkach nadwyżki zawsze istnieje możliwość szybkiego zwiększenia wydatków (na nowe zadania lub na spłatę wcześniejszych zobowiązań), gdyby władza publiczna uznała to za stosowne. W takiej sytuacji zwiększenie wydatków nie wymaga poszukiwania źródeł ich finansowania, ponieważ istniejąca już nadwyżka umożliwia finansowanie dodatkowych wydatków. Trzeba bowiem pamiętać, że pozyskanie dodatkowych dochodów wymaga z reguły podniesienia podatków, co jest zawsze ekonomicznie kosztowne[6] i politycznie niepopularne. Wzrost wydatków w warunkach posiadania nadwyżki budżetowej jest więc łatwiejszy do zrealizowania i może nastąpić szybciej niż wobec braku nadwyżki.

W przypadku występowania deficytu sytuacja jest bardziej skomplikowana. Brak wystarczających dochodów na pokrycie wszystkich zaplanowanych wydatków stwarza konieczność znalezienia dodatkowych źródeł finansowania. Wymaga to z reguły albo podniesienia podatków, albo zaciągnięcia pożyczek (na rynku pieniężnym lub – znacznie rzadziej – w banku centralnym). Różne sposoby finansowania deficytu będą przedmiotem uwagi w dalszej części rozdziału. W tym miejscu należy jedynie podkreślić, że niezależnie od wyboru sposobu pozyskania środków, finansowanie deficytu pozostaje zawsze procesem złożonym i obarczonym ryzykiem. Gdy źródłem pokrycia deficytu mają być dodatkowe podatki, ich pozyskanie w warunkach demokracji wymaga z reguły zgody parlamentu, co zawsze wydłuża

[6] Nałożenie podatku wywołuje – oprócz efektu fiskalnego – spadek skali działalności objętej podatkiem (w wyniku działania efektu dochodowego i efektu substytucji). Koszt utraconej produkcji z tego tytułu stanowi ekonomiczny koszt związany z nałożeniem podatku. Ponadto, stosowanie podatków wymaga nakładów na ich pobór i rozliczanie (jest to administracyjny koszt podatku), co dodatkowo obniża efektywność.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

proces decyzyjny[7]. Co więcej, podnoszenie podatków (lub nałożenie nowych danin publicznych) jest ekonomicznie i politycznie kosztowne. Z kolei zaciąganie pożyczek na pokrycie deficytu będzie zawsze obciążone ryzykiem wynikającym z sytuacji na rynku – potencjalni pożyczkodawcy muszą chcieć pożyczyć pieniądze rządowi, a warunki tych pożyczek powinny być odpowiednio korzystne. W przeciwnym razie może okazać się, że rząd nie jest w stanie sfinansować powstałego deficytu, co niesie za sobą konieczność ograniczenia wydatków lub sięgnięcia po niekonwencjonalne sposoby finansowania potrzeb pożyczkowych. Wszystko to sprawia, że nierównowaga budżetowa w postaci deficytu budżetowego stwarza trudniejsze wyzwanie dla polityki gospodarczej niż nierównowaga w postaci nadwyżki.

Jak wcześniej wspomniano, równowaga budżetowa występuje we współczesnych państwach stosunkowo rzadko. Można stwierdzić, że to nierównowaga budżetowa jest regułą, a nie wyjątkiem, przy czym w ogromnej większości przypadków nierównowaga przyjmuje postać deficytu, raczej, niż nadwyżki. Ta szczególna skłonność do deficytu (*deficit bias*) wynika z działania pewnych szczególnych mechanizmów, które zostaną omówione w następnym punkcie. Zanim do nich przejdziemy, trzeba jednak zaznaczyć, że w ramach współczesnej makroekonomii wcale nie zaleca się utrzymywania równowagi budżetowej w każdym okresie. Wręcz odwrotnie, ze względu na cykliczny charakter funkcjonowania gospodarek rynkowych i związane z tym „falowanie" produkcji, polityka fiskalna powinna mieć właściwości antycykliczne, aby łagodzić wahania koniunktury i utrzymywać wysoki poziom produkcji, dochodu oraz zatrudnienia. Antycykliczny charakter polityki fiskalnej oznacza, że gdy tempo wzrostu gospodarczego spada poniżej tempa potencjalnego, w budżecie powinien występować deficyt, a gdy tempo wzrostu jest wyższe niż tempo potencjalne, w budżecie powinna pojawiać się nadwyżka. W ten sposób polityka fiskalna uzupełnia niedostatek popytu efektywnego w okresach spowolnienia koniunkturalnego i przyczynia się do zmniejszenia popytu efektywnego w okresach zbyt szybkiego wzrostu.

Gdy kolejne cykliczne deficyty i nadwyżki są symetryczne, czyli zbliżone co do wielkości, a różniące się jedynie znakiem, konsekwencją tak prowadzonej polityki antycyklicznej byłby generalnie zrównoważony budżet w średnim horyzoncie

[7] Podnoszenie podatków, o ile nie ma być zwykłą konfiskatą, wymaga zgody podatników. W krajach demokratycznych ich reprezentantem staje się wybieralny Parlament. Fundamentalna zasada *no taxation without representation* jest podstawą systemów podatkowych w większości krajów europejskich już od XIII wieku.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

czasowym, a więc w ramach całego cyklu koniunkturalnego[8]. Zerowe saldo budżetu w długim okresie implikuje z kolei stopniowe obniżenie poziomu długu publicznego aż do zera. Mimo że nie brakuje zwolenników takiego właśnie scenariusza rozwoju, można wykazać, że stan braku zadłużenia nie jest stanem optymalnym w warunkach rozwijającej się gospodarki i że umiarkowane zadłużenie (nieprzekraczające pewnego progowego poziomu relacji długu do PKB) może przyczyniać się do przyspieszenia wzrostu gospodarczego.

3.2.3. Federalizm fiskalny

Polityka fiskalna jest prowadzona na różnych szczeblach władzy publicznej, w tym na szczeblu centralnym, czyli rządowym, oraz na szczeblach regionalnych i lokalnych. Ustalenie pożądanego zakresu koordynacji polityki pomiędzy poszczególnymi szczeblami powinno wynikać z zestawienia kosztów i korzyści związanych z przeniesieniem kompetencji decyzyjnych z niższego szczebla na wyższy szczebel władzy publicznej. Z jednej strony wiąże się to z ograniczeniem swobody decyzyjnej niższego szczebla oraz z gorszym rozpoznaniem potrzeb lokalnych na wyższym szczeblu, z drugiej zaś strony pozwala odnosić korzyści zewnętrzne (*positive externalities*) wynikające z koordynacji i z większej skali działania. Teoria federalizmu fiskalnego dotyczy rozkładu kompetencji fiskalno-budżetowych w ramach sektora publicznego i odpowiada na pytania: czy, w jaki sposób i na podstawie jakich kryteriów powinno się rozdzielać te kompetencje pomiędzy różne szczeble władzy publicznej.

Rozpatrywana w najbardziej podstawowym ujęciu, teoria federalizmu fiskalnego głosi, że szczebel centralny władzy publicznej (rząd) powinien być odpowiedzialny za politykę stabilizacji makroekonomicznej i redystrybucję dochodów do najuboższych członków społeczeństwa (Musgrave 1959; Oates 1972, 1998). W obu przypadkach uzasadnieniem jest występowanie istotnych ograniczeń w wypełnianiu tych funkcji przez niższe szczeble władzy. Pozbawiony własnej polityki pieniężnej oraz kursowej i dysponujący jedynie polityką fiskalno-budżetową, niższy szczebel władzy (na przykład województwo, powiat czy gmina) sprawuje ograniczoną kontrolę nad sytuacją gospodarczą swojego regionu głównie dlatego, że im mniejsze

[8] Jedną z obowiązujących w Unii Europejskiej reguł dyscypliny fiskalnej jest właśnie obowiązek utrzymania przez państwa członkowskie zrównoważonego budżetu w średnim okresie (*Medium-Term Objective* – zob. rozdział 8, punkty 8.8 i 8.9).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jest terytorium, tym większa część wydatków publicznych „wycieka" ostatecznie poza region i tym większy wpływ ma otoczenie gospodarcze. Dlatego też polityka stabilizacyjna na niższym szczeblu władzy okazuje się nieskuteczna. Podobnie, wysoka międzyregionalna mobilność ludzi i kapitału silnie ogranicza możliwości opodatkowania czynników produkcji oraz redystrybucji dochodów. Na przykład relatywnie wysokie opodatkowanie osób uzyskujących wysokie dochody może spowodować ich migrację do innych regionów. Z kolei relatywnie hojne programy wspierania najuboższych mogą doprowadzić do napływu osób o niskich dochodach z innych regionów. W związku z tym polityka redystrybucji sprawdza się lepiej na szczeblu centralnym. Ponadto rząd centralny powinien dostarczać dobra publiczne o charakterze „ogólnonarodowym", a więc takie, które są konsumowane przez całe społeczeństwo. Do dóbr takich należą, przykładowo, obrona narodowa, jednolita waluta czy przepisy ruchu drogowego.

Z kolei niższe szczeble władzy publicznej powinny ponosić odpowiedzialność za dostarczanie dóbr publicznych o charakterze „lokalnym", czyli takich, których konsumpcja jest ograniczona do danego regionu. Władza lokalna jest w stanie lepiej dostosować ilość oraz strukturę tych dóbr do potrzeb i preferencji społeczności lokalnych. Wśród dóbr tego typu wymienia się lokalną infrastrukturę drogową, edukacją na poziomie przedszkolnym i podstawowym bądź też lokalne obiekty sportowe i wypoczynkowe.

Ogólnie rzecz biorąc, jeśli nie pojawiają się oszczędności wynikające ze scentralizowanego systemu dostarczania dóbr publicznych (na przykład w postaci zatrudnienia mniejszej liczby urzędników) oraz nie występują międzyregionalne efekty zewnętrzne (na przykład pozytywne skutki harmonizacji przepisów w skali wielu regionów), to dostarczanie dóbr publicznych w sposób zdecentralizowany (i zakładający realizację zasady Pareto) będzie zawsze społecznie bardziej opłacalne. Ta prawidłowość określana jest mianem twierdzenia o decentralizacji (*Decentralization Theorem* – Oates 1998). Z kolei zasada, zgodnie z którą dostarczanie dóbr publicznych powinno dokonywać się na najniższym możliwym szczeblu władzy, nazywa się zasadą pomocniczości lub subsydiarności (*subsidiarity principle*). Zarówno twierdzenie o decentralizacji, jak i zasada pomocniczości znajdują zastosowanie także w podejmowaniu decyzji w organizacjach ponadnarodowych.

Aby wykonywać swoje funkcje, niższe szczeble władzy gospodarczej muszą mieć do dyspozycji określone instrumenty fiskalne. Należą do nich przede wszystkim podatki, a także pożyczki i inne instrumenty dłużne. Jednakże w obrębie danego państwa, zarówno federalnego, jak i unitarnego, ma miejsce także redystrybucja

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

międzyregionalna, w postaci transferów jednostronnych, dokonywanych bezpośrednio z regionów bogatszych do biedniejszych albo za pośrednictwem budżetu centralnego (federalnego).

Określenie proporcji, w jakich rozdzielane są dochody podatkowe pomiędzy różne szczeble władzy publicznej, nosi w nauce o finansach nazwę problemu przyporządkowania podatkowego (*tax-assignment problem* – McLure 1998). Chodzi o znalezienie odpowiedzi na pytania o to, które sposród podatków najlepiej służą zaspokojaniu potrzeb lokalnych szczebli władzy publicznej, czy i jak różnicować podatki w zakresie rozmaitych jurysdykcji terytorialnych oraz jaki powinien być udział niższych szczebli władzy w ogólnych dochodach podatkowych. Ponieważ między regionami danego kraju występuje duża mobilność jednostek gospodarczych, mniejsza zaś między państwami, różnicowanie stawek podatków dochodowych w obrębie jednego kraju powodowałoby migrację podatników z regionów o wyższych podatkach do regionów o niższych podatkach. Podobne ograniczenie dotyczy podatków pośrednich, takich jak VAT i akcyza. W tym przypadku terytorialne zróżnicowanie stawek prowadziłoby do nieefektywnych podróży w poszukiwaniu tańszych dóbr i usług. Dlatego też zróżnicowanie podatków dochodowych i pośrednich w obrębie jednego kraju jest ekonomicznie nieuzasadnione, chyba że równolegle występują terytorialnie zróżnicowane korzyści, które odnoszą jednostki gospodarcze w poszczególnych regionach – wówczas różnicowanie może okazać się zasadne[9]. Z kolei typowym podatkiem lokalnym odpornym na mobilność podatników oraz ważnym źródłem dochodu władz lokalnych jest podatek od nieruchomości, w tym także podatek rolny i podatek leśny, potrącany zatem od aktywów nieprzenośnych.

Jednostronne transfery fiskalne dokonywane między szczeblami władzy publicznej stanowią ważny instrument federalizmu fiskalnego. U ich podstaw mogą znajdować się powody polityczne, ważne względy społeczne (takie jak wyrównywanie poziomu życia mieszkańców) lub ekonomiczne (związane z rozwojem potencjału gospodarczego poszczególnych regionów). W Polsce transfery takie noszą nazwę subwencji dla samorządów i mogą mieć charakter warunkowy lub bezwarunkowy. Subwencje warunkowe przeznacza się jedynie na ściśle określone cele (np. na oświatę), podczas gdy subwencje bezwarunkowe są instrumentem wyrównywania dochodów.

[9] Na przykład może to być uzasadnione potrzebą wsparcia regionów o dużym bezrobociu lub upośledzonych pod względem geograficznym (wyspy, obszary izolowane) lub klimatycznym.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.2.4. Unia fiskalna

Zjawisko federalizmu fiskalnego prowadzi do powstania unii fiskalnej, czyli systemu określonych, szczególnych powiązań fiskalno-budżetowych pomiędzy wybranymi regionami danego państwa. Uważa się, że unię fiskalną tworzą takie kluczowe składowe jak: co najmniej dwuszczeblowa struktura składająca się z centralnej władzy budżetowej i budżetów regionalnych, system podatków zapewniający dochody budżetowi centralnemu, system bezzwrotnych transferów pomiędzy poszczególnymi regionami oraz z budżetu centralnego do budżetów regionalnych, system reguł fiskalnych zapewniający właściwą dyscyplinę budżetową szczebli regionalnych oraz zdolność do zaciągania pożyczek przez budżet centralny na rachunek i w imieniu całej unii (IMF 2013). Tak skonstruowana unia fiskalna ma gwarantować niezbędną równowagę między odpowiedzialnością fiskalną z jednej strony a solidarnością między regionami z drugiej strony.

Unie fiskalne występują najczęściej w ramach poszczególnych państw, gdzie poczucie wzajemnej solidarności i współodpowiedzialności za losy oraz rozwój poszczególnych regionów jest odpowiednio silne. Lecz unia fiskalna (bądź jej elementy) może też funkcjonować w skali międzynarodowej, gdy państwa zrzeszone w ramach organizacji międzynarodowej decydują się przekazać część swoich kompetencji w zakresie prowadzenia polityki fiskalno-budżetowej szczeblowi ponadnarodowemu. Z elementami takiej unii fiskalnej mamy do czynienia w przypadku Unii Europejskiej, gdzie istnieje wspólny (choć bardzo niewielki) budżet europejski, finansowany częściowo podatkami krajowymi, a także system reguł fiskalnych, mający zapewnić dyscyplinę budżetową w poszczególnych państwach.

Międzynarodowe unie fiskalne nie są jednak częstym zjawiskiem. Wiąże się z nimi ryzyko jednostronnych, asymetrycznych transferów zasobów z krajów bardziej odpowiedzialnych finansowo do krajów mniej odpowiedzialnych. Dlatego też suwerenne państwa niechętnie zawierają unie fiskalne, głównie z uwagi na obawę przed zjawiskiem pokusy nadużycia i ze względu na polityczną niepopularność transferów finansowych kierowanych do innych krajów. Kryzys zadłużenia w strefie euro w latach 2010–2012 pokazał, że unia walutowa składająca się z suwerennych krajów nierespektujących reguł fiskalnych i niemająca wspólnego budżetu centralnego może być narażona na częste załamania. Więcej na ten temat piszemy w rozdziale 8.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.2.5. Finanse samorządów terytorialnych

Zarówno w państwach federalnych, jak i w państwach unitarnych o rozwiniętym systemie samorządowym lokalne oraz regionalne szczeble władzy publicznej realizują szereg zadań i funkcji, które wymagają finansowania. Jak wspomniano wyżej, środki na te zadania mogą pochodzić z podatków, długu lub z transferów płynących od wyższych szczebli władzy publicznej. W Polsce wśród dochodów jednostek samorządu terytorialnego wyróżnia się dochody własne, subwencję ogólną i dotacje celowe z budżetu państwa.

Dochody własne obejmują dochody uzyskiwane z działalności władz jednostek samorządu terytorialnego (JST), wpływy z podatków lokalnych i opłat, dochody finansowe (np. z odsetek) oraz udział we wpływach z podatku dochodowego od osób prawnych (CIT) i podatku dochodowego od osób fizycznych (PIT). Subwencja ogólna jest dochodem transferowym przekazywanym z budżetu państwa i przeznaczona zostaje na pokrycie wydatków związanych z wypełnianiem przez JST ich ustawowych zadań i funkcji. O szczegółowym podziale tych środków na wybrane cele decydują władze JST. Subwencja ogólna dla gmin i powiatów obejmuje część wyrównawczą, część równoważącą i część oświatową, podczas gdy subwencja ogólna dla województw składa się z części wyrównawczej, części regionalnej i części oświatowej. Z kolei dotacje celowe są dochodem transferowym przekazywanym z budżetu państwa z przeznaczeniem na realizację konkretnych zadań, zleconych władzom lokalnym przez władze centralne. Do zadań tych należą działania wynikające z ustaw, działania związane z usuwaniem bezpośrednich zagrożeń dla porządku i bezpieczeństwa publicznego oraz skutków klęsk żywiołowych czy też działania wynikające z umów międzynarodowych. Szczegółowy opis dochodów JST oraz ich przeznaczenia zawiera ramka 3.2.

Ramka 3.2. Dochody jednostek samorządu terytorialnego w Polsce – stan na 2015 rok (na podstawie ustawy o dochodach jednostek samorządu terytorialnego)

Dochodami jednostek samorządu terytorialnego są:

1) dochody własne,
2) subwencja ogólna,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3) dotacje celowe z budżetu państwa.

Ad.1) Dochody własne:

1. Źródłami dochodów własnych JST są:

a) dochody uzyskiwane przez samorządowe jednostki budżetowe oraz wpłaty od samorządowych zakładów budżetowych;
b) dochody z majątku JST;
c) spadki, zapisy i darowizny na rzecz JST;
d) dochody z kar pieniężnych i grzywien określonych w odrębnych przepisach;
e) 5,0% dochodów uzyskiwanych na rzecz budżetu państwa w związku z realizacją zadań z zakresu administracji rządowej oraz innych zadań wynikających z ustaw, o ile odrębne przepisy nie stanowią inaczej;
f) odsetki od pożyczek udzielanych przez JST, o ile odrębne przepisy nie stanowią inaczej;
g) odsetki od nieterminowo przekazywanych należności stanowiących dochody JST;
h) odsetki od środków finansowych gromadzonych na rachunkach bankowych JST, o ile odrębne przepisy nie stanowią inaczej;
i) dotacje z budżetów innych jednostek samorządu terytorialnego;
j) inne dochody należne JST na podstawie odrębnych przepisów.

2. Źródłami dochodów własnych gminy są ponadto:

a) wpływy z podatków: od nieruchomości, rolnego, leśnego, od środków transportowych, dochodowego od osób fizycznych, opłacanego w formie karty podatkowej, od spadków i darowizn, od czynności cywilnoprawnych;
b) wpływy z opłat: skarbowej, targowej, miejscowej, uzdrowiskowej i od posiadania psów, eksploatacyjnej (w części określonej w ustawie z dnia 4 lutego 1994 roku – Prawo geologiczne i górnicze, DzU 1994, nr 27, poz. 96, z późn. zm.), oraz innych opłat uiszczanych na podstawie odrębnych przepisów;
c) udział we wpływach z podatku dochodowego od osób fizycznych od płatników tego podatku zamieszkałych na obszarze gminy (do wysokości 39,34%,);
d) udział we wpływach z podatku dochodowego od osób prawnych od płatników tego podatku posiadających siedzibę na obszarze gminy (6,71%).

3. Źródłami dochodów własnych powiatu są ponadto:

a) wpływy z opłat stanowiących dochody powiatu, uiszczanych na podstawie odrębnych przepisów;
b) udział we wpływach z podatku dochodowego od osób fizycznych od podatników tego podatku zamieszkałych na obszarze powiatu (10,25%);

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

c) udział we wpływach z podatku dochodowego od osób prawnych od podatników tego podatku posiadających siedzibę na obszarze powiatu (1,40%).

4. Źródłem dochodów województwa jest natomiast:

a. udział we wpływach z podatku dochodowego od osób fizycznych od podatników tego podatku zamieszkałych na obszarze województwa (1,60%);

b. udział we wpływach z podatku dochodowego od osób prawnych od podatników tego podatku posiadających siedzibę na obszarze województwa (14,75%).

Ad. 2) Subwencja ogólna:

1. Subwencja ogólna składa się z części:

a) dla gmin i powiatów:
   - wyrównawczej,
   - równoważącej,
   - oświatowej,

b) dla województw:
   - wyrównawczej,
   - regionalnej,
   - oświatowej.

2. O przeznaczeniu środków otrzymanych z tytułu subwencji ogólnej decyduje organ stanowiący jednostki samorządu terytorialnego.

3. Struktura:

a) ogólna kwota i część oświatowa wynikają z ustawy budżetowej;

b) tworzy się rezerwę subwencji ogólnej na cele infrastruktury drogowej;

c) część wyrównawcza dla gmin składa się z kwoty podstawowej i kwoty uzupełniającej;
   - kwotę podstawową otrzymuje gmina, w której wskaźnik dochodów podatkowych (z podatków oraz opłaty skarbowej i eksploatacyjnej) na jednego mieszkańca w gminie („wskaźnik G") jest niższy niż 92% wskaźnika dochodów podatkowych w przypadku wszystkich gmin, zwanego „wskaźnikiem Gg";
   - kwota podstawowa jest zróżnicowana w przypadku gmin poniżej 40% Gg, 40–75% i 75–92%;
   - kwotę uzupełniającą otrzymuje gmina, w której gęstość zaludnienia jest niższa od średniej gęstości zaludnienia w kraju, chyba że jej wskaźnik G jest wyższy od 150% wskaźnika Gg;

d) część równoważąca subwencji ogólnej w przypadku gmin jest przeznaczona dla gmin o niższych dochodach podatkowych lub bardziej obciążonych wydatkami na dodatki mieszkaniowe;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

e) część wyrównawczą subwencji ogólnej w przypadku powiatów stanowi suma kwoty podstawowej i kwoty uzupełniającej;
   - kwotę podstawową otrzymuje powiat, w którym wskaźnik dochodów podatkowych (z podatków dochodowych) na jednego mieszkańca w powiecie („wskaźnik P") jest niższy od wskaźnika dochodów podatkowych w przypadku wszystkich powiatów, zwanego „wskaźnikiem Pp";
   - kwota podstawowa jest obliczana tak samo niezależnie od stopnia powyższej różnicy;
   - kwotę uzupełniającą otrzymuje powiat, w którym wskaźnik bezrobocia, liczony jako iloraz stopy bezrobocia w powiecie i stopy bezrobocia w kraju, zwany dalej „wskaźnikiem B", jest wyższy od 1,10%;
   - kwota ta jest zróżnicowana wedle progów 1,10–1,25%, 1,25–1,40%, 1,40%+;

f) część równoważąca subwencji ogólnej w przypadku powiatów jest przeznaczona dla jednostek o niższych dochodach oraz z większymi problemami związanymi z zadaniami na rynku pracy, rodzinami zastępczymi i większą siecią drogową;

g) w województwie kwota podstawowa części wyrównawczej jest obliczana tak samo jak w powiecie;

h) kwotę uzupełniającą otrzymuje województwo, w którym liczba mieszkańców nie przekracza 3 milionów;

i) kwota ta jest zróżnicowana wg progów: > 3%, 3–2,5%, 2,5–2, 2% >;

j) część regionalną otrzymują w częściach województwa o bezrobociu przekraczającym o 10% średnią krajową (1/5), o większej powierzchni dróg, PKB niższym niż 75% średniej krajowej, większych wydatkach na kolejowe przewozy regionalne oraz o niższych dochodach.

4. Ministerstwo Finansów przekazuje kwoty w miesięcznych ratach (rata subwencji oświatowej za marzec jest podwójna); za nieprzekazanie środków nalicza się odsetki podobnie jak w przypadku zaległości podatkowych (wzbliżony sposób rozlicza się także wpłaty wyrównawcze); spory rozpatruje sąd powszechny.

5. Wpłaty wyrównawcze – tworzą fundusz na część równoważącą/regionalną:

a) gminy z G większym niż 150% Gg dokonują wpłat do budżetu państwa, z przeznaczeniem na część równoważącą subwencji ogólnej dla gmin. Kwota ta wynosi:
   - 20% nadwyżki – jeżeli G jest nie większe niż 200% wskaźnika Gg;
   - dodatkowo 25% nadwyżki przy ponad 200% wskaźnika Gg – gdy G jest większy niż 200% i nie większy niż 300% Gg;
   - dodatkowo 30% nadwyżki wskaźnika G przy ponad 300% wskaźnika Gg;

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

b) powiaty z P większym niż 110% Pp dokonują wpłat do budżetu państwa, z przeznaczeniem na część równoważącą subwencji ogólnej dla powiatów. Kwota ta wynosi:
   – 80% nadwyżki wskaźnika P przy ponad 110% wskaźnika Pp – w przypadku powiatów, w których wskaźnik P jest nie większy niż 120% wskaźnika Pp;
   – dodatkowo 95% nadwyżki wskaźnika P przy ponad 120% wskaźnika Pp – w przypadku powiatów, w których wskaźnik P jest większy niż 120% i nie większy niż 125% wskaźnika Pp;
   – dodatkowo 98% nadwyżki wskaźnika P przy ponad 125% wskaźnika Pp;

c) województwa ze wskaźnikiem W (wskaźnik dochodów podatkowych na jednego mieszkańca w województwie) większym niż 110% wskaźnika Ww (wskaźnik dochodów podatkowych odnośnie wszystkich województw) dokonują wpłat do budżetu państwa, z przeznaczeniem na część regionalną subwencji ogólnej dla województw. Kwota ta wynosi:
   – 80% nadwyżki – w przypadku województw, w których wskaźnik W jest nie większy niż 170% wskaźnika Ww;
   – dodatkowo 95% nadwyżki przy ponad 170% Ww – w przypadku województw, w których wskaźnik W jest większy niż 170% wskaźnika Ww.

Ad. 3) Dotacje celowe:

1. Dochodami jednostek samorządu terytorialnego mogą być dotacje celowe z budżetu państwa (do 80% wydatków) na:

a) realizację zadań z zakresu administracji rządowej oraz na realizację innych zadań wynikających z ustaw;
b) realizację zadań realizowanych przez jednostki samorządu terytorialnego na mocy porozumień zawartych z organami administracji rządowej;
c) usuwanie bezpośrednich zagrożeń w przypadku bezpieczeństwa i porządku publicznego, skutków powodzi i osuwisk ziemnych oraz skutków innych klęsk żywiołowych;
d) finansowanie lub dofinansowanie realizacji określonych w ustawach zadań własnych;
e) realizację zadań wynikających z umów międzynarodowych (UE).

2. Dochodami powiatu są dotacje celowe z budżetu państwa przeznaczane na realizacje zadań straży oraz inspekcji.
3. Dochodami jednostek samorządu terytorialnego mogą być środki z funduszy celowych oraz od innych JST.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W 2013 roku łączne dochody JST wyniosły 183 458,3 mln PLN, co stanowiło równowartość około 29% wielkości dochodów całego sektora finansów publicznych w Polsce i około 61% wielkości dochodów budżetu państwa (budżetu centralnego). Udział dochodów JST w PKB wyniósł w 2013 roku 11,2%, i w porównaniu z 2009 rokiem uległ niewielkiemu obniżeniu (z 11,5%). Najzasobniejszą kategorię dochodów stanowiły dochody własne JST, które wynosiły 91 800,3 mln PLN. Subwencja ogólna osiągnęła poziom 51 257,1 mln PLN, a dotacje celowe 40 400,9 mln PLN. Zgromadzone dane dowodzą, że skala redystrybucji środków między szczeblami władzy publicznej jest znacząca. Transfery z budżetu państwa stanowiły niemal 50% wszystkich dochodów otrzymywanych przez JST. Jednocześnie należy zauważyć, że system redystrybucji w Polsce jest bardzo skomplikowany. Jak pokazują informacje zgromadzone w ramce 3.2, subwencja ogólna obejmuje kilka rodzajów transferów, których zasadniczym celem jest zmniejszanie różnic w dochodach na mieszkańca między poszczególnymi gminami, powiatami i województwami. Oprócz stopnia zamożności społeczności regionalnych i lokalnych (mierzonego wielkością dochodów podatkowych na mieszkańca), wielkość subwencji zależy także od skali obciążenia JST wydatkami związanymi z przeciwdziałaniem bezrobociu, utrzymaniem dróg i wsparciem dla rodzin.

Szczególny charakter ma część regionalna subwencji ogólnej, obejmująca transfery przekazywane od województw i powiatów bogatszych do biedniejszych. Jest to klasyczny mechanizm międzyregionalnej redystrybucji dochodów, charakterystyczny w przypadku państw unitarnych. W Polsce budzi on jednak spore kontrowersje ze względu na skalę redystrybucji[10].

Z punktu widzenia struktury podmiotowej największymi dochodami dysponowały powiaty i miasta na prawach powiatu – łącznie 87 294,2 mln PLN, następnie gminy – 80 043,4 mln PLN, oraz województwa – 16 120,7 mln PLN. Struktura rodzajowa dochodów własnych JST została przedstawiona w tabeli 3.3.

Jak widać, największą pozycję w dochodach własnych stanowią dochody z podatku PIT (około 35%). Duży udział mają też pozostałe dochody (takie jak dochody finansowe, dochody ze sprzedaży majątku – około 23%) oraz dochody

[10] Najbogatsze województwa wpłacają do budżetu państwa część swoich dochodów, które następnie przekazywane są województwom uboższym. Skala redystrybucji jest na tyle duża, że województwa zajmujące czołowe miejsca pod względem poziomu dochodów podatkowych na mieszkańca (Mazowsze, Dolny Śląsk) po dokonaniu transferów spadają w tej klasyfikacji do dolnej części tabeli. Dlatego też system międzyregionalnych transferów w Polsce nazywany bywa często „janosikowym".

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

z podatku od nieruchomości (około 20% całości). Z kolei tradycyjne podatki, takie jak rolny i leśny czy opłata targowa, odgrywają znacznie mniejszą rolę.

Tabela 3.3. Wielkość i struktura dochodów własnych JST w Polsce w 2013 roku (dane w mln PLN)

Kategoria dochodu	Mln PLN
Udział w dochodach z podatku CIT	6 300,1
Udział w dochodach z podatku PIT	32 460,8
Podatek rolny	1 665,2
Podatek leśny	227,8
Podatek od nieruchomości	18 729,4
Podatek od środków transportu	940,9
Podatek opłacany w formie karty podatkowej przez osoby fizyczne prowadzące działalność gospodarczą	82,6
Podatek od spadków i darowizn	255,5
Podatek od czynności cywilnoprawnych	1 628,7
Opłaty skarbowe	488,8
Opłata eksploatacyjna	280,7
Opłata targowa	187,3
Dochody z majątku	7 563,3
Pozostałe dochody własne	20 898,1
Razem dochody własne JST	91 800,3

Źródło: *Finanse jednostek samorządu terytorialnego w 2013 roku*, GUS, Warszawa 2014, s. 34.

3.3. Skłonność do deficytu (*deficit bias*)

Deficyty budżetowe są zjawiskiem powszechnym we współczesnych gospodarkach. Pojawiają się i pogłębiają nie tylko w sytuacjach szczególnych, takich jak wojny czy kryzysy, ale także w normalnych czasach, najczęściej na skutek cyklicznych wahań wzrostu gospodarczego oraz jako konsekwencja interwencjonizmu państwowego i realizacji funkcji państwa w gospodarce oraz w pozostałych obszarach życia społecznego. Oprócz ekonomicznych (np. spadek dochodów podatkowych w okresach recesji) czy społecznych (wzrost wydatków na opiekę społeczną) przesłanek występowania deficytu, we współczesnych państwach demokratycznych funkcjonują także pewne mechanizmy polityczne wywołujące i potęgujące skłonność do deficytu budżetowego (*deficit bias*).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W systemie demokracji przedstawicielskiej funkcjonują dwa takie mechanizmy powodujące silną tendencję do występowania deficytu w budżetach narodowych – jednym z nich jest polityczny cykl wyborczy, a drugim tzw. problem wspólnego zasobu. W ramach cyklu wyborczego (*political-electoral cycle*) politycy okresowo zwiększają wydatki, dążąc do uzyskania poparcia wyborców, co powoduje nałożenie się na regularny gospodarczy cykl koniunkturalny „politycznego cyklu koniunkturalnego" (*political business cycle* – PBC – zob. Nordhaus 1975, Rogoff i Siebert 1988, Schultz 1995). Urzędujący politycy, którzy ubiegają się o reelekcję, mają skłonność do zwiększania wydatków w okresie bezpośrednio przed i w czasie kampanii wyborczej. Z kolei politycy, którzy dopiero chcą być wybrani, przejawiają tendencję do formułowania obietnic zgodnych z oczekiwaniami wyborców. Ponieważ wyborcy z reguły oczekują zwiększonych wydatków publicznych i mniejszych obciążeń podatkowych, politycy są skłonni składać takie obietnice, aby uzyskać ich poparcie[11]. Część z nich jest następnie realizowana, co przekłada się z reguły na pogorszenie stanu finansów publicznych (Maskin i Tirole 2014). Zachowania takie są oczywiście dyktowane chęcią utrzymania się przy władzy.

Badania empiryczne prowadzone przez wielu ekonomistów potwierdzają obecność „politycznego cyklu koniunkturalnego". Na przykład Alesina i in. (1997), analizując wpływ wyborów na sytuację budżetową w 13 krajach OECD w latach 1961–1993, oszacowali, że w latach wyborczych deficyt budżetowy powiększa się średnio o 0,6% PKB. Podobny wynik uzyskali Shi i Svensson (2002, 2003), którzy badali zmiany sytuacji budżetowej na podstawie danych panelowych obejmujących kraje rozwinięte i kraje rozwijające się za lata 1975–1995. Uzyskane przez nich wyniki wskazują, że saldo budżetu w latach wyborczych ulega pogorszeniu średnio o 0,6% PKB w krajach rozwiniętych i o ponad 1,0% PKB w krajach rozwijających się. Z kolei Schuknecht (1996) oszacował cykliczne pogorszenie salda budżetu w 35 krajach rozwijających się średnio na ponad 0,6% PKB. Ogólnie rzecz biorąc, ta silna tendencja do zwiększania wydatków w okresach przed- i powyborczych (oraz ewentualnie do obniżania podatków) nie jest kompensowana odpowiednimi oszczędnościami w czasie sprawowania władzy między kolejnymi okresami wyborczymi. Rodzi to systematyczną skłonność do utrzymywania się deficytu w zakresie finansów publicznych[12].

[11] Obietnice składane przez polityków w kampanii wyborczej nazywane są często „kiełbasą wyborczą" (*pork barrel*).

[12] Należy przy tym zauważyć, że zwiększanie wydatków budżetowych w latach wyborczych nie powoduje bynajmniej zauważalnego przyspieszenia wzrostu gospodarczego (zob. Canes-Wrone i Ponce de Leon

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Z kolei problem wspólnego zasobu (*common pool problem*) wynika z redystrybucyjnego charakteru polityki fiskalnej (Buchanan i Tullock 1962). Wydatki rządowe, finansowane z podatków pobieranych od ogółu społeczeństwa, przeznaczane są z reguły na pokrycie potrzeb wybranych grup społecznych (zawodowych, etnicznych, lokalnych, regionalnych). Oznacza to, że ci, którzy korzystają z określonych rządowych programów wydatkowych, nigdy nie ponoszą pełnego kosztu realizacji tych strategii, ponieważ ich wkład we wspólną pulę zasobów podatkowych jest mniejszy niż suma świadczeń otrzymanych z budżetu. W takiej sytuacji polityczni przedstawiciele grup korzystających z wydatków publicznych w sposób uprzywilejowany będą zawsze zainteresowani pozyskaniem dodatkowych środków na rzecz tych programów. Powstaje zatem dodatkowy koszt zewnętrzny (*externality*) tego rodzaju strategii, ponieważ ich całkowity koszt społeczny przewyższa indywidualny koszt ich finansowania przez daną grupę społeczną. Gdy w ten sposób zachowują się wszyscy wybieralni politycy (lub ich większość), powstaje silna, trwała tendencja do nadmiernych wydatków i do utrzymywania się deficytu w zakresie finansów publicznych (Von Hagen i Harden 1996, Hallerberg i von Hagen 1999, Eichengreen i in. 2011).

W ramach wzmocnionej współzależności poszczególnych krajów w wyniku postępu globalizacji i integracji rynków powstają dodatkowe bodźce do powiększania deficytu. Jest to zjawisko szczególnie charakterystyczne dla unii walutowej i wyraźnie widoczne w przypadku europejskiej unii walutowej. W ugrupowaniu integracyjnym, w ramach którego państwa posługują się wspólną walutą, skłonność do deficytu budżetowego nasila się na skutek występowania dodatkowych dwóch zjawisk: zjawiska „pasażera na gapę" (*free riding*) i zjawiska pokusy nadużycia (*moral hazard*). Problem „pasażera na gapę" pojawia się w sytuacji, gdy państwo członkowskie unii walutowej narusza wspólnie ustalony, dopuszczalny poziom deficytu i zwiększa wydatki, mając jednocześnie świadomość, że dodatkowy koszt z tym związany (wzrost stóp procentowych wynikający ze wzrostu popytu przy stałej podaży pieniądza) nie będzie obciążał wyłącznie tego jednego państwa, lecz ze względu na funkcjonowanie wspólnej waluty zostanie rozłożony na wszystkie państwa członkowskie unii. W takiej sytuacji państwo naruszające dopuszczalny poziom deficytu pokrywa tylko niewielką część ogólnego kosztu – znacznie mniejszą, niż gdyby pozostawało poza unia walutową.

2015). Jedną z przyczyn takiego stanu rzeczy może być zwiększona niepewność co do kierunków polityki gospodarczej po wyborach.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 3.4. Saldo sektora finansów publicznych w państwach członkowskich UE27w latach 1998–2014

Kraj	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Liczba deficytów	Liczba nadwyżek
BE	–0,9	–0,6	0,0	0,4	–0,1	–0,1	–0,3	–2,8	0,2	–0,3	–1,3	–5,9	–4,1	–3,9	–4,0	–2,9	–3,1	14	1
BG	1,2	0,1	–0,5	1,1	–1,2	–0,4	1,8	1,0	1,9	1,1	1,7	–4,7	–3,2	–2,1	–0,8	–0,8	–5,8	9	8
CZ	–5,0	–3,7	–3,7	–5,6	–6,8	–6,6	–3,0	–3,6	–2,6	–0,7	–2,7	–5,9	–4,7	–3,1	–4,4	–1,3	–1,9	17	0
DK	0,0	1,3	2,3	1,5	0,4	0,1	2,1	5,2	5,2	4,8	3,2	–2,7	–2,7	–1,9	–4,1	–1,3	1,5	5	11
DE	–2,2	–1,5	1,3	–2,8	–3,7	–4,0	–3,8	–3,3	–1,6	0,3	0,1	–3,0	–3,3	–1,0	0,1	–0,1	0,3	12	5
EE	–0,7	–3,5	–0,2	–0,1	0,3	1,7	1,6	1,6	2,4	2,5	–2,8	–1,7	0,1	1,0	–0,2	–0,1	0,7	8	9
IE	2,4	2,7	4,8	0,9	–0,3	0,4	1,4	1,6	2,9	0,1	–7,3	–14,3	–32,4	–13,0	–8,2	–5,7	–3,9	8	9
GR	–	–	–3,7	–4,5	–4,8	–5,6	–7,5	–5,2	–5,7	–6,4	–9,8	–15,4	–10,5	–9,2	–9,0	–12,4	–3,6	15	0
ES	–3,2	–1,4	–1,0	–0,6	–0,5	–0,2	–0,3	1,0	2,0	1,9	–4,2	–11,1	–9,2	–8,5	–10,6	–6,9	–5,9	14	3
FR	–2,6	–1,8	–1,5	–1,5	–3,1	–4,1	–3,6	–2,9	–2,3	–2,7	–3,3	–7,5	–7,0	–5,2	–4,8	–4,1	–3,9	17	0
IT	–2,8	–1,7	–0,8	–3,1	–2,9	–3,5	–3,5	–4,3	–3,4	–1,5	–2,7	–5,4	–4,6	–3,8	–5,0	–2,9	–3,0	17	0
CY	–4,1	–4,3	–2,3	–2,2	–4,4	–6,5	–4,1	–2,4	–1,2	3,4	0,9	–6,0	–5,3	–6,3	–3,0	–4,9	–8,9	15	2
LV	0,0	–3,9	–2,8	–1,9	–2,3	–1,6	–1,0	–0,4	–0,5	–0,3	–4,2	–9,7	–7,7	–3,5	–6,4	–0,9	–1,5	16	0
LT	–3,0	–2,8	–3,2	–3,6	–1,9	–1,3	–1,5	–0,5	–0,4	–1,0	–3,3	–9,5	–7,1	–5,5	–1,3	–2,6	–0,7	17	0
LU	3,4	3,4	6,0	6,1	2,1	0,5	–1,1	0,0	1,4	3,7	3,0	–0,9	–1,7	–0,6	–3,2	0,7	1,4	5	11
HU	–7,8	–5,4	–3,0	–4,0	–8,9	–7,2	–6,4	–7,9	–9,3	–5,0	–3,7	–4,5	–4,2	4,2	–0,6	–2,5	–2,5	16	1
MT	–9,9	–7,7	–6,2	–6,4	–5,5	–9,9	–4,7	–3,0	–2,7	–2,4	–4,5	–3,7	–3,6	–2,7	–2,0	–2,6	–2,1	17	0
NE	–0,9	0,4	2,0	–0,2	–2,1	–3,1	–1,7	–0,3	0,5	0,2	0,6	–5,5	–5,4	–4,6	–3,3	–2,4	–2,4	12	5
AT	–2,4	–2,3	–1,7	0,0	–0,7	–1,4	–4,4	–1,7	–1,5	–0,9	–0,9	–4,1	–4,6	–2,6	–4,1	–1,3	–2,7	17	0
PL	–4,3	–2,3	–3,0	–5,3	–5,0	–6,2	–5,4	–4,1	–3,6	–1,9	–3,7	–7,3	–7,9	–5,1	–2,5	–4,0	–3,3	17	0
PT	–3,5	–2,7	–2,9	–4,3	–2,9	–3,0	–3,4	–5,9	–4,1	–3,1	–3,5	–10,1	–9,1	–4,2	–3,9	–4,8	–7,2	17	0
RO	–3,2	–4,4	–4,7	–3,5	–2,0	–1,5	–1,2	–1,2	–2,2	–2,6	–5,7	–8,5	–6,4	–5,2	–6,4	–2,2	–1,4	17	0
SI	–2,4	–3,0	–3,7	–4,0	–2,5	–2,7	–2,2	–1,4	–1,3	–0,1	–1,8	–6,0	–5,6	–6,4	–3,0	–15,0	–5,0	17	0
SR	–5,3	–7,4	–12,3	–6,5	–8,2	–2,8	–2,4	–2,8	–3,2	–1,8	–2,1	–8,0	–7,9	–4,8	–3,8	–2,6	–2,8	17	0
FI	1,5	1,6	6,8	5,0	4,0	2,4	2,3	2,7	4,0	5,2	4,2	–2,6	–2,5	–0,9	–4,5	–2,5	–3,3	6	11
SE	0,7	0,9	3,6	1,5	–1,3	–1,0	0,6	2,2	2,3	3,6	2,2	–0,7	0,0	0,1	–1,8	–1,4	–1,7	6	10
UK	–0,1	0,9	3,6	0,5	–2,1	–3,4	–3,4	–3,4	–2,7	–2,7	–5,0	–11,4	–10,4	–8,2	–0,2	–5,7	–5,7	14	3
UE27																		362	89

Źródło: opracowanie własne na podstawie danych Eurostatu.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 80981 23A35373237

Z kolei pokusa nadużycia występuje wówczas, gdy państwo członkowskie zadłuża się ponad ustalony limit, wiedząc, że w sytuacji zagrożenia niewypłacalnością pozostałe państwa członkowskie będą zmuszone przyjść mu z pomocą, ponieważ straty spowodowane potencjalnym bankructwem (w tym między innymi utratą wartości aktywów w sektorze bankowym) przewyższałyby koszt ewentualnej pomocy. Pokusa nadużycia może przejawiać się też w działaniach danego państwa w przypadku dopuszczenia do inflacji wyższej niż poziom inflacji tolerowany przez pozostałe kraje unii walutowej, ponieważ wyższa inflacja zmniejsza realną wartość długu (Buti i in. 2003, Pisani-Ferry 2003). Na przykład, gdy jedno z państw członkowskich strefy euro jest bardzo zadłużone, jego przedstawiciel w Europejskim Banku Centralnym prawdopodbnie będzie skłonny tolerować wyższą inflację w unii walutowej, w związku z czym może też głosować za utrzymywaniem niższych stóp procentowych.

W tabeli 3.4 przedstawiono roczne salda sektora finansów publicznych w 27 państwach członkowskich UE w latach 1998–2014 (deficyty przekraczające limit 3% PKB zostały zaznaczone na szaro)[13]. Dane potwierdzają istnienie generalnej skłonności do deficytu w państwach członkowskich UE. Na łączną liczbę 459 rocznych obserwacji w przypadku 27 państw członkowskich w okresie obejmującym 17 lat (1998–2014) deficyty budżetowe wystąpiły 362 razy, czyli w 77,2% przypadków, nadwyżki pojawiły się 89 razy, czyli w 19,4% przypadków, a w 8 przypadkach budżety były (względnie) zrównoważone. Znaczna przewaga deficytów utrzymywała się w całym analizowanym okresie mimo obowiązywania w UE dyscyplinujących reguł fiskalnych, wśród których wspomniana wyżej reguła celu średniookresowego nakładała wymóg utrzymywania równowagi budżetowej (lub utrzymywania nadwyżki) w średnim okresie.

Skłonność do deficytu, a w konsekwencji i tendencja do zadłużania się, jest więc powodowana zarówno przez przyczyny związane z funkcjonowaniem systemu demokracji przedstawicielskiej, jak i czynniki charakterystyczne dla pogłębionej integracji, w tym występujące zwłaszcza w ramach unii walutowej. Zadłużenie nie może jednak rosnąć w nieskończoność – po przekroczeniu pewnego poziomu wzrasta bowiem ryzyko niewypłacalności i kryzysu zadłużenia. Potrzeba uniknięcia nadmiernego przyrostu deficytu i długu w tych warunkach zmusza do sięgania po prawne oraz instytucjonalne narzędzia ograniczania wydatków publicznych

[13] W zestawieniu nie uwzględniono Chorwacji, która stała się 28 państwem członkowskim UE 1 lipca 2013 roku.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i/lub deficytu. Instrumenty te przyjmują najczęściej postać reguł fiskalnych (często w postaci numerycznej) określających kształtowanie się ważnych zmiennych fiskalnych, takich jak różne kategorie wydatków i dochodów budżetowych, poziom deficytu lub poziom długu (Kopits i Symansky 1998)[14]. Niekiedy powołuje się do życia specjalne niezależne instytucje stojące na straży dyscypliny fiskalnej – są to tzw. rady polityki fiskalnej (zob. Eichengreen i in. 1999; Wyplosz 2002a, 2002b, 2005). Wobec braku stałych reguł lub instytucji postulat zachowania dyscypliny fiskalnej realizowany jest – z lepszym lub gorszym efektem – poprzez centralizację decyzji budżetowych w rękach ministra finansów lub każdorazowe negocjacje między przedstawicielami różnych grup politycznych czy sektorów instytucjonalnych gospodarki w ramach parlamentu (lub innych ciał), co jednak skutkuje najczęściej narastaniem długu (von Hagen 2002).

3.4. Finansowanie deficytu

Finansowanie deficytu to pozyskiwanie środków pieniężnych umożliwiających zrealizowanie wydatków nieznajdujących pokrycia w dochodach budżetowych. Aby sfinansować deficyt powstały w wyniku nadmiernych wydatków (lub niedoboru dochodów), rząd może sięgnąć po następujące rozwiązania:

a) zaciąganie pożyczek przez rząd w banku centralnym, co oznacza dodatkową kreację pieniądza, czyli zwiększanie bazy monetarnej przez bank centralny – dlatego też ten sposób nazywany jest monetyzacją deficytu;
b) zaciąganie pożyczek i kredytów w bankach komercyjnych;
c) zaciąganie pożyczek na krajowym rynku kapitałowym poprzez sprzedaż papierów wartościowych (obligacje skarbowe, bony skarbowe) uczestnikom rynku, głównie bankom, innym instytucjom finansowym, a także przedsiębiorstwom i osobom prywatnym;
d) zaciąganie pożyczek i kredytów za granicą (w formie pożyczek bankowych lub poprzez emisję i sprzedaż papierów wartościowych);
e) przychody ze sprzedaży majątku (np. prywatyzacja), transfery zagraniczne i inne dochody jednorazowe (zwrot udzielonych pożyczek, odszkodowania, sprzedaż koncesji).

[14] Reguły fiskalne są omówione w punkcie 3.10.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.4.1. Finansowanie deficytu przy pomocy zwiększania bazy monetarnej (monetyzacja deficytu)

Zaciąganie pożyczek przez rząd w banku centralnym powoduje zwiększenie bazy monetarnej, czyli ilości pieniądza rezerwowego banku centralnego (*monetary base, reserve money* – zob. rozdział 5). W praktyce przekłada się to na wzrost podaży pieniądza i w konsekwencji jego drukowanie, ponieważ udzielenie pożyczki rządowi przez bank centralny oznacza uznanie rachunku rządowego w banku centralnym, co umożliwia rządowi dokonywanie wypłat z tego rachunku i finansowanie dodatkowych wydatków. Bank centralny wykorzystuje tu swój sankcjonowany prawem przywilej kreacji pieniądza gotówkowego. Gdy deficyt finansowany jest bezpośrednio dodatkową emisją pieniądza (który dostarcza bank centralny), koszt finansowania w krótkim okresie albo wynosi zero (wyłącznie koszt druku banknotów i bicia monet), albo pozostaje niski, gdy deficyt finansowany jest pożyczką z banku centralnego po stopie procentowej niższej od stóp rynkowych. Różnica między (zerowym lub obniżonym) kosztem emisji pieniądza a kosztem zaciągnięcia długu na rynku stanowi zysk z emisji – tzw. seniorat (*seigniorage*).

Gdy skala deficytu jest duża i zwiększanie bazy monetarnej przekracza ilość pieniądza, zapewniającą utrzymanie stabilności cen, pojawia się inflacja, która obniża wartość realną pieniądza oraz realną wartość długu. Tę utratę wartości realnych zasobów pieniężnych znajdujących się w obiegu nazywa się często podatkiem inflacyjnym. Monetyzacja deficytu – jeśli wiąże się z utratą wartości pieniądza – przerzuca zatem ciężar obsługi długu z dłużników, w tym rządu, na posiadaczy aktywów pieniężnych (nominalnych) i konsumentów, którzy ponoszą koszty rosnącej inflacji. Pożyczanie pieniędzy w banku centralnym i finansowanie deficytu dodatkową emisją pieniądza jest więc bardzo atrakcyjnym sposobem zmniejszania realnej wartości długu, w związku z czym rządy poszczególnych państw chętnie korzystały z tej możliwości w przeszłości. Wielkie inflacje występującce wśród państwach europejskich w latach 20. oraz inflacyjne epizody w państwach Ameryki Łacińskiej w latach 80. i 90. XX wieku były spowodowane właśnie dużymi deficytami budżetowymi, finansowanymi emisją krajowego pieniądza.

Finansowanie przyrostu deficytu budżetowego wzrostem bazy monetarnej wywołuje przejściowy wzrost produkcji i dochodu do momentu, aż uczestnicy rynku nie dostosują swoich oczekiwań do rosnącej inflacji oraz dopóty, dopóki nie wprowadzi się pełnej indeksacji dochodów. Wpływ monetyzacji deficytu na stopę

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

procentową jest jednak *ex ante* nieokreślony – w modelu IS-LM obie krzywe przesuwają się w prawo, a nowy punkt ich przecięcia (punkt równowagi) może znaleźć się zarówno powyżej, jak i poniżej dotychczasowego poziomu stopy procentowej. Ogólnie rzecz biorąc, gdy bank centralny dostosowuje podaż pieniądza do rosnących potrzeb wydatkowych rządu, stopy procentowe nie wzrastają. Taka polityka pieniężna nazywana jest polityką akomodacyjną lub akomodacją pieniężną (*accomodative policy*, *monetary accommodation*).

W wyniku takiej polityki na dłuższą metę musi jednak pojawić się inflacja, która powoduje wzrost stóp procentowych i ogranicza efekt zwiększania wydatków rządowych. Na skutek utrwalenia się oczekiwań inflacyjnych wśród uczestników rynku, krzywa Philipsa staje się pionowa i wzrost wydatków rządowych nie wywołuje już wzrostu produkcji i zatrudnienia, a jedynie wzrost inflacji. Inflacyjne skutki finansowania deficytu przez zwiększanie masy pieniądza rezerwowego zostały opisane w 1981 roku przez amerykańskich ekonomistów Thomasa Sargenta i Neila Wallace'a w ich słynnej pracy o „nieprzyjemnej arytmetyce monetarnej" (zob. rozdział 4.10.1).

Inflacyjne skutki finansowania deficytu poprzez monetyzację skłoniły wiele państw do wprowadzenia prawnego zakazu stosowania tej metody. Stosowany jest on m.in. w Polsce oraz we wszystkich krajach Unii Europejskiej. Art. 220, ustęp 2 Konstytucji RP stanowi, że „ustawa budżetowa nie może przewidywać pokrywania deficytu budżetowego przez zaciąganie zobowiązania w centralnym banku państwa". Z kolei art. 123 Traktatu o funkcjonowaniu Unii Europejskiej stanowi w ustępie 1, co nastepuje:

Zakazane jest udzielanie przez Europejski Bank Centralny lub banki centralne Państw Członkowskich, zwane dalej „krajowymi bankami centralnymi", pożyczek na pokrycie deficytu lub jakichkolwiek innych kredytów instytucjom, organom lub jednostkom organizacyjnym Unii, rządom centralnym, władzom regionalnym, lokalnym lub innym władzom publicznym, innym instytucjom lub przedsiębiorstwom publicznym Państw Członkowskich, jak również nabywanie bezpośrednio od nich przez Europejski Bank Centralny lub krajowe banki centralne ich papierów dłużnych"[15].

[15] Oznacza to także, że bieżący rachunek rządu w banku centralnym musi wykazywać saldo dodatnie na koniec każdego dnia obrachunkowego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.4.2. Finansowanie deficytu z długu

Finansowanie z długu może polegać albo na zaciąganiu pożyczek przez rząd bezpośrednio w bankach komercyjnych, albo na sprzedaży przez rząd dłużnych skarbowych papierów wartościowych (obligacje skarbowe, bony skarbowe) inwestorom na rynku kapitałowym w zamian za środki pieniężne. W pierwszym przypadku zaciągnięty przez rząd dług pozostaje w rękach jednej instytucji finansowej aż do momentu jego całkowitej spłaty. W szczególności zaciągnięty dług nie może być (w normalnych warunkach) przedmiotem obrotu wtórnego na rynkach finansowych. Pożyczki takie, ze względu na duże kwoty wchodzące w grę i związane z tym ryzyko, były najczęściej udzielane rządom przez konsorcja bankowe, składające się z kilku lub kilkunastu instytucji. Udzielone w ten sposób pożyczki stanowią aktywa niepłynne, co ogranicza swobodę kształtowania przez banki struktury portfela aktywów stosownie do zmieniających się warunków rynkowych. Również krąg potencjalnych pożyczkodawców jest zasadniczo ograniczony do banków i nie obejmuje innych inwestorów instytucjonalnych oraz indywidualnych, którzy mogliby być zainteresowani lokowaniem środków finansowych na dłuższe okresy (np. fundusze emerytalne i ubezpieczeniowe). Z powodu tych ograniczeń koszt takich pożyczek był z reguły dość wysoki, a ich dostępność zależała od sytuacji płynnościowej banków. Dlatego też ta metoda finansowania deficytu była stosowana w czasach, gdy nie istniały rozwinięte rynki kapitałowe, a obecnie sięga się po nią tylko w szczególnych sytuacjach i w krajach najmniej rozwiniętych.

Najpowszechniej stosowaną metodą finansowania deficytu jest pozyskiwanie środków pieniężnych poprzez sprzedaż przez rząd skarbowych papierów wartościowych (SPW) na rynku kapitałowym. Tego typu dokument daje nabywcy prawo do otrzymywania określonego dochodu z tytułu udzielonej rządowi pożyczki w określonym czasie. Ten dochód, który dla rządu stanowi koszt obsługi długu, może przyjmować postać wypłat stałych kwot w danym okresie. W przypadku obligacji o stałym oprocentowaniu i o kilku- lub kilkunastoletnim okresie do wykupu, tytułem do uzyskania dochodu w danym okresie jest tzw. kupon (*coupon*), uprawniający do otrzymywania kwot odsetek raz na rok lub raz na pół roku[16]. W przypadku

[16] Historycznie rzecz biorąc, kupon przyjmował postać papierowych odcinków, stanowiących część papierowej obligacji. Posiadacz obligacji odrywał kupon od obligacji i przedkładał go instytucji emitującej środki w celu otrzymania odsetek. Obecnie, w związku z upowszechnieniem się niematerialnej formy obligacji, kupony coraz częściej mają postać zapisów elektronicznych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

obligacji o zmiennym oprocentowaniu (np. obligacje indeksowane według stopy inflacji) kwoty płatności odsetkowych są zmienne w czasie, określone przez warunki emisji danej obligacji. Niekiedy, zwłaszcza w odniesieniu do krótkoterminowych papierów skarbowych (z reguły z okresem wykupu poniżej jednego roku), płatności należnych odsetek mają charakter wypłat jednorazowych kwot w formie odsetek zaktualizowanych, płatnych na początku (dyskonto), lub odsetek skapitalizowanych płatnych wraz z kapitałem na końcu okresu pożyczki.

Koszt finansowania jest równy stopie oprocentowania SPW sprzedawanych przez rząd. Kształtuje się on najczęściej poniżej rynkowej stopy oprocentowania kredytów komercyjnych, ponieważ rząd jest z reguły bardziej wiarygodnym dłużnikiem niż pozostali krajowi kredytobiorcy. Dzieje się tak dlatego, że rząd ma wyłączne prawo poboru podatków w danym kraju, stanowiących stałe źródło dochodów budżetu państwa i zapewniających środki na spłatę zobowiązań. Dlatego też w normalnych warunkach rynkowych (tzn. kiedy rządy nie przeżywają trudności finansowych) żadna instytucja finansowa mająca siedzibę w danym kraju nie ma wyższego ratingu niż rząd. Reguła ta oczywiście przestaje obowiązywać w sytuacji zagrożenia kryzysem płynności lub kryzysem niewypłacalności.

W przeciwieństwie do zaciągania pożyczek bankowych, emisja papierów skarbowych zapewnia większą elastyczność pozyskiwania środków finansowych, poszerza krąg potencjalnych inwestorów i obniża koszty finansowania. SPW stanowią aktywa płynne, które są przedmiotem obrotu na rynku wtórnym. Daje to inwestorom możliwość dostosowania poziomu i struktury czasowej posiadanych obligacji do swoich aktualnych preferencji. Rządowi umożliwia zaś dostosowanie oferty obligacji i bonów do struktury popytu, z uwzględnieniem zróżnicowanych poziomów awersji do ryzyka i zróżnicowanych horyzontów inwestowania wśród poszczególnych przedsiębiorców. Papiery skarbowe o okresach zapadalności od 2 do 5 lat są z reguły nabywane przez fundusze inwestycyjne, prywatne i publiczne (*sovereign wealth funds*), przy czym papiery o dużej dochodowości, lecz obciążone zwiększonym ryzykiem trafiają do tzw. funduszy inwestycyjnych wysokiego ryzyka (*hedge funds*). Z kolei papiery bezpieczne i o dłuższym terminie zapadalności są najchętniej nabywane przez fundusze emerytalne i ubezpieczeniowe, które pozostają zainteresowane inwestowaniem na najdłuższe okresy.

Wysoka płynność oraz wysokie zróżnicowanie SPW pod względem dochodowości, ryzyka i okresów zapadalności sprawia, że stanowią one atrakcyjną formę aktywów dla banków komercyjnych i inwestycyjnych. Dodatkowym czynnikiem zwiększającym popyt na SPW ze strony banków pozostaje to, że są one

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wykorzystywane jako zabezpieczenie (*collateral*) w transakcjach z bankiem centralnym. W ramach tzw. operacji otwartego rynku (*open market operations*) banki komercyjne pożyczają pieniądz rezerwowy od banków centralnych pod zastaw posiadanych SPW (zob. rozdział 5). Banki centralne mogą też kupować i sprzedawać SPW w ramach operacji bezwarunkowych (*outright*). Kupując i sprzedając SPW w ramach warunkowych i bezwarunkowych operacji otwartego rynku, banki centralne regulują poziom rezerw pieniężnych (*reserve money, monetary base*) znajdujących się w posiadaniu banków komercyjnych i w ten sposób kształtują obieg pieniężny w kraju. Dlatego też operacje te, mimo że polegają na zakupie (lub sprzedaży) instrumentów dłużnych rządu, są traktowane jako instrument polityki pieniężnej, a nie polityki fiskalnej, ponieważ ich celem jest regulowanie ilości pieniądza w obiegu, a nie finansowane deficytu i długu publicznego.

Nie oznacza to jednak, że operacje otwartego rynku nie mają w ogóle wpływu na deficyt i dług. Skup SPW przez bank centralny zwiększa rezerwy pieniężne znajdujące się w posiadaniu banków komercyjnych, które dzięki temu mogą następnie kupować nowe instrumenty dłużne, sprzedawane przez rząd na rynku pierwotnym. Niewątpliwie ułatwia to rządowi finansowanie bieżącego deficytu. Ponadto bezwarunkowy skup SPW podnosi ich cenę i obniża stopy rentowności, co może skłanić rząd do zwiększania emisji obligacji i w efekcie do zwiększania zadłużenia. Efekt ten może też wystąpić w skali międzynarodowej, gdy banki centralne – w ramach zarządzania rezerwami walutowymi kraju – kupują obligacje emitowane przez rządy innych krajów. Przykładowo banki centralne wielu krajów Azji Wschodniej (Japonia, Chiny, Republika Korei) i krajów regionu Zatoki Perskiej (Arabia Saudyjska, Zjednoczone Emiraty Arabskie, Kuwejt, Katar) od wielu lat lokują swoje rezerwy walutowe głównie w obligacjach rządu Stanów Zjednoczonych. Stały popyt ze strony tych banków pomaga utrzymać wysokie ceny obligacji i niskie stopy rentowności, co z pewnością jest czynnikiem ułatwiającym rządowi USA prowadzenie ekspansywnej polityki fiskalnej[17].

Trzeba pamiętać, że finansowanie deficytu poprzez zaciąganie pożyczek również może prowadzić do monetyzacji. Gdy polityka pieniężna ma charakter akomodacyjny, bank centralny zwiększa podaż pieniądza, żeby sprostać zwiększonemu popytowi na pieniądz i utrzymać stopy procentowe na niezmienionym poziomie. Tak dzieje się zwłaszcza w warunkach recesji i w sytuacji zagrożenia kryzysem.

[17] Trzeba jednak zaznaczyć, że z przywileju tego mogą korzystać tylko kraje cieszące się największą wiarygodnością na rynkach finansowych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W pozostałych sytuacjach, a w szczególności, gdy tempo wzrostu PKB jest zbliżone do tempa potencjalnego, bank centralny na ogół stara się utrzymać stałe tempo wzrostu podaży pieniądza, co w warunkach zwiększonych wydatków publicznych finansowanych długiem przekłada się na wzrost stóp procentowych. Gdy zwiększony popyt rządowy konkuruje o ten sam zasób pieniądza z danym popytem sektora prywatnego, możemy mieć do czynienie z tzw. efektem „wypierania" (*crowding-out*) wydatków prywatnych przez wydatki publiczne.

3.4.3. Zaciąganie pożyczek za granicą

Zaciąganie pożyczek za granicą ma ekonomicznie niemal identyczny charakter jak zaciąganie pożyczek w kraju. Zamiast sprzedawać obligacje na rynku krajowym za walutę krajową, rząd sprzedaje obligacje na rynkach zagranicznych za waluty zagraniczne (dolary, euro, jeny). Różnica polega na tym, że obligacje sprzedawane za granicą są denominowane w walutach obcych, co oznacza, że spłata odsetek i samego zobowiązania musi także nastąpić w walutach obcych. Podobnie wygląda sytuacja, gdy rząd, zamiast sprzedawać obligacje, zaciąga pożyczki bezpośrednio w bankach zagranicznych. Tak samo, jak w przypadku zaciągania pożyczek bezpośrednio w bankach krajowych, ten sposób finansowania jest mniej dogodny i obecnie rzadko stosowany[18].

Zaciąganie pożyczek w walucie obcej wywołuje powstanie ryzyka kursowego[19]. Ponieważ spłata zobowiązania musi nastąpić w walucie obcej (walucie pożyczki), a ogromna większość dochodów i przychodów budżetu pozostaje w walucie krajowej, rząd musi kupować walutę obcą, aby spłacić zobowiązania wobec wierzycieli zagranicznych. Wyrażony w walucie krajowej przyszły koszt nabycia walut obcych

[18] W latach 1971–1980 Polska zaciągała pożyczki w bankach zagranicznych w celu sfinansowania zwiększonego importu inwestycyjnego i konsumpcyjnego. Część tych pożyczek była gwarantowana przez rządy państw zachodnich, a część udzielana przez banki na czysto komercyjnych zasadach. W tym czasie zadłużenie zagraniczne Polski wzrosło z niewiele ponad 1,5 mld do ponad 40 mld USD. Ciężar obsługi tego długu w walutach obcych okazał się tak duży w stosunku do możliwości finansowych państwa, że w połowie 1980 roku polski rząd zawiesił spłaty rat i odsetek od części długu gwarantowanej przez rządy zachodnie, a od pozostałej części długu płacił tylko odsetki, lecz nie spłacał kapitału. W rezultacie Polska znalazła się w stanie technicznego bankructwa (*default*). Problem został rozwiązany dopiero w 1994 roku, gdy polski rząd wynegocjował formalne umorzenie około połowy naszego zadłużenia wobec rządów państw-wierzycieli (Klub Paryski) oraz banków komercyjnych (Klub Londyński). Spłata pozostałej części zadłużenia została rozłożona na kolejne 18 lat. Ostatnie płatności z tego tytułu Polska uiściła w końcu 2012 roku.

[19] Zob. rozdział 5.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jest zawsze obciążony ryzykiem zmiany kursu walutowego. Na ogół ryzyko to jest tym większe, im dłuższy staje się okres kredytowania. W szczególności, gdy kurs waluty krajowej wobec walut obcych ulega osłabieniu, wielkość obciążeń z tytułu obsługi zadłużenia zagranicznego wzrasta. Gdy skala osłabienia kursu jest znaczna, spowodowany tym wzrost kosztów obsługi długu może być tak duży, że rząd nie będzie w stanie obsługiwać zadłużenia w ustalonych kwotach i terminach i stanie w obliczu niewypłacalności.

Mimo ryzyka kursowego, jest kilka istotnych powodów, dla których rządy decydują się zaciągać pożyczki za granicą. Po pierwsze, pozwala to na większą dywersyfikację źródeł kapitału, co zmniejsza zależność rządu od krajowego rynku kapitałowego. Po drugie, chodzi o większą dostępność kapitału. Emitowanie obligacji za granicą umożliwia dostęp do zagranicznych inwestorów instytucjonalnych, posiadających znaczne środki finansowe. Zasoby kapitału na rynkach zagranicznych są wielokrotnie większe niż na rynku pojedynczego kraju. W przypadku krajów mniej zamożnych, rozwijających się, zasób dostępnych oszczędności krajowych może być niewystarczający w celu sfinansowania krajowych potrzeb inwestycyjnych i deficytu budżetowego. Gdy na dodatek waluta krajowa jest mało wiarygodna i obciążona dużym ryzykiem kursowym, rząd będzie miał trudności ze sprzedażą obligacji na rynku krajowym, na skalę zapewniającą pokrycie potrzeb pożyczkowych. Po trzecie, z tego samego powodu oprocentowanie pożyczek denominowanych w walucie obcej jest z reguły niższe niż oprocentowanie pożyczek w walucie krajowej. Nawet gdy uwzględni się ryzyko kursowe, bardziej opłcalne może się okazać zaciągnięcie pożyczki za granicą niż w kraju. Wreszcie po czwarte, rządy państw często emitują obligacje denominowane w walutach obcych w sytuacji dobrej koniunktury, aby uzyskać rynkową wycenę wiarygodności kredytowej kraju. Obligacje takie, nazywane często obligacjami benchmarkowymi, stanowią podstawę służącą do ustalenia tzw. ratingu inwestycyjnego danego kraju przez międzynarodowe agencje ratingowe. Rządowe obligacje benchmarkowe stanowią także punkt odniesienia (*benchmark*) dla krajowych banków i przedsiębiorstw lokujących własne emisje instrumentów dłużnych na rynkach międzynarodowych.

3.4.4. Przychody ze sprzedaży aktywów

Przychody ze sprzedaży składników majątku oraz inne przychody, takie jak odszkodowania czy spłata udzielonego przez rząd kredytu, mają z definicji charakter jednorazowy i dlatego nie mogą być traktowane jako stałe źródło finansowania

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wydatków budżetowych. Na przykład przychody z prywatyzacji majątku państwowego nie mogą być stałym źródłem finansowania, ponieważ majątek państwowy jest ograniczony. Również przychody ze sprzedaży praw majątkowych, takich jak koncesje i licencje, nie mogą stanowić stałego źródła finansowania, gdyż zasób tych praw pozostaje również ograniczony.

Z tych przyczyn uznaje się, że przychody ze sprzedaży aktywów (składników majątku państwowego) nie powinny stanowić pokrycia wydatków budżetowych. Przychody tego typu należy raczej traktować jako źródło pokrycia zobowiązań, takich jak dług publiczny, czy też źródło finansowania przejściowych, jednorazowych kosztów związanych z wdrażaniem pojedynczych reform i zmian systemowych, takich na przykład jak reforma emerytalna.

Nieco inny charakter mają transfery zagraniczne w postaci środków otrzymywanych z budżetu Unii Europejskiej. Są one powiązane z konkretnymi wydatkami, a ich dokonanie wymaga podejęcia decyzji ze strony właściwych organów UE. Dlatego też, mimo że transfery takie mogą pojawiać się w kolejnych budżetach danego kraju przez kilka lub kilkanaście lat, ich każdorazowa wielkość jest zawsze niepewna i dlatego nie powinno się ich traktować jako stałego źródła finansowania deficytu.

3.5. Rachunkowe metody obliczania salda budżetu

Aby obliczyć wielkość salda budżetu w danym okresie, trzeba przyjąć pewne założenia dotyczące zakresu podmiotowego i przedmiotowego operacji finansowych sektora publicznego. W zależności od tego, jakie instytucje publiczne i jakie przepływy finansowe są objęte rachunkiem, wielkość salda budżetowego będzie bowiem różna. Stosowanie odmiennych metod liczenia tej zmiennej uzasadnia się tym, że różne ujęcia wielkości salda niosą różne informacje, przydatne w przypadku rozmiatych celów. Wybór konkretnej metody liczenia salda budżetowego zależy od tego, jaki aspekt zjawiska staje się przedmiotem zainteresowania.

3.5.1. Metoda kasowa i metoda memoriałowa

Najbardziej podstawowe rozróżnienie uwzględnia moment zarejestrowania danego przepływu finansowego. Z tej perspektywy rejestracja operacji budżetowych może być dokonana metodą kasową lub metodą memoriałową. W przypadku metody kasowej dana operacja jest rejestrowana w sprawozdawczości budżetowej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w momencie faktycznego przepływu środków pieniężnych. Z kolei zgodnie z metodą memoriałową dana operacja jest rejestrowana w momencie jej wystąpienia, to znaczy w chwili powstania (lub zmiany) zobowiązania bądź należności finansowej, najczęściej o charakterze wyprzedzającym w stosunku do przepływu środków pieniężnych. Ponieważ moment przepływu środków (moment dokonania płatności) z reguły nie pokrywa się z momentem powstania zobowiązania bądź należności (którym jest z reguły moment wystawienia faktury lub rachunku), wielkość operacji finansowych rejestrowanych metodą kasową w danej chwili będzie różnić się od wielkości operacji rejestrowanych metodą memoriałową.

Należy od razu zaznaczyć, że obie metody są rachunkowo i metodologicznie poprawne, choć dają różne wyniki. A zatem wybór jednej spośród nich powinien być podyktowany potrzebami informacyjnymi oraz względami praktycznymi. W Polsce od wielu lat w finansach publicznych obowiązuje metoda kasowa, zgodnie z przepisami zawartymi w ustawie o finansach publicznych z 2009 roku. Z kolei w UE stosuje się metodę memoriałową, na podstawie ujednoliconego i wprowadzonego w życie w 2014 roku Systemu Rachunków Narodowych i Regionalnych (*System of National and Regional Accounts* – tzw. ESA 2010)[20]. Prawo unijne nie narzuca jednak krajom członkowskim obowiązku harmonizacji przepisów dotyczących sposobu rejestrowania operacji w sektorze finansów publicznych. Każde państwo członkowskie Unii Europejskiej może więc dla własnych potrzeb stosować osobną, krajową metodologię, w tym metodę kasową rejestracji operacji finansowych. Wymogi dotyczące członkostwa w Unii Europejskiej, w tym zwłaszcza potrzeby jednolitej sprawozdawczości budżetowej prowadzonej przez służby statystyczne UE (Eurostat), sprawiają jednak, że równolegle, obok metody kasowej, w Polsce do rozliczania operacji budżetowych stosowana jest także metoda memoriałowa. Najważniejsze dane statystyczne dotyczące polskich finansów publicznych są więc podawane do ogólnej wiadomości w dwóch ujęciach rachunkowych, nazywanych czasami metodologią krajową i metodologią unijną (ESA 2010).

Należy pamiętać, że zawartość informacyjna wielkości salda budżetowego, obliczanego na podstawie obu metod, jest różna. Zasady ESA 2010 definiują deficyt budżetowy konsekwentnie jako przyrost zadłużenia netto (*net borrowing*), a nadwyżkę budżetową jako przyrost należności netto (*net lending*) w danym okresie. W efekcie metoda ESA 2010 prezentuje zmianę sytuacji majątkowej (stanu aktywów

[20] Metodologia ESA 2010 zastąpiła stosowaną wcześniej w UE metodologię ESA 1995 (zob. też punkt 3.2).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

netto) danego podmiotu (np. Skarb Państwa), podczas gdy metoda kasowa ukazuje różnice w faktycznych przepływach finansowych w danym okresie. Metoda kasowa określa więc skalę potrzeb pożyczkowych budżetu (więcej na ten temat pisze Wernik 2011).

3.5.2. Zakres operacji finansowych objętych rachunkiem

Stosowanie metody kasowej w polskiej praktyce budżetowej nie jest jedyną różnicą między krajową metodologią liczenia salda budżetu a metodologią unijną (ESA 2010). Różny jest także sposób rozliczenia przychodów z prywatyzacji. Teoretycznie – jako przychody jednorazowe – powinny one stanowić źródło pokrycia potrzeb pożyczkowych, lecz w praktyce są przeznaczane na różne ściśle określone wydatki, takie jak dotacja do FUS oraz subwencje dla innych funduszy. Ponieważ suma tych wydatków czasami przekracza przychody z prywatyzacji, w polskiej rachunkowości budżetowej funkcjonuje dość kuriozalna kategoria – „ujemne saldo przychodów z prywatyzacji". Ten „niedobór" środków na pokrycie niektórych wydatków nie jest jednak, według metodologii krajowej, wliczany do deficytu budżetowego, lecz do deficytu obliczanego metodą unijną. Taki sposób rejestracji tych transakcji niewątpliwie zaciemnia obraz sytuacji w finansach publicznych.

Oprócz tego między metodologią krajową a metodologią unijną występują różnice zakresowe. Innymi słowy, zakres przedmiotowy sektora finansów publicznych według definicji krajowej nie pokrywa się z zakresem „sektora instytucji rządowych i samorządowych" według definicji unijnej (*general government*). Zasadnicza rozbieżność polega na tym, że metodologia krajowa nie zalicza do sektora finansów publicznych rachunków Krajowego Funduszu Drogowego, który od 2009 roku przejął od budżetu państwa zadania związane z finansowaniem budowy dróg krajowych i autostrad[21]. Z kolei według ESA 2010 (i dodajmy, zgodnie z logiką), KFD jest częścią sektora instytucji rządowych i samorządowych (*general government*). Całkowite zadłużenie KFD z tytułu zaciągniętych pożyczek i wyemitowanych

[21] Krajowy Fundusz Drogowy to fundusz celowy, stanowiący wsparcie w przypadku realizacji rządowego Programu Budowy Dróg i Autostrad poprzez gromadzenie środków finansowych na przygotowanie, budowę, przebudowę, remonty, utrzymanie i ochronę autostrad, dróg ekspresowych, a także innych dróg krajowych. Fundusz został utworzony w Banku Gospodarstwa Krajowego na mocy ustawy z dnia 27 października 1994 roku o autostradach płatnych oraz o Krajowym Funduszu Drogowym (z późn. zm.). Fundusz rozpoczął funkcjonowanie z dniem 1 stycznia 2004 roku. Od 2009 roku KFD finansuje wszystkie inwestycje drogowe przewidziane w rządowym wieloletnim programie budowy dróg krajowych do 2015 roku.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

obligacji wynosiło na koniec 2014 roku około 44,6 mld PLN. Dług ten nie jest zaliczany do długu publicznego na podstawie metodologii krajowej, ale stanowi część długu publicznego według metodologii ESA 2010.

3.5.3. Salda budżetowe poszczególnych części sektora finansów publicznych

Powracając do różnych sposobów liczenia salda budżetowego pod względem zakresu operacji objętych sprawozdawczością, niewątpliwie największe znaczenie ekonomiczne ma saldo ogólne całego sektora finansów publicznych. Ukazuje ono najpełniej stopień nierównowagi budżetowej oraz jej wpływ na poziom zadłużenia państwa jako całości. Dlatego wielkość tego salda, a zwłaszcza wielkość deficytu całego sektora finansów publicznych, stanowi jeden z najważniejszych parametrów określających bieżącą sytuację makroekonomiczną danej gospodarki. Często potrzebna jest jednak bardziej szczegółowa informacja o strukturze tego deficytu, w tym zwłaszcza w podziale na deficyt budżetu centralnego oraz deficyt pozostałych części sektora finansów publicznych. Dane dotyczące deficytu całego sektora mogą bowiem przesłaniać istotne różnice występujące w zakresie sytuacji budżetowej poszczególnych podsektorów, a takie informacje są potrzebne w celu prawidłowego zaprogramowania działań przywracających równowagę budżetową. Co więcej, potrzeba tego rozróżnienia wynika także z innej procedury prawnej i politycznej, stosowanej przy uchwalaniu budżetu państwa oraz budżetów pozostałych części sektora. W większości krajów, w tym w Polsce, największe znaczenie ma budżet państwa (budżet centralny), który jest przygotowywany przez Radę Ministrów i uchwalany przez Sejm w formie ustawy. Budżety jednostek samorządu terytorialnego są uchwalane przez organy stanowiące tych jednostek na podstawie ogólnych wytycznych określających ramy dyscypliny finansowej i zasady dobrego gospodarowania środkami publicznymi. Z kolei budżety funduszy i agencji budżetowych w formie planów finansowych, po zatwierdzeniu przez Radę Ministrów, stanowią załączniki do ustawy budżetowej i wraz z nią są uchwalane przez Sejm.

3.5.4. Saldo pierwotne, saldo kapitałowe, saldo zagraniczne

Oprócz podziału na saldo całego sektora i salda cząstkowe poszczególnych części sektora finansów publicznych, stosuje się także inne klasyfikacje. Najczęściej z salda całego sektora finansów publicznych wyodrębnia się salda cząstkowe dotyczące

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pewnych wybranych kategorii operacji budżetowych. Z tego punktu widzenia najbardziej istotne okazuje się rozróżnienie między następującymi kategoriami:

a) Różnica między saldem rzeczywistym (*actual balance*) a saldem pierwotnym (*primary balance*). Saldo rzeczywiste to faktyczna różnica pomiędzy wszystkimi otrzymanymi dochodami a wszystkimi zrealizowanymi wydatkami budżetu w danym okresie czasu. Z kolei saldo pierwotne to różnica między dochodami budżetowymi a wydatkami pomniejszonymi o koszty obsługi długu. Ponieważ koszty obsługi długu (płatności odsetkowe) są w danym okresie wydatkiem, na który rząd nie ma wpływu, zmiany w poziomie salda pierwotnego dostarczają bardziej wiarygodnej informacji o rzeczywistych zmianach w polityce fiskalnej rządu.
b) Różnica między saldem bieżącym a saldem kapitałowym budżetu. Pierwsze z nich jest wynikiem operacji bieżących, drugie zaś wynika z operacji kapitałowych budżetu. Rozróżnienie to pozwala upewnić się, że bieżące wydatki budżetu znajdują pełne pokrycie w bieżących dochodach, a ewentualne pożyczki są przeznaczane jedynie na pokrycie wydatków inwestycyjnych. Reguła ta bywa nazywana „złotą regułą" finansów publicznych (*golden rule*). Zgodnie z tą zasadą, bieżące saldo budżetu powinno być równe zeru, a pożyczki mogą być zaciągane jedynie na pokrycie wydatków inwestycyjnych. Zasada ta znajduje uzasadnienie w tym, że inwestycje tworzą majątek trwały (produkcyjny lub w postaci infrastruktury), który będzie w przyszłości albo sam generował określony dochód, albo przyczyniał się do wzrostu dochodu wytwarzanego przez sektor prywatny. Z jednej strony oznacza to, że rosnącym zobowiązaniom rządu towarzyszy taki sam wzrost posiadanych aktywów, z drugiej zaś zapewnia to strumień dochodów podatkowych w przyszłości, mogących stanowić źródło spłaty zaciągniętych kredytów[22].
c) Różnica między saldem zagranicznym a saldem krajowym. Pierwsze z nich z reguły jest ujemne, ponieważ rządy mają na ogół niewielkie dochody zagraniczne, często zaciągają natomiast pożyczki na rynkach zagranicznych, które potem muszą regularnie obsługiwać. Skala deficytu w transakcjach zagranicznych pozostaje ważnym wskaźnikiem wrażliwości budżetu na zmiany kursu walutowego.

[22] „Złota reguła" była stosowana m.in. w Niemczech od lat 60. XX wieku, na podstawie odpowiedniego zapisu w niemieckiej konstytucji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.5.5. Saldo strukturalne budżetu

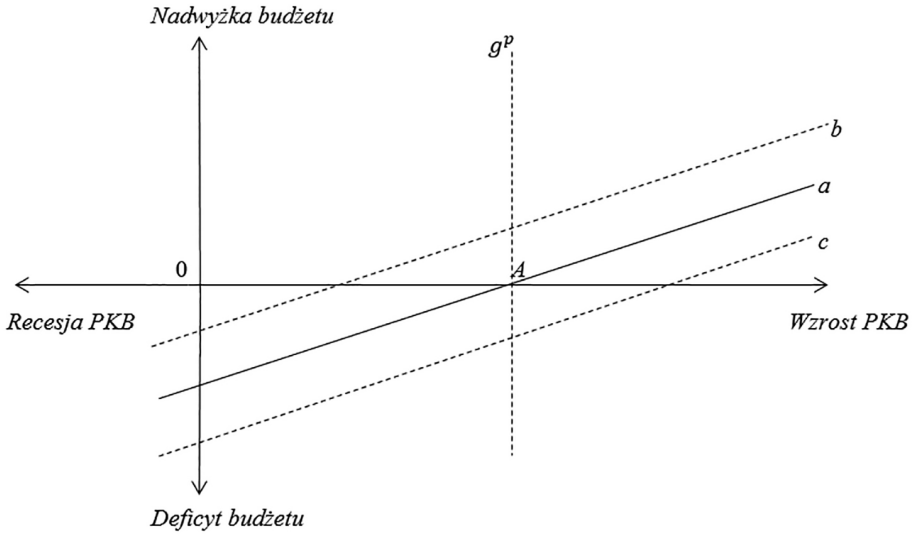
Saldo strukturalne budżetu (*structural balance, cyclically-adjusted balance*) to saldo uwzględniające wpływ cyklu koniunkturalnego na saldo faktyczne (*actual balance*). Rozróżnienie to jest bardzo ważne przy określaniu rzeczywistego stopnia restrykcyjności lub ekspansywności polityki fiskalnej, a więc w celu oceny nastawienia polityki fiskalnej (*fiscal policy stance*). Punkt wyjścia stanowi przyjęcie, że wielkość i znak faktycznego salda budżetu zależy nie tylko od samych decyzji rządu w sprawie wielkości dochodów i wydatków budżetowych, ale także od stanu koniunktury gospodarczej, ponieważ zarówno dochody jak i wydatki budżetowe reagują na zmiany koniunktury i przebieg cyklu koniunkturalnego. W czasie recesji dochody podatkowe spadają na skutek obniżania się zysków przedsiębiorstw, spadku dochodów indywidualnych i spadku sprzedaży, wydatki mogą natomiast rosnąć (na przykład ze względu na wzrost wydatków na zasiłki dla bezrobotnych). W rezultacie pojawia się tendencja do deficytu, lecz deficyt ten nie jest wynikiem niezależnych decyzji władz dążących do zwiększenia wydatków budżetowych, a raczej ma charakter endogeniczny, spowodowany zmianą koniunktury. Podobnie w okresie dobrej koniunktury dochody podatkowe szybko rosną, podczas gdy wydatki mogą maleć, i w efekcie pojawia się tendencja do występowania nadwyżki w budżecie. Zasadniczo przy braku znacznych dyskrecjonalnych zmian na poziomie budżetu dochody i wydatki budżetowe zachowują się w sposób antycykliczny w przebiegu cyklu, przy czym zmiany po stronie dochodów są relatywnie silniejsze niż po stronie wydatków. Antycykliczny – czyli łagodzący przebieg cyklu – wpływ endogenicznych zmian dochodów i wydatków na zagregowany popyt w przebiegu cyklu koniunkturalnego oznacza, że ograniczają one spadek popytu w czasie recesji oraz wzrost popytu w okresie ekspansji. Zmiany te pełnią zatem funkcję tzw. automatycznych stabilizatorów koniunktury (*automatic stabilizers*).

Ocena stopnia restrykcyjności lub ekspansywności polityki fiskalnej nie może zatem być dokonywana jedynie na podstawie znaku i wielkości salda faktycznego, ponieważ nie wiadomo, w jakim stopniu odzwierciedla ono politykę budżetową rządu, a w jakim jest wynikiem aktualnego stanu koniunktury gospodarczej. W tym celu należy określić saldo strukturalne, tj. *hipotetyczne* saldo budżetu, jakie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wystąpiłoby w warunkach „normalnej" koniunktury, to znaczy gdyby gospodarka rosła w tempie równym potencjalnej stopie wzrostu PKB[23, 24].

Rysunek 3.2. Automatyczne stabilizatory koniunktury a stopień restrykcyjności polityki fiskalnej



Źródło: opracowanie własne.

Zależność salda budżetu od fazy cyklu koniunkturalnego przedstawia rysunek 3.2. Automatyczne stabilizatory koniunktury sprawiają, że im szybszy jest wzrost PKB, tym większe dochody budżetowe i tym większe saldo budżetu. W warunkach równowagi, kiedy PKB rośnie w tempie potencjalnym g^p, saldo budżetu powinno być równe zeru – gdy tempo wzrostu PKB jest niższe, budżet wykazuje deficyt,

[23] Zgodnie z interpretacją obowiązującą w Unii Europejskiej, saldo strukturalne budżetu jest równe saldu rzeczywistemu skorygowanemu o wpływ fazy cyklu koniunkturalnego oraz o jednorazowe i przejściowe wydatki i dochody. Tak zdefiniowane saldo strukturalne wykorzystywane jest w Unii Europejskiej podczas oceny stopnia realizacji średniookresowego celu budżetowego, będącego częścią procedury prewencyjnej kontroli i nadzoru nad polityką fiskalną państw członkowskich w ramach Paktu Stabilności i Wzrostu (więcej na ten temat w rozdziale 8). Tak określone saldo strukturalne jest równe zeru (lub jest zbliżone do zera), gdy gospodarka rośnie w tempie równym potencjalnemu tempu wzrostu.

[24] Potencjalne tempo wzrostu PKB (lub potencjalny PKB) to takie tempo wzrostu, które pozwala wykorzystać w pełni zasoby produkcyjne, w tym zwłaszcza zasoby siły roboczej, ale nie powoduje jeszcze napięć inflacyjnych. W przypadku potencjalnego PKB bezrobocie jest równe bezrobociu naturalnemu. Różnica pomiędzy rzeczywistym tempem wzrostu a potencjalnym tempem wzrostu nazywana jest luką podażową lub produktową (*output gap*).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

a kiedy tempo wzrostu staje się wyższe, budżet wykazuje nadwyżkę. Tę zależność ilustruje linia *a* na rysunku 3.2. Gdy saldo budżetu zmienia się wzdłuż linii *a*, polityka fiskalna jest neutralna względem wzrostu gospodarczego na poziomie potencjalnym, to znaczy ani go nie pobudza, ani go nie hamuje. Linia *b* odpowiada restrykcyjnej polityce fiskalnej, ponieważ w warunkach potencjalnego tempa wzrostu budżet wykazuje nadwyżkę (co oznacza, że sektor rządowy obniża popyt łączny w gospodarce), z kolei linia *c* odzwierciedla ekspansywną politykę fiskalną, gdyż w warunkach potencjalnego tempa wzrostu budżet wykazuje deficyt (a zatem sektor rządowy zwiększa łączny popyt w gospodarce).

Definiując saldo cykliczne jako część salda faktycznego wynikającą z danej fazy koniunktury (jest ono równe zeru w przypadku potencjalnego PKB), można przedstawić saldo faktyczne (rzeczywiste) jako sumę salda cyklicznego i salda strukturalnego. W normalnych warunkach deficyt strukturalny wykazuje mniejsze wahania z okresu na okres niż deficyt faktyczny. Teoretycznie saldo strukturalne powinno odzwierciedlać wyłącznie skutki dyskrecjonalnych decyzji rządu, dotyczących dochodów i wydatków w zakresie budżetu.

Koncepcja strukturalnego salda budżetu opiera się na pojęciu potencjalnego PKB i luki podażowej jako różnicy między rzeczywistym a potencjalnym PKB. Pojęcia te, choć pozostają bardzo przydatne w analizie ekonomicznej, nie są bezpośrednio obserwowalne, w związku z czym muszą być estymowane na podstawie modeli ekonomicznych, przy przyjęciu określonych założeń (zob. ramka 3.3 i ramka 3.4).

Ramka 3.3. Deficyt strukturalny

Strukturalne saldo budżetu to saldo rzeczywiste skorygowane o wpływ cyklu koniunkturalnego. Innymi słowy, saldo strukturalne to takie saldo, jakie wystąpiłoby gdyby PKB był na poziomie potencjalnym, czyli zapewniającym pełne zatrudnienie i stabilność cen. Oznaczmy saldo faktyczne budżetu jako *b*, saldo strukturalne jako *s*, a saldo cykliczne jako *c* (wszystkie zmienne jako % PKB). Wówczas z definicji otrzymujemy:

$$b \equiv s + c. \tag{3.10}$$

Aby oszacować saldo strukturalne w danym roku, należy najpierw określić, w której fazie cyklu koniunkturalnego znajduje się gospodarka, a więc na ile rzeczywiste PKB odchyla się od potencjalnego PKB. Różnica pomiędzy PKB rzeczywistym

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

a PKB potencjalnym stanowi lukę podażową (lub lukę produktową – *output gap*) – oznaczmy ją jako $(y - \bar{y})$. Saldo rzeczywiste ulega zmianom wraz ze zmianami luki podażowej, ponieważ zmienia się saldo cykliczne. Wrażliwość salda rzeczywistego budżetu wobec zmian luki podażowej ε wynosi:

$$\varepsilon = \frac{db}{d\left(y - \bar{y}\right)} > 0, \tag{3.11}$$

gdzie współczynnik ε wskazuje, o ile % PKB wzrośnie (spadnie) saldo budżetowe, gdy luka podażowa wzrośnie (spadnie) o 1% PKB. Ponieważ saldo cykliczne $c = \varepsilon(y - \bar{y})$, saldo strukturalne będzie wynosić:

$$s = b - \varepsilon(y - \bar{y}). \tag{3.12}$$

Gdy rzeczywisty PKB równa się potencjalnemu PKB $y = \bar{y}$, a $s = b$. Z kolei gdy $y > \bar{y}$, wyrażenie w nawiasie jest dodatnie, wówczas $s < b$, i odwrotnie – gdy $y < \bar{y}$, to $s > b$.

Zmienna c pozostaje zależna od wielkości luki podażowej, która jest zmienną nieobserwowalną. Komisja Europejska w sposób systematyczny szacuje wielkości luki podażowej oraz potencjalnego PKB odnośnie do poszczególnych krajów UE na podstawie dość skomplikowanej metodologii. Inne międzynarodowe organizacje, w tym Międzynarodowy Fundusz Walutowy oraz OECD, stosują własne metodologie szacowania potencjalnego PKB i luki podażowej.

Ramka 3.4. Produkt potencjalny i luka podażowa*

Produkt potencjalny to taki poziom produkcji (PKB), który pozwala na pełne wykorzystanie zasobów produkcyjnych, a zarazem nie wywołuje inflacji. Przyjmijmy, że produkcja w danej gospodarce jest funkcją zasobu pracy L, zasobu kapitału K i produktywności łącznej obu tych zasobów A (*total factor productivity* – TFP). Wygodnie jest przedstawić tę zależność w postaci popularnej funkcji produkcji Cobba–Douglasa:

$$Y = A \cdot L^{\alpha} K^{1-\alpha}, \tag{3.13}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gdzie α i $(1 - \alpha)$ oznaczają elastyczność produkcji kolejno względem zasobu pracy i zasobu kapitału. Aby określić poziom potencjalnego PKB, czyli $Y(P)$, należy oszacować potencjalny poziom zatrudnienia $L(P)$. Można go zdefiniować, wykorzystując następujące równanie:

$$L(P)_t = WP_t PR_t^T (1 - NAWRU_t), \tag{3.14}$$

gdzie WP_t jest liczbą osób w wieku produkcyjnym w okresie t, PR_t^T wyraża długookresową stopę aktywności zawodowej (superskrypt T oznacza wartość trendu), a $NAWRU_t$ stanowi stopę bezrobocia nieprzyspieszającą inflacji płacowej (*non-accelerating wage* (*inflation*) *rate of unemployment*). Wówczas potencjalny PKB może być przedstawiony w następującej postaci:

$$Y(P)_t = \left[WP_t PR_t^T \left(1 - NAWRU_t\right)\right]^{\alpha} K_t^{1-\alpha} A_t^T. \tag{3.15}$$

Lukę podażową (*output gap*) daje się więc zdefiniować jako różnicę między rzeczywistym a potencjalnym PKB, wyrażoną w procentach. Oznaczając cykliczny składnik stopy bezrobocia jako $C_t^U = U_t - NAWRU_t$, możemy przedstawić lukę podażową $Y(GAP)_t$ jako:

$$Y(GAP)_t = \alpha\left[PR_t - PR_t^T\right) + \left(NAWRU_t - U_t\right] + A(GAP)_t. \tag{3.16}$$

Równanie (3.16) pokazuje, że źródło luki podażowej stanowią odchylenia rzeczywistej stopy aktywności zawodowej, stopy bezrobocia i łącznej produktywności zasobów od odpowiednich wartości trendu[25].

3.5.6. Deficyt faktyczny i deficyt strukturalny w Polsce w latach 1995–2014

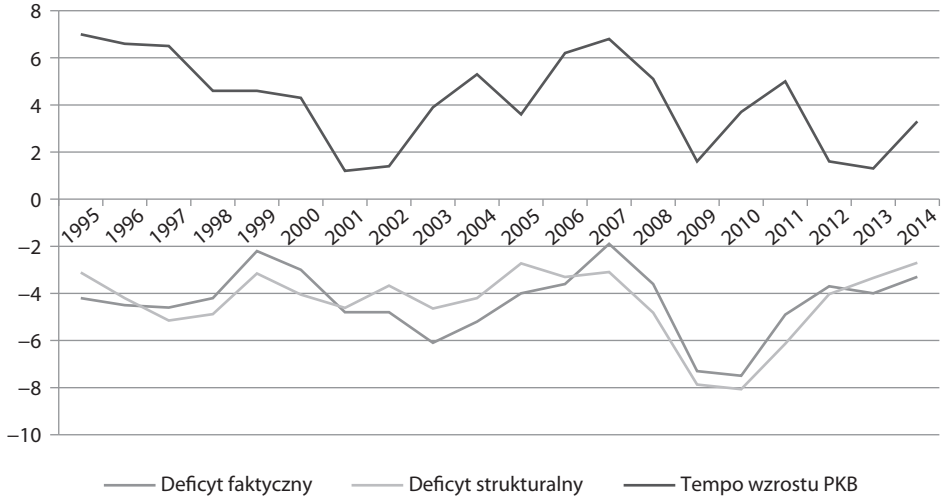
Zgodnie z teoretyczną zależnością opisaną na rysunku 3.2, automatyczne stabilizatory koniunktury sprawiają, że w okresach lepszej koniunktury faktyczne saldo budżetu powinno być dodatnie, a w okresach słabej koniunktury – ujemne.

[25] Szczegółowy opis metodologii liczenia produktu potencjalnego i luki podażowej można znaleźć w pracy Denis, McMorrow i Roeger (2002).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Po uwzględnieniu składnika cyklicznego saldo strukturalne powinno być zbliżone do zera w całym okresie cyklu koniunkturalnego. Czy kształtowanie się salda faktycznego i salda strukturalnego w Polsce pozostaje zgodne z tymi prawidłowościami? Wykres 3.6 przedstawia wielkości deficytu faktycznego i deficytu strukturalnego w Polsce w okresie 1995–2014, na tle danych dotyczących tempa wzrostu PKB w poszczególnych latach.

Wykres 3.6. Deficyt faktyczny, deficyt strukturalny i tempo wzrostu PKB w Polsce w latach 1995–2014 (dane w % PKB)



Uwaga: deficyt strukturalny zdefiniowany jako wielkość potrzeb pożyczkowych sektora finansów publicznych, skorygowana o wpływ cyklu koniunkturalnego na podstawie wartości trendu PKB, jako % PKB.

Źródło: opracowanie własne na podstawie danych bazy AMECO.

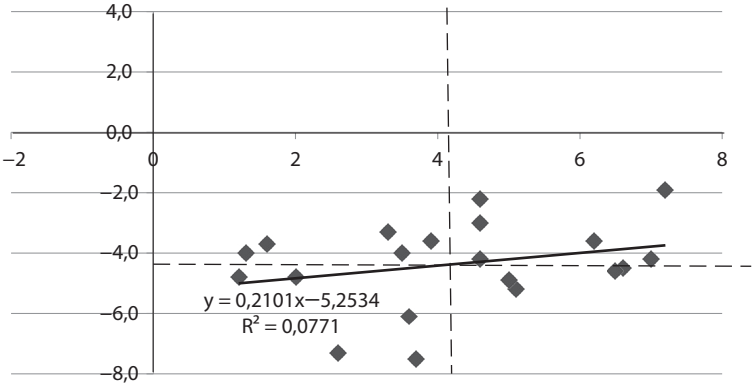
Jak można zauważyć, w całym analizowanym dwudziestoleciu saldo budżetowe, zarówno faktyczne, jak i strukturalne, było ujemne, niezależnie od fazy cyklu koniunkturalnego, przy czym wielkość deficytu faktycznego wahała się w przedziale między –1,9% a –7,5% PKB, podczas gdy wielkość deficytu strukturalnego w przedziale między –2,7% a –8,1% PKB. Nawet w okresach bardzo szybkiego wzrostu gospodarczego (lata 1995–2000 oraz 2004–2008) nie tylko ani razu nie pojawiła się nadwyżka budżetowa, lecz występował także spory deficyt strukturalny (średnio 4,1% i 3,6% PKB). Jak pamiętamy, saldo strukturalne mierzy wpływ dyskrecjonalnych decyzji rządu w sprawie dochodów i wydatków budżetowych. Dlatego też wysokie ujemne saldo strukturalne oznacza, że nastawienie polityki fiskalnej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(*fiscal stance*) było w całym tym okresie nadmiernie ekspansywne. Jeśli przyjąć, że potencjalne tempo wzrostu w tym okresie było równe średniej geometrycznej rocznych stóp wzrostu PKB i wynosiło około 4,2%[26], to przy tempie wzrostu PKB na poziomie potencjalnym deficyt budżetowy dodatkowo zwiększał popyt globalny ponad poziom równowagi średnio o wielkość równą 4,4% PKB.

Aby ocenić *fiscal stance*, istotne jest również ustalenie, czy polityka fiskalna miała charakter procykliczny czy antycykliczny, a zatem czy wzmacniała wahania cyklu, czy raczej je łagodziła. Można się o tym przekonać, badając korelację między zmianami faktycznego deficytu a zmianami PKB. W przypadku Polski korelacja między tempem wzrostu PKB a wielkością salda budżetu była wprawdzie dodatnia, lecz jednocześnie niezbyt silna – współczynnik korelacji między deficytem faktycznym a tempem wzrostu PKB wynosi 0,28. Można zatem wnioskować, że polityka fiskalna w Polsce była w badanym okresie łagodnie antycykliczna.

Wykres 3.7. Zależność między saldem sektora finansów publicznych a tempem wzrostu PKB w Polsce w okresie 1995–2014



Źródło: opracowanie własne na podstawie danych bazy AMECO.

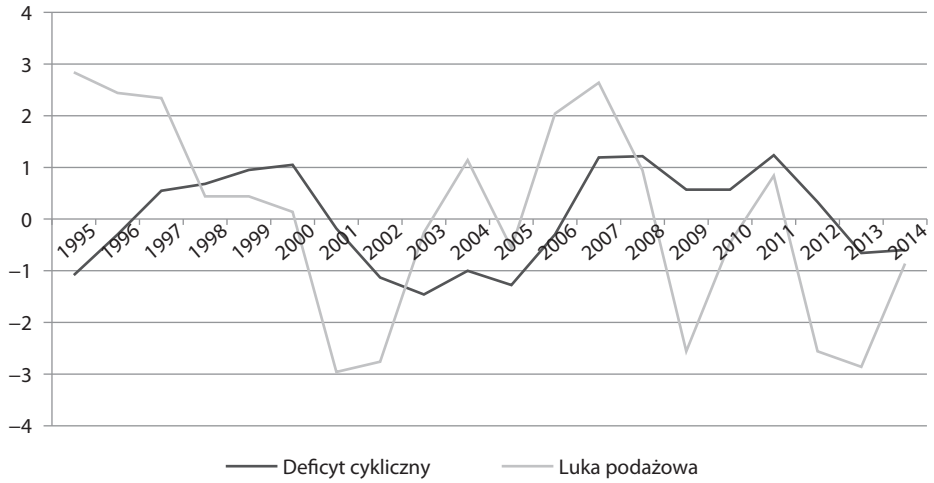
Nadmiernie ekspansywne nastawienie polityki fiskalnej w całym omawianym okresie jest również wyraźnie widoczne, gdy zestawi się roczne wartości salda sektora finansów publicznych i roczne wartości tempa wzrostu PKB. Wykres 3.7 przedstawia pary obserwacji dla lat 1995–2014. Na osi poziomej odłożono roczne tempo wzrostu PKB, a na osi pionowej saldo sektora finansów publicznych jako % PKB.

[26] Średnia geometryczna stóp wzrostu PKB w latach 1995–2014 wyniosła 4,16% rocznie. Przybliżanie potencjalnego tempa wzrostu za pomocą średniej wartości za dany okres jest oczywiście uproszczeniem, ponieważ potencjalne tempo wzrostu ulega zmianom w czasie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Jak widać, zależność między tymi zmiennymi jest rosnąca, podobna do opisanej na rysunku 3.2, ale krzywa dotycząca Polski przebiega znacznie niżej niż krzywa *a* zapewniająca równowagę średniookresową i zerowe saldo strukturalne przy potencjalnym wzroście PKB. W przypadku Polski potencjalnemu tempu wzrostu PKB na poziomie 4,2% odpowiada – średnio – strukturalny deficyt na poziomie 4,4% PKB. Oznacza to, że w stosunku do średniego nastawienia polityki fiskalnej w latach 1995–2014, w Polsce powinno nastąpić trwałe zacieśnienie polityki fiskalnej o około 3–4% PKB.

Wykres 3.8. Deficyt cykliczny i luka podażowa w Polsce w latach 1995–2014



Uwaga: deficyt cykliczny odpowiadający różnicy między deficytem faktycznym a deficytem strukturalnym, zgodnie z tożsamością (3.10). Wielkości deficytu strukturalnego zaczerpnięto z bazy danych statystycznych Komisji Europejskiej (baza AMECO), a wielkości luki podażowej zostały oszacowane jako różnica między rzeczywistym tempem wzrostu PKB a tempem potencjalnym, przy czym za wartość potencjalnego tempa wzrostu PKB przyjęto średnią geometryczną stóp wzrostu PKB za lata 1995–2014, wynoszącą 4,16%.

Źródło: opracowanie własne na podstawie bazy danych AMECO.

Siłę działania automatycznych stabilizatorów koniunktury mierzy stopień zależności deficytu cyklicznego od luki podażowej (zgodnie z równaniem 3.12, deficyt cykliczny jest funkcją luki podażowej). Jak pokazuje wykres 3.8, zależność ta w przypadku Polski była bardzo słaba[27].

[27] Zgodnie z równaniem regresji salda cyklicznego względem luki podażowej oszacowanym dla lat 1995–2014 o postaci:

$$y = a + bx \Rightarrow y = 0{,}016 + 0{,}089x,\ R^2 = 0{,}036,$$

wzrostowi luki podażowej o 1% PKB towarzyszył wzrost salda cyklicznego zaledwie o 0,09% PKB.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.5.7. Potrzeby pożyczkowe sektora publicznego

Inną jeszcze miarę nierównowagi budżetowej stanowi wielkość potrzeb pożyczkowych sektora publicznego. Jest to kwota pożyczek, które trzeba zaciągnąć w danym okresie, aby zarówno pokryć planowany deficyt, jak i uzyskać środki na spłatę zapadających pożyczek, zaciągniętych w okresach poprzednich. Te ostatnie operacje finansowe nazywa się rolowaniem długu lub refinansowaniem długu. Z tego powodu potrzeby pożyczkowe pozostają oczywiście zawsze większe niż deficyt budżetu. Trzeba przy tym pamiętać, że tak określone potrzeby pożyczkowe zależą nie tylko od wielkości deficytu budżetowego w danym roku, lecz także od terminowej struktury zapadających płatności. Tabela 3.5 przedstawia wielkości deficytu sektora finansów publicznych oraz potrzeb pożyczkowych w Polsce w latach 2012–2016. Jak widać, kwoty związane z wykupem zapadającego długu były z reguły znacznie większe niż wartość bieżącego deficytu, a jego udział w całości potrzeb pożyczkowych stanowił średnio około 37%. Można także zauważyć, że to głównie wahania wielkości spłat zapadającego długu wpływały na zmiany potrzeb pożyczkowych brutto.

Tabela 3.5. Deficyt sektora finansów publicznych i planowane potrzeby pożyczkowe w Polsce w latach 2012–2016 (dane w mld PLN)

Rok	Deficyt	Wykup długu i pozostałe potrzeby pożyczkowe netto	Potrzeby pożyczkowe brutto*	Udział deficytu w potrzebach pożyczkowych
2012	38,2	116,6	154,8	24,7%
2013	48,3	90,9	139,2	34,7%
2014	62,5	70,1	132,6	47,1%
2015	57,6	97,2	154,8	37,2%
2016	69,8**	112,9	182,7	38,3%

* Suma planowanych potrzeb netto oraz przypadającego do wykupu długu w ustawie budżetowej według metodologii ESA 2010.
** Wg projektu budżetu.

Źródło: opracowanie własne na podstawie danych Ministerstwa Finansów.

Z punktu widzenia zarządzania płynnością budżetu potrzeby pożyczkowe pozostają wskaźnikiem ważniejszym niż sam deficyt. Również rynki finansowe obserwują z uwagą zmiany potrzeb pożyczkowych, ponieważ ich skala decyduje o wielkości podaży SPW na rynku kapitałowym, a z kolei wielkość podaży SPW ma

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wpływ na ich cenę. Gdy potrzeby pożyczkowe rządu są duże, to, nawet gdy bieżący deficyt budżetowy okazuje się niewielki, rząd zmuszony jest sprzedawać większe ilości obligacji, co wywołuje presję na spadek cen i wzrost rentowności obligacji.

3.6. Ekonomiczne konsekwencje deficytu budżetowego

Ogólnie rzecz biorąc, deficyty budżetowe nie cieszą się uznaniem. Występowanie deficytu świadczy o tym, że wydatki publiczne są większe niż dochody, a to z kolei jest powszechnie traktowane jako przejaw życia ponad stan, lekkomyślnej rozrzutności i złej polityki. Często tak w rzeczywistości bywa, lecz ta jednostronna opinia pozostaje nieuzasadniona. Deficyt budżetowy może powstać na skutek nieudolnej, a nawet nieodpowiedzialnej polityki gospodarczej, ale w wielu sytuacjach miewa też charakter endogeniczny i znajduje swoje ekonomiczne uzasadnienie. Podobnie jak ma to miejsce w przypadku wielu innych kwestii, ostateczna ocena deficytu zależy od danej sytuacji.

Przede wszystkim trzeba podkreślić, że deficyt budżetowy sam z siebie nie musi być czymś złym. Pierwszą okolicznością, jaką należy wziąć pod uwagę, jest nierównomierny rozkład dochodów i wydatków w czasie oraz związane z tym okresowo pojawiające się przejściowe deficyty i nadwyżki. W przypadku gospodarstw domowych i poszczególnych jednostek rygorystyczne, każdorazowe dostosowanie wydatków do zmieniających się dochodów powodowałoby duże wahania poziomu konsumpcji w czasie. Racjonalne gospodarstwa domowe unikają takiej sytuacji, starając się wygładzać konsumpcję w dłuższym horyzoncie czasowym. Uzasadnieniem w przypadku takiego postępowania jest malejąca krańcowa użyteczność konsumpcji. Dlatego też z punktu widzenia gospodarstw domowych i poszczególnych jednostek warto zaciągać pożyczki na utrzymanie osiągniętego poziomu konsumpcji w okresach niskich dochodów i spłacać te pożyczki w okresach wysokich dochodów, tak aby w cyklu życia konsumpcja ulegała wygładzeniu. Oznacza to, że gospodarstwa domowe mogą w pewnych okresach notować nadwyżkę wydatków nad dochodami (czyli deficyt) i finansować ten deficyt długiem pod warunkiem, że w innych okresach będą realizować nadwyżkę dochodów nad wydatkami i przeznaczać ją na spłatę długu.

W przeciwieństwie do gospodarstw domowych i poszczególnych jednostek cykl „życia" państwa pozostaje jednak nieskończony. Oznacza to, że można porównywać krańcowe użyteczności dochodu i konsumpcji między kolejnymi pokoleniami.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Jeśli w dodatku uwzględniony zostanie stały rozwój gospodarczy, przejawiający się systematycznym wzrostem PKB na mieszkańca, można łatwo dostrzec, że użyteczność krańcowej jednostki konsumpcji stopniowo maleje z upływem czasu, ponieważ kolejne pokolenia stają się coraz zamożniejsze. Oznacza to, że stopa procentowa, według której kolejne pokolenia dyskontują przyszłe strumienie dochodów, maleje z upływem czasu (ta stopa procentowa reprezentuje tzw. czystą preferencję czasową – *pure time preference*). Przy tych założeniach pożyczanie (w pewnym zakresie) środków od przyszłych pokoleń w celu sfinansowania bieżących potrzeb jest więc ekonomicznie uzasadnione także tym, że jednostka dodatkowej konsumpcji dla bieżącego pokolenia ma społecznie większą wartość (użyteczność) niż jednostka dodatkowej konsumpcji w przypadku następnego pokolenia.

Spostrzeżenie dotyczące tego, że opłaca się mieć nadwyżkę wydatków nad dochodami, gdy bieżące dochody są niskie, oraz nadwyżkę dochodów nad wydatkami, kiedy bieżące dochody pozostają wysokie, leży u podstaw polityki wykorzystania deficytu budżetowego w celu pobudzania gospodarki w okresach słabej koniunktury. W istocie jest to jeden z najbardziej fundamentalnych wniosków ekonomii keynesowskiej, sformułowany pod wpływem doświadczeń Wielkiego Kryzysu lat 1929–1933. Keynes uważał, że przyczyną cyklicznych spadków produkcji i zatrudnienia w gospodarce kapitalistycznej były okresowe niedobory popytu prywatnego, w związku z czym rekomendował, aby rządy nie obawiały się zwiększać wydatków finansowanych deficytem w czasie recesji. Jego pogląd, zgodnie z którym deficyt budżetowy stanowi czynnik wspomagający gospodarkę w okresach słabej koniunktury, jest współcześnie powszechnie akceptowany.

We wszystkim trzeba jednak zachować umiar. Ustawiczne deficyty kumulują się i przekładają na rosnący dług publiczny. Dlatego też polityka utrzymywania stałego deficytu prowadzi często do nadmiernego zadłużenia i problemów z wypłacalnością. W takiej sytuacji ocena ekonomicznych skutków deficytu, w tym zwłaszcza tego, czy ma on pozytywny wpływ na gospodarkę, zależy od dodatkowych czynników i okoliczności, tj. m.in. od tego, czy deficyt jest przejściowy, czy trwały, niski czy wysoki (w stosunku do PKB), czy ma miejsce w okresie dobrej, czy złej koniunktury, oraz czy występuje w warunkach niskiego czy wysokiego zadłużenia. Gdy deficyt jest zjawiskiem trwałym, pozostaje względnie wysoki w relacji do PKB, ma miejsce w okresie silnego wzrostu gospodarczego lub gdy występuje w warunkach znacznego zadłużenia, będzie miał szkodliwe konsekwencje w przypadku gospodarki. Należy do nich ryzyko wzrostu inflacji na skutek monetyzacji deficytu, wzrost kosztów obsługi długu, co ogranicza środki na rozwój i grozi

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wpadnięciem w „pułapkę zadłużenia" (*debt trap*), absorpcja oszczędności prywatnych i związany z tym efekt „wypychania" (*crowding out*) oraz wzrost rynkowych stóp procentowych i spowolnienie wzrostu gospodarczego.

Gdy suma oszczędności krajowych jest w danym momencie określona, absorpcja oszczędności prywatnych poprzez pożyczki zaciągane przez sektor publiczny ogranicza możliwości finansowania inwestycji prywatnych, w tym inwestycji produkcyjnych. Stopy procentowe muszą wzrosnąć i część projektów inwestycyjnych sektora prywatnego nie zostanie sfinansowana. Skutki tego przesunięcia w przypadku wzrostu gospodarczego zależą przede wszystkim od tego, czy sektor publiczny wydaje pożyczone pieniądze na inwestycje, czy na wydatki bieżące. Doświadczenia większości krajów pokazują, że wydatki inwestycyjne sektora publicznego stanowią średnio nie więcej niż 5–10% ogółu wydatków budżetowych, a pozostała część to wydatki bieżące. Z kolei wydatki bieżące mogą powiększyć zagregowany popyt i w ten sposób przejściowo poprawić stan koniunktury, lecz nie zwiększą one potencjalnego poziomu PKB i – same z siebie – nie przyspieszają wzrostu gospodarczego. A zatem przeważająca część oszczędności prywatnych, jakie sektor publiczny pożycza na sfinansowanie deficytu, idzie na pokrycie wydatków bieżących. W ten sposób wydatki rządowe na cele bieżące „wypychają" z godpodarki inwestycje prywatne. Można więc przyjąć, że wpływ deficytu na skalę inwestycji – a zatem i na wzrost gospodarczy – jest w zasadzie negatywny[28].

Jak przedstawiają się zaś skutki deficytu w przypadku wzrostu gospodarczego, gdy wydatki rządowe są przeznaczane w większości na inwestycje? Tu odpowiedź jest bardziej złożona i zależy głównie od tego, jaka jest efektywność inwestycji publicznych w porównaniu z inwestycjami prywatnymi. Jeśli inwestycje publiczne mają charakter produkcyjny, nie można z góry założyć, że będą one mniej efektywne niż inwestycje prywatne, choć doświadczenia wielu krajów pokazują, że tak na ogół się dzieje. Państwo nie jest ani dobrym właścicielem, ani efektywnym inwestorem, ani sprawnym menadżerem. Jeśli inwestycje publiczne dokonywane są natomiast w obszarze szeroko rozumianej infrastruktury (transportowej, energetycznej, prawno-instytucjonalnej), ich wpływ na wzrost będzie zazwyczaj pozytywny, głównie dlatego, że sektor prywatny może w ogóle nie przeprowadzać takich inwestycji ze względu na zbyt duże ryzyko i długi okres zwrotu. Wówczas publiczne inwestycje infrastrukturalne, mając charakter uzupełniający wobec inwestycji prywatnych,

[28] Choć niektóre wydatki bieżące mają silny wpływ prorozwojowy (edukacja, badania naukowe, wzmocnienie praworządności).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

mogą pobudzać inwestycje prywatne poprzez zwiększanie ich produktywności krańcowej, i w ten sposób przyczyniać się do przyspieszenia wzrostu.

Jak już wspomniano, opisane zjawisko zmniejszenia inwestycji prywatnych na skutek wzrostu wydatków publicznych nazywa się efektem wypychania (lub wypierania). Efekt ten jest najsilniejszy, gdy podaż pieniądza pozostaje stała, a wzrost popytu na pieniądz w związku z inwestycjami publicznymi prowadzi do podwyższenia rynkowych stóp procentowych. Efekt wypychania nie nastąpi, gdy bank centralny zwiększy podaż pieniądza (akomodacja pieniężna), aby uniknąć wzrostu stóp procentowych. Skutkiem takiej polityki w dłuższym okresie może być jednak wzrost inflacji, który ostatecznie także doprowadzi do spadku tempa wzrostu gospodarczego.

Należy pamiętać, że efekt wypychania powstaje generalnie w przypadku deficytu finansowanego pożyczkami. Gdy wzrost wydatków finansowany jest wzrostem podatków, efekt wypychania zasadniczo nie występuje, ponieważ spadkowi wydatków prywatnych towarzyszy odpowiedni wzrost wydatków publicznych. Pewne znaczenie może mieć tylko efekt redystrybucji dochodu – od grup społecznych charakateryzujących się większą skłonnością do oszczędzania po grupy reprezentujące mniejszą skłonność w tym zakresie. Wówczas, nawet gdy wydatki finansowane są z podatków, może wystąpić efekt wypychania na skutek spadku oszczędności krajowych i spowodowanego tym wzrostu stóp procentowych.

Deficyt budżetowy ma z reguły także ujemny wpływ na saldo obrotów bieżących. Wzrost wydatków rządowych oznacza bowiem również wzrost wydatków na import, co przy danym eksporcie przekłada się na pogorszenie rachunku obrotów bieżących (zob. ramka 3.5). Mówi się w związku z tym o tzw. deficytach bliźniaczych (*twin deficits*). Występowanie obu deficytów jest więc najczęściej symptomem pojawienia się nierównowagi makroekonomicznej w gospodarce.

Ramka 3.5. Deficyty bliźniacze: deficyt budżetowy i deficyt bieżący bilansu płatniczego

W ramach budżetu, zarówno po stronie wydatków, jak i dochodów, występują operacje zagraniczne. Ogólnie rzecz biorąc saldo tych operacji (saldo zagraniczne budżetu) jest z reguły ujemne, ponieważ wśród większości krajów rządy więcej kupują (i pożyczają) za granicą, niż otrzymują dochodów z zagranicy. Te składniki salda zagranicznego są włączane do poszczególnych części bilansu płatniczego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Pomijając dochody i rozpatrując wpływ deficytu na rachunek bieżący od strony wydatków, zależność tę opisuje podstawowa tożsamość makroekonomiczna:

$$S + CA = (G - T) + I. \tag{3.17}$$

A zatem, na mocy definicji, suma oszczędności krajowych S i zagranicznych CA (są to źródła finansowania) równa jest sumie deficytu budżetowego $(G - T)$ oraz inwestycji prywatnych I (są to przedmioty finansowania). Wówczas saldo obrotów bieżących CA pozostaje równe niedoborowi oszczędności publicznych $(G - T)$ i oszczędności prywatnych $(I - S)$:

$$CA = (G - T) + (I - S). \tag{3.18}$$

Według tej interpretacji, wzrost deficytu budżetowego wywołuje pogorszenie salda obrotów bieżących, co wymaga albo napływu kapitału, albo odpływu rezerw walutowych. Lecz konkretny kształt zależności miedzy $(G - T)$ i CA zależy od systemu kursu walutowego w danym kraju. W warunkach stałego kursu wzrost deficytu przekłada się na zwiększenie stóp procentowych, napływ kapitału i wzrost rezerw walutowych, wzrost podaży pieniądza, spadek stóp procentowych oraz pogorszenie salda bilansu bieżącego. Przeciwdziała temu poprawa salda bilansu kapitałowego dzięki napływowi funduszy – efekt netto w przypadku całego bilansu płatniczego zależy od względnej siły oddziaływania tych efektów[29].

Wzrost deficytu powoduje z jednej strony wzrost wydatków na import i pogorszenie salda bieżącego, z drugiej zaś – wzrost stóp procentowych i napływ kapitału, co przekłada się na poprawę salda rachunku kapitałowego. W warunkach płynnego kursu walutowego kurs waluty krajowej ulega zmianie (tym razem rezerwy walutowe pozostają na niezmienionym poziomie, jeśli bank centralny nie reaguje dostosowaniem polityki pieniężnej). Gdy przeważa efekt kapitałowy, waluta krajowa ulega aprecjacji, co powoduje obniżenie konkurencyjności i dalszy spadek eksportu netto, a dodatkowo pogarsza saldo rachunku bieżącego. Gdy przeważa efekt handlowy, waluta krajowa ulega deprecjacji, co z kolei wpływa na poprawę

[29] Gdy silniejszy będzie efekt kapitałowy, wystąpi napływ kapitału netto i zwiększenie bazy monetarnej; wywołany tym wzrost wydatków przełoży się na pogorszenie salda handlowego i przywrócenie równowagi w bilansie płatniczym. Gdy silniejszy stanie się zaś efekt handlowy, nastąpi odpływ kapitału netto i zmniejszenie bazy monetarnej, a wywołany tym spadek wydatków przywróci równowagę w bilansie płatniczym (zob. więcej na ten temat w rozdziale 6).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

konkurencyjności i zwiększa eksport netto. Łączny efekt w przypadku całego bilansu płatniczego będzie zależał od względnej siły oddziaływania tych efektów[30].

Jak wspomniano wcześniej, deficyt budżetowy, mimo że często wywołuje nienajlepsze skojarzenia, w pewnych sytuacjach może być pożądany. Co w takim razie z nadwyżką budżetową? Czy jest ona przejawem racjonalnej polityki, czy może świadczy o nieefektywnym wykorzystaniu posiadanych środków?

Trwała nadwyżka dochodów nad wydatkami oznacza tworzenie rezerw i niewątpliwie gwarantuje bezpieczeństwo finansowe. To, co jest zagrożeniem w przypadku deficytu – czyli narastanie zadłużenia i związane z tym ryzyko niewypłacalności – w zasadzie nie występuje w przypadku nadwyżki budżetowej. Umożliwia ona również wcześniejszą spłatę zaciągniętych pożyczek, co przyczynia się do poprawy wiarygodności kraju na rynkach finansowych. Nadwyżka stwarza wreszcie pożądany margines bezpieczeństwa na wypadek nagłego pogorszenia koniunktury – rząd może zwiększyć wydatki w celu stymulowania wzrostu bez powstawania negatywnych konsekwencji w postaci wzrostu kosztu obsługi długu.

Nadwyżka budżetowa niesie jednak także negatywne konsekwencje. Utrzymywana zbyt długo, oznacza nadmierny w stosunku do potrzeb poziom podatków i w efekcie może obniżać potencjalne tempo wzrostu PKB. W okresach gorszej koniunktury nadwyżka ma charakter procykliczny i może mieć niepożądane skutki deflacyjne. Nie istnieje jednak pełna symetria ujemnych konsekwencji nadwyżki i deficytu – z reguły negatywne skutki dużego deficytu okazują się znacznie groźniejsze niż efekty dużej nadwyżki.

W świetle tych obserwacji można uznać, że optymalną polityką w cyklu koniunkturalnym jest umiarkowany deficyt w okresach spowolnienia gospodarczego i umiarkowana nadwyżka w okresach szybkiego wzrostu. W perspektywie średniookresowej saldo strukturalne budżetu (skorygowane o efekt cyklu) powinno kształtować się tak, aby zapewniać stabilizację relacji długu do PKB na bezpiecznym poziomie, to znaczy takim, przy którym nie występują negatywne skutki uderzające w tempo wzrostu gospodarczego i które pozostaje spójne z niskim i stabilnym poziomem

[30] Gdy efekt kapitałowy przeważa, waluta ulega aprecjacji, co powoduje obniżenie konkurencyjności i spadek eksportu netto. W warunkach pełnej mobilności kapitału spadek eksportu netto będzie równy początkowym wydatkom – polityka fiskalna stanie się całkowicie nieskuteczna (całkowity *crowding-out*). Z kolei gdy przeważa efekt handlowy, waluta ulega deprecjacji, co powoduje wzrost eksportu netto i poprawia skuteczność polityki fiskalnej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

stóp procentowych. Kraje Unii Europejskiej przyjęły zasadę traktującą o tym, że w średnim okresie budżet powinien być zrównoważony, choć kraje członkowskie odnotowujące szybsze tempo wzrostu i posiadające niski poziom zadłużenia mogą realizować deficyt strukturalny średnio na poziomie nieprzekraczającym 1% PKB.

Warto zauważyć, że wymóg utrzymania równowagi budżetowej w średnim okresie jest dość rygorystyczny, ponieważ implikuje, że dług publiczny w długim okresie zmierza do zera (zob. równanie 3.21 w następnym podrozdziale). Zerowy poziom zadłużenia nie musi być jednak stanem optymalnym ze względu na omówioną wyżej malejącą krańcową użyteczność dochodu i konsumpcji oraz międzypokoleniowe różnice na poziomie zamożności.

3.7. Dług publiczny – podstawowe zależności

3.7.1. Wprowadzenie

Dług publiczny to suma zobowiązań finansowych jednostek wchodzących w skład sektora finansów publicznych, określona na dany moment. Ponieważ zobowiązania te powstają wówczas, gdy sektor publiczny zaciąga pożyczki, aby sfinansować nadwyżkę wydatków nad dochodami, dług publiczny można też postrzegać – w pewnym dopuszczalnym przybliżeniu – jako skumulowaną sumę deficytów budżetowych z poprzednich okresów. Okresowe deficyty finansowane pożyczkami kumulują się więc w postaci rosnącego długu publicznego.

Przyrost długu w danym okresie nie jest jednak dokładnie równy deficytowi budżetowemu w tym okresie. Łatwo to zauważyć, gdy zestawi się przyrost długu w poszczególnych okresach z wielkością zarejestrowanych deficytów faktycznych w tym czasie. Między otrzymanymi wielkościami występują różnice, które są spowodowane tym, że poziom długu jest „stanem" zobowiązań na daną chwilę, a deficyt stanowi „strumień" nowych zobowiązań w wybranym okresie. Dlatego też w rachunkowości budżetowej różnice te noszą nazwę efektu dostosowania „stanów i strumieni" (*stock-flow adjustment*) lub dostosowania „długu i deficytu" (*debt-deficit adjustment*). Źródłem tych różnic są przede wszystkim zmiany kursów walutowych w danym okresie. Ta część długu, która została zaciągnięta w walutach obcych, jest przeliczana na walutę krajową według kursu walutowego obowiązującego w dniu kończącym okres sprawozdawczy (np. rok). Ponieważ kursy walutowe zmieniają się w czasie, wzrost poziomu długu w danym okresie wynika

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

nie tylko z pojawienia się wówczas nowego deficytu, lecz także ze zmian krajowej wartości części zagranicznej długu, spowodowanych zmianami kursów. W szczególności osłabienie kursu waluty krajowej przekłada się na wzrost poziomu zadłużenia większy niż faktyczny deficyt budżetowy w danym roku, umocnienie kursu wywołuje zaś efekt odwrotny.

Innym źródłem rozbieżności występujących między przyrostem zadłużenia a wielkością deficytu jest sposób finansowania deficytu. Ponieważ deficyt budżetowy może być częściowo finansowany nie z pożyczek, lecz ze sprzedaży aktywów (np. przychody z prywatyzacji) lub transferów bezzwrotnych, przyrost zadłużenia będzie w danym okresie niższy niż faktyczny deficyt.

Tabela 3.6. Dług publiczny i deficyt sektora finansów publicznych w Polsce w latach 2002–2013 wg metodologii ESA 2010 (dane w mld PLN)

Rok	Dług publiczny	Przyrost długu	Deficyt sektora	Różnica (3–4)
1	2	3	4	5
2002	338,7	–	39,3	–
2003	394,1	55,4	51,4	4,0
2004	419,8	25,7	47,8	–22,1
2005	459,6	39,8	39,3	0,5
2006	501,9	42,3	38,2	4,1
2007	524,4	22,5	22,0	0,5
2008	595,4	71,0	46,4	24,6
2009	678,3	82,9	99,7	–16,8
2010	770,6	92,3	108,8	–16,5
2011	851,7	81,1	76,0	5,1
2012	878,9	27,2	60,1	–32,9
2013	926,6	47,7	66,6	–18,9
2003–2013	–	587,9	656,3	–68,4

Źródło: opracowanie własne na podstawie danych Eurostatu.

Tabela 3.6 przedstawia wielkości długu publicznego i deficytu sektora finansów publicznych w Polski w latach 2002–2013 w walucie krajowej. Ostatnia kolumna tabeli ukazuje różnice występujące między faktycznym przyrostem zadłużenia a faktycznym deficytem, wynikające z efektu *stock-flow adjustment*. Jak widać, w analizowanym okresie obserwowano zarówno dodatnie, jak i ujemne wartości różnic, w niektórych latach dość znaczne (np. lata 2004, 2008 i 2012), przy czym wyraźnie przeważały wartości ujemne, co pozwala zauważyć, że faktyczne poziomy

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

deficytu były większe niż oficjalnie zarejestrowane przyrosty długu, z powodów wyjaśnionych wyżej. W całym okresie 2003–2013 całkowity przyrost zadłużenia pozostawał niemal o 70 mld PLN niższy niż skumulowana wartość rocznych faktycznych deficytów.

Dług wzrastający w relacji do PKB zwiększa zagrożenie niewypłacalności, ponieważ jego obsługa wymaga coraz wiekszych płatności odsetkowych. Wiele krajów świata na przestrzeni ostatnich kilku dekad dopuściło do nadmiernego wzrostu zadłużenia publicznego i wpadło w pułapkę długu. Kraje te, nie będąc w stanie spłacać rosnących rat i odsetek, zmuszone były zawiesić obsługę długu i poprosić wierzycieli o restrukturyzację zadłużenia. Zawieszenie obsługi długu (*default*) stawia zarówno zadłużony kraj, jak i jego wierzycieli w bardzo niekorzystnej sytuacji. Z jednej strony dane państwo zostaje odcięte od źródeł finansowania i często musi dokonywać bardzo drastycznych cięć wydatków rządowych. Ich zmniejszenie pociąga za sobą z kolei spadek produkcji i wzrost bezrobocia. Z drugiej strony wierzyciele tracą zainwestowane pieniądze i sami mogą znaleźć się na krawędzi bankructwa. Dlatego też stan zadłużenia poszczególnych krajów i dane na ten temat stanowią zawsze jedną z najważniejszych informacji rynkowych, będących podstawą dla podejmowania decyzji przez inwestorów finansowych, którzy udzielają pożyczek rządom. Należy przy tym podkreślić, że informacja o poziomie zadłużenia jest znacznie istotniejsza niż informacja o bieżącym deficycie. Kraj może odnotowywać bowiem przejściowo nawet bardzo duży deficyt budżetowy, który nie musi jednak oznaczać wcale rosnącego ryzyka niewypłacalności, jeśli poziom zadłużenia pozostaje niski.

Jeśli rosnący dług zwiększa jednak ryzyko niewypłacalności, dlaczego rządy poszczególnych państw się zadłużają? Podstawowe ekonomiczne uzasadnienie w przypadku zadłużania się ma związek z dążeniem do optymalnego wygładzania (wyrównywania) konsumpcji w czasie, zarówno w cyklu życia jednego pokolenia, jak i pomiędzy pokoleniami. Jak już wspomniano, pożyczanie pieniędzy pozwala osobom fizycznym wygładzać konsumpcję w czasie w warunkach zmieniających się dochodów w okresie życia danej osoby. Na przykład młodzi ludzie rozpoczynający kariery zawodowe często zaciągają kredyty na zakup mieszkania lub samochodu, ponieważ nie zgromadzili jeszcze dostatecznych oszczędności. W ten sposób kredyt pozwala im podnieść poziom bieżącej konsumpcji. W miarę upływu czasu i rozwoju karier zawodowych ludzie spłacają kredyty i coraz większą część rosnącego dochodu zaczynają przeznaczać na oszczędności. Po przejściu na emeryturę bieżące dochody maleją, lecz ludzie wykorzystują zgromadzone oszczędności w celu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

utrzymania poziomu konsumpcji. Z kolei zaciąganie kredytów w przypadku przedsiębiorstw pozwala wygładzać inwestycje i produkcję w warunkach zmieniających się w cyklu koniunkturalnym bieżących przychodów ze sprzedaży.

Podobnie jest w przypadku władzy publicznej. Zaciąganie pożyczek daje rządom możliwość kształtowania wydatków publicznych stosownie do bieżących potrzeb w warunkach zmieniających się w czasie dochodów podatkowych. Amerykański ekonomista Robert Barro zauważył, że zmiany wydatków publicznych w poszczególnych okresach wymagałyby zmiany poziomu opodatkowania. Można jednak wykazać, że optymalną polityką fiskalną jest utrzymywanie stałej stopy opodatkowania dochodu, co oznacza, że okresowe deficyty spowodowane przejściowo wyższymi wydatkami należy finansować długiem, a okresowe nadwyżki powinny służyć spłacie długu. Barro opracował w ten sposób model optymalnej polityki długu publicznego, oparty na zasadzie „wygładzania podatków" (*tax-smoothing*) w ujęciu wielookresowym. Skrócony opis tego modelu przedstawiono w ramce 3.6.

Ramka 3.6. Model „wygładzania podatków" według Roberta Barro*

Punktem wyjścia dla Roberta Barro jest stwierdzenie, że wydatki rządowe G mogą być w poszczególnych okresach finansowane z podatków lub z długu, lecz w długim okresie jedyne źródło finansowania stanowią wyłącznie podatki, ponieważ wszystkie długi muszą być prędzej czy później spłacone. Jednocześnie Barro za kluczowe przyjął założenie, że nakładanie i ściąganie podatków pociąga za sobą koszty w postaci zniekształceń rynkowych i kosztów poboru oraz redystrybucji. Wówczas ograniczenie budżetowe rządu w okresie t przyjmuje następującą postać (Barro 1979):

$$G_t - rb_{t-1} = \tau_t + \left(b_t - b_{t-1}\right), \tag{3.19}$$

gdzie r jest realną stopą procentową, b realnym poziomem długu publicznego, a τ sumą pobieranych podatków. Barro przyjmuje, że stopa procentowa pozostaje niezmienna w czasie, a wielkość wydatków budżetowych G_t stanowi zmienną egzogeniczną. Rozwijając ograniczenie budżetowe (3.19) na wszystkie okresy $t = 1,2,\ldots,\infty$, otrzymujemy ograniczenie budżetowe rządu w nieskończonym horyzoncie:

$$\sum_{t=1}^{\infty}\left[\frac{G_t}{(1+r)^t}\right] + b_0 = \sum_{t=1}^{\infty}\left[\frac{\tau_t}{(1+r)^t}\right]. \tag{3.20}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Równanie pokazuje, że w nieskończonym horyzoncie czasowym suma zdyskontowanych wydatków rządowych powiększona o początkowy stan zadłużenia musi być równa sumie zdyskontowanych dochodów podatkowych. Jednocześnie równanie to nie przesądza o konkretnym rozkładzie dochodów podatkowych w czasie.

Jak stwierdziliśmy wyżej, pobór podatków kosztuje. Barro zakłada, że ogólna suma tych wydatków jest rosnącą funkcją sumy pobieranych podatków τ oraz malejącą funkcją wielkości dochodu krajowego Y. Wówczas społeczny koszt poboru podatków w okresie t można zapisać jako:

$$Z_t = F(\tau_t, Y_t) = \tau_t f\left(\frac{\tau_t}{Y_t}\right), \quad \text{przy czym } f' > 0. \tag{3.21}$$

W równaniu zakłada się, że funkcja Z_t jest jednorodna w stopniu pierwszym względem argumentów τ i Y[31]. Zdyskontowaną wartość wszystkich kosztów poboru podatków we *wszystkich* okresach można więc zapisać następująco:

$$Z = \sum_{t=1}^{\infty} \tau_t f\left(\frac{\tau_t}{Y_t}\right) / (1+r)^t. \tag{3.22}$$

Zmienne G_t i Y_t są egzogeniczne. Zadanie rządu polega na znalezieniu takich wartości $\tau_1, \tau_2, \ldots,$ aby zminimalizować zdyskontowaną wartość kosztów poboru podatków Z. Należy więc znaleźć minimum funkcji (3.22) przy ograniczeniu budżetowym (3.20). W tym celu budujemy funkcję Lagrange'a o postaci:

$$L = \sum_{t=1}^{\infty} \tau_t f\left(\frac{\tau_t}{Y_t}\right) / (1+r)^t - \lambda \left\{ \sum_{t=1}^{\infty} \left[\frac{G_t}{(1+r)^t} \right] + b_0 - \sum_{t=1}^{\infty} \left[\frac{\tau_t}{(1+r)^t} \right] \right\}. \tag{3.23}$$

Różniczkując funkcję L względem τ_t i przyrównując pochodne cząstkowe do zera, otrzymujemy[32]:

$$\frac{\delta L_t}{\delta \tau_t} = \frac{f'\left(\frac{\tau_t}{Y_t}\right)}{(1+r)^t} + \frac{\lambda}{(1+r)^t} = \frac{\delta Z_t}{\delta \tau_t} + \lambda = 0,$$

[31] Oznacza to, że jeśli argument funkcji przemnoży się przez pewną stałą, to wartość funkcji zostanie także przemnożona o tę samą stałą.

[32] Ponieważ $f'\left(\frac{\tau_t}{Y_t}\right) = \frac{\delta Z_t}{\delta \tau_t}$, to wyrażenie $\left(\frac{1}{(1+r)^t}\right)$ (po obu stronach równania możemy pominąć.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

czyli

$$\frac{\delta Z_t}{\delta \tau_t} = -\lambda, \text{ dla wszystkich } t = 1,2, \ldots, \infty. \tag{3.24}$$

Równanie (3.24) pokazuje, że warunkiem minimum funkcji (3.22) jest wyrównanie się krańcowych kosztów poboru podatków w przypadku wszystkich okresów. Lecz stałość krańcowych kosztów poboru oznacza, w nawiązaniu do równania (3.21) i warunku jednorodności funkcji *Z*, że również relacja $\frac{\tau_t}{Y_t}$ jest stała w czasie. Optymalna polityka fiskalna polega więc utrzymywaniu stałej stopy opodatkowania, niezależnie od zmieniających się wartości wydatków *G*.

Zaproponowany przez Barro model „wygładzania podatków" wyjaśnia, dlaczego – zamiast zmieniać poziom obciążeń podatkowych – rządowi opłaca się finansować zmieniające się wydatki z długu, lecz nie uwzględnia jeszcze jednego ważnego argumentu na rzecz zadłużania się. Chodzi o zasadnicze zjawisko długofalowego wzrostu zamożności społecznej w warunkach malejącej krańcowej użyteczności dochodu. Jak już zauważono, dług publiczny pozwala optymalnie wygładzać konsumpcję społeczną nie tylko w horyzoncie życia jednego pokolenia, lecz także między pokoleniami. Zakładając, że w długim okresie wzrasta dochód wytworzony na mieszkańca, można przyjąć, że przyszłe pokolenia będą zamożniejsze niż obecne. Oprócz tego, że dochód na mieszkańca będzie wyższy, zasobniejszy okaże się też zgromadzony majątek na mieszkańca. Wynika to z postępu technologicznego oraz przyrostu kapitału ludzkiego i intelektualnego. Oznacza to także, że krańcowa jednostka dochodu ma większą wartość dla obecnego pokolenia niż dla przyszłych pokoleń. A zatem, uwzględniając malejącą krańcową wartość dochodu, transfer środków od przyszłych do obecnych pokoleń podnosi poziom międzypokoleniowego dobrobytu całego społeczeństwa.

Wskazuje się także, że umiarkowany dług publiczny dostarcza podmiotom rynkowym bezpiecznych aktywów, co pogłębia i zwiększa płynność rynków finansowych. To z kolei łagodzi ograniczenia w dostępie sektora prywatnego do kredytu i w ten sposób ułatwia realizację prywatnych inwestycji. Bez możliwości zadłużania się współczesne gospodarki musiałyby rozwijać się wolniej i w warunkach większej niestabilności makroekonomicznej. Wielkim entuzjastą długu publicznego jest między innymi amerykański noblista Paul Krugman (zob. ramka 3.7).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 3.7. Zalety długu publicznego według Paula Krugmana

W krótkim felietonie opublikowanym w 2015 roku w „The International New York Times" (22–23 sierpnia 2015 roku) Paul Krugman przekonuje, że „dług jest dobry" (*debt is good*), krytykuje politykę oddłużania, jaką stosowały kraje rozwinięte w okresie kryzysu finansowego i gospodarczego w latach 2009–2014, i wręcz nawołuje do zwiększania zadłużenia. Jakie są jego argumenty?

Po pierwsze, z długu można finansować wiele pożytecznych inwestycji publicznych, takich jak budowa dróg, kolei, portów czy lotnisk. Jest to szczególnie uzasadnione wtedy, gdy stopy procentowe są bliskie zeru – wówczas zaciąganie pożyczek niemal nic nie kosztuje. Po drugie, dług publiczny sprzyja lepszemu funkcjonowaniu gospodarki, ponieważ dostarcza inwestorom bezpiecznych aktywów w postaci obligacji rządowych. Pozwalają one lepiej zarządzać ryzykiem i zapewniają wysoką płynność, co jest szczególnie ważne w okresach stresu finansowego. Wprawdzie sektor prywatny także dostarcza aktywa w postaci instrumentów dłużnych i udziałowych, lecz pozostają one zawsze obciążone większym ryzykiem niż obligacje rządowe. Szczególnie bolesnym doświadczeniem, według Krugmana, było załamanie się rynku papierów wartościowych emitowanych pod zastaw niepłynnych aktywów, takich jak ABS (*asset-backed securities*) i MBS (*mortgage-backed securities*). Po trzecie, w warunkach superniskich stóp procentowych inwestorzy staną się być może bardziej skłonni do inwestowania w ryzykowne aktywa, co może zaskutkować kryzysem finansowym. Aby ograniczyć czynniki skłaniające do podejmowania tego ryzyka, rząd powinien jeszcze bardziej się zadłużyć, ponieważ dzięki temu ceny obligacji będą spadać, a rentowność wzrośnie.

Wywód Krugmana jest oczywiście niepełny i tendencyjny. Niepełny, ponieważ nie uwzględnia sytuacji krajów zadłużonych, którym rynki finansowe albo w ogóle odmawiają dalszych pożyczek, albo domagają się bardzo wysokich stóp procentowych. Owszem, niektóre państwa – takie jak Stany Zjednoczone czy Wielka Brytania – mimo dużego zadłużenia, nadal są w stanie na korzystnych warunkach zadłużać się na rynkach kapitałowych, ponieważ ciągle cieszą się zaufaniem inwestorów. Lecz w przypadku wielu innych krajów dalsze zaciąganie pożyczek albo jest niemożliwe, albo wymaga wzrostu stóp procentowych do poziomu, który sprawia, że cała operacja staje się nieopłacalna. Wywód ten jest również tendencyjny, ponieważ eksponuje jedynie korzyści płynące z zaciąganych pożyczek, a zupełnie pomija koszty obsługi długu. Tymczasem wysokie koszty obsługi długu należą do kategorii

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wydatków sztywnych i ograniczają swobodę w kształtowaniu polityki wydatków publicznych. W skrajnym przypadku wysokie koszty obsługi mogą doprowadzić do bankructwa suwerennego. Przykładem są tu kraje takie jak Grecja czy Argentyna, których obligacje trudno obecnie traktować jako „bezpieczne aktywa".

Zadłużanie się nie działa jednak jak czarodziejska różdżka i wiąże się z określonym ryzykiem. W miarę jak wzrasta dług, zdolność dłużnika do terminowej jego obsługi staje się w coraz większym stopniu wrażliwa na negatywne szoki dochodowe i wzrost stóp procentowych. Dla danej skali szoku im wyższe okaże się zadłużenie wybranego kraju, tym większe stanie się ryzyko jego niewypłacalności. W przypadku bardzo zadłużonych krajów nawet stosunkowo łagodny szok dochodowy może spowodować całkowitą utratę zaufania inwestorów i wstrzymanie napływu finansowania. Z kolei brak finansowania zmusza do dokonywania drastycznych oszczędności, które przekładają się na spadek produkcji, konsumpcji i inwestycji. Im wyższy będzie poziom zadłużenia, tym głębsza będzie recesja wywołana wstrzymaniem finansowania. Jednocześnie im głębsza będzie recesja, tym bardziej wzrośnie ryzyko tego, że zadłużony kraj może nie być w stanie obsługiwać długu na bieżąco. Oznacza to, że wysoki poziom zadłużenia w relacji do PKB zwiększa zmienność produkcji i obniża średnie tempo wzrostu gospodarczego (Reinhard i Rogoff 2010). Ponadto, zagraża stabilności sektora finansowego, który przecież utrzymuje część swoich aktywów w rządowych papierach wartościowych. Kłopoty z wypłacalnością rządu mogą więc pogrążyć także krajowy system bankowy. Kryzys zadłużenia w Grecji w latach 2010–2015 przebiegał właśnie według takiego scenariusza.

Ogólny wniosek płynący z dotychczasowych rozważań jest więc taki, że zadłużenie może być korzystne dla wzrostu i stabilności gospodarczej, o ile nie przekracza pewnego bezpiecznego poziomu. Bardzo niski poziom zadłużenia lub wręcz jego brak oznacza rezygnację z wykorzystania istniejących możliwości przyspieszenia rozwoju za pomocą oszczędności zagranicznych. Lecz jednocześnie wysoki poziom zadłużenia zwiększa ryzyko niewypłacalności i w efekcie hamuje rozwój gospodarczy. Ekonomiści szacowali „bezpieczny" poziom zadłużenia – okazuje się, że zależy on między innymi od stopnia rozwoju danego państwa. Im bardziej rozwinięty kraj, tym wyższy może być taki poziom zadłużenia, który nie wywołuje negatywnych konsekwencji w przypadku wzrostu ekonomicznego (Cecchetti i in. 2011). Wrócimy do tego zagadnienia poniżej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Mówiąc o występowaniu ogólnej zależności pomiędzy rosnącym poziomem długu a rosnącym ryzykiem niewypłacalności, musimy dokonać kilku ważnych uściśleń. Po pierwsze, sam wzrost zadłużenia w wielkościach absolutnych nie musi jeszcze stanowić powodu do niepokoju, ponieważ wraz z postępującym rozwojem gospodarczym wszystkie wielkości nominalne rosną. Dlatego też, gdy mówimy o poziomie zadłużenia, mamy na myśli z reguły relatywną wielkość zadłużenia najczęściej w postaci relacji długu do PKB, ponieważ to wielkość PKB określa wielkość bieżącego dochodu i bieżące możliwości płatnicze kraju, w tym zdolność dokonywania terminowej obsługi długu. Ponieważ wielkość nominalnego PKB zazwyczaj rośnie z okresu na okres (wzrost ten jest sumą realnego zwiększenia produkcji oraz wzrostu cen) tak długo, jak nominalna wielkość zadłużenia rośnie wolniej lub co najwyżej w takim samym tempie co nominalny poziom PKB, relacja długu do PKB spada lub nie wzrasta. Wówczas zdolność kraju do obsługi zadłużenia nie ulega pogorszeniu. Innym sposobem oceny poziomu zadłużenia w relacji do możliwości jego obsługi jest porównanie go z wartością aktywów, w tym zwłaszcza aktywów płynnych, znajdujących się w posiadaniu sektora publicznego. Sprzedaż aktywów pozwala zdobyć środki na spłatę długu. Dlatego też często dokonujemy rozróżnienia między zadłużeniem brutto i zadłużeniem netto. Wielkość całkowitego zadłużenia wyrażonego w wartości bezwzględnej stanowi zadłużenie brutto, po wyłączeniu aktywów publicznych otrzymuje się natomiast zadłużenie netto. Oczywiście, zadłużenie netto jest z reguły o wiele niższe niż zadłużenie brutto, a czasami zadłużenie netto może być nawet ujemne, co oznacza, że kraj jest pożyczkodawcą netto[33, 34].

Po drugie, ten sam poziom zadłużenia w relacji do PKB będzie miał odmienne konsekwencje ekonomiczne w przypadku różnych państw. W krajach rozwiniętych relacja ta może pozostawać wysoka, nawet na poziomie 80–100% PKB, bez wywoływania negatywnych skutków gospodarczych i wzrastającego ryzyka niewypłacalności. Na przykład w wielu państwach członkowskich Unii Europejskiej zadłużenie publiczne kształtuje się na takim właśnie poziomie, mimo że przyjęte w ramach UE reguły fiskalne zakładają, że zadłużenie publiczne nie powinno

[33] W niektórych krajach dług netto jest niewiele niższy niż dług brutto. Na przykład na koniec 2009 roku w Belgii kategorie te wynosiły odpowiednio 80,7% PKB i 101,0% PKB, a we Włoszech odpowiednio 101,0% PKB i 128,8% PKB. Lecz w innych krajach różnica jest bardzo duża – np. w Norwegii dług brutto w 2009 roku wynosił 49,2% PKB, podczas gdy dług netto był ujemny i wynosił –153,4% PKB. Ujemny dług netto odnotowano również w Finlandii, Szwecji i Luksemburgu (dane na podstawie „OECD Economic Outlook", no. 87, April 2010).

[34] Oczywiście należy pamiętać, że nie wszystkie aktywa publiczne mają charakter płynny i mogą być sprzedane.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przekraczać 60% PKB. Z kolei w krajach rozwijających się zadłużenie na poziomie 30–40% PKB może już powodować trudności z obsługą długu i hamować rozwój gospodarczy. Kluczowe znaczenie mają tu różnice dotyczące kosztów obsługi długu, wynikające z odmiennego ukształtowania poziomów wiarygodności kredytowej poszczególnych państw. W przypadku danego poziomu relacji długu do PKB zdolność krajów rozwiniętych do obsługi zadłużenia jest znacznie większa niż krajów rozwijających się, ponieważ kraje rozwinięte są zamożniejsze i posiadają zgromadzony znacznie większy majątek – także w formie aktywów finansowych – który w dużej mierze równoważy posiadanie zobowiązania. Ponadto kraje rozwinięte cieszą się z reguły większą stabilnością makroekonomiczną i instytucjonalną, a polityka gospodarcza w tych jest tu bardziej przewidywalna. Czynniki te prowadzą do spadku premii za ryzyko na rynkach finansowych tych państw, co zaś sprawia, że mogą one zaciągać pożyczki po niższych stopach procentowych. Z kolei niższe stopy procentowe zapewniają niskie koszty obsługi długu w relacji do PKB, a to umożliwia utrzymywanie wyższego zadłużenia.

Po trzecie, wiele zależy od tego, na co przeznaczone są pożyczki zaciągane przez rząd. Gdy za pomocą dodatkowych zasobów finansuje się rozwój infrastruktury transportowej lub badania geologiczne mające na celu odkrycie nowych zasobów surowców i paliw, działania te mogą przyczyniać się poprawy warunków funkcjonowania sektora prywatnego. Zwiększają one krańcową produktywność kapitału prywatnego i w ten sposób mogą przyspieszać tempo wzrostu oraz zwiększać potencjalny poziom PKB. Z kolei szybszy wzrost PKB zmniejsza relatywny ciężar obsługi zadłużenia, ponieważ obniża relację długu do PKB. Inwestycje takie przekładają się też na wzrost wielkości posiadanych przez sektor publiczny aktywów (autostrady, porty, koncesje na wydobycie surowców), które można w przyszłości odpłatnie udostępniać lub prywatyzować, co z kolei może także ułatwić obsługę długu. Jeśli zaciągnięte pożyczki przeznaczane są natomiast na finansowanie wydatków bieżących (np. transfery społeczne lub wydatki związane z utrzymaniem administracji), zwiększają one zadłużenie, lecz nie przyczyniają się do zwiększenia potencjalnego poziomu PKB, co oznacza, że nie zwiększają również zdolności do obsługi długu.

Niekiedy dokonuje się rozróżnienia między długiem jawnym a długiem ukrytym. Ten drugi stanowi takie zobowiązanie, które formalnie nie występuje w postaci papieru wartościowego lub umowy oficjalnie zaciągniętej pożyczki, lecz ma charakter zobowiązania potencjalnego, które w pewnym momencie stanie się zobowiązaniem wymagalnym. Przykładem takiego długu są prawa do emerytury, które nabywają

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

obywatele płacący składki na ubezpieczenia społeczne do państwowego (publicznego) zakładu ubezpieczeń społecznych. W tradycyjnym systemie emerytalnym środki uzyskane ze składek opłacanych przez obecnie pracujące osoby są na bieżąco wypłacane w formie świadczeń emerytalnych dla aktualnych świadczeniobiorców. W takim systemie, nazywanym systemem repartycyjnym, osoby płacące składki emerytalne nie gromadzą żadnego rzeczywistego kapitału, lecz stopniowo nabywają prawo do emerytury. Prawo to, gwarantowane przez państwo, pozostaje właśnie zobowiązaniem ukrytym, ponieważ jest w istocie przyszłym zobowiązaniem finansowym rządu wobec tych osób[35]. Publikowane obecnie oficjalne statystyki dotyczące zadłużenia publicznego obejmują wprawdzie wyłącznie zobowiązania jawne, lecz coraz częściej w analizach ekonomicznych w celu oceny stabilności i perspektyw finansów publicznych uwzględnia się także zobowiązania ukryte.

W systemie finansów publicznych występują także inne powszechne zobowiązania niewymagalne, stanowiące potencjalne źródło długu, takie jak gwarancje i poręczenia rządowe czy nieprawomocne orzeczenia sądowe dotyczące określonych płatności. Są to tzw. pozycje pozabilansowe (*off-balance liabilities*), które w momencie zaistnienia wydarzenia objętego gwarancją lub poręczeniem bądź w chwili uprawomocnienia wyroku sądowego, przekształcają się w zobowiązania wymagalne, zwiększające dług publiczny (więcej na ten temat zob. Owsiak 2006, s. 331–335).

3.7.2. Płynność a wypłacalność

Zaciąganie pożyczek przez rząd powoduje wzrost zadłużenia państwa. Czy można określić, jakie są bezpieczne granice długu publicznego z punktu widzenia zdolności państwa do jego obsługi? Oraz co dokładnie oznacza niewypłacalność w przypadku suwerennego podmiotu, jakim jest państwo? Aby odpowiedzieć na te pytania, należy najpierw dokonać rozróżnienia między płynnością a wypłacalnością dłużnika oraz między zadłużeniem prywatnych podmiotów rynkowych a zadłużeniem publicznym. Kwestie te zostaną kolejno omówione.

Dłużnik zachowuje *płynność*, gdy posiada płynne środki finansowe do pokrycia zapadających zobowiązań. Brak płynności nie oznacza jeszcze niewypłacalności,

[35] Trzeba jednocześnie zaznaczyć, że źródłem pokrycia tych przyszłych zobowiązań są przyszłe dochody z tytułu składek ubezpieczeniowych, które będą wpłacali do systemu kolejni jego uczestnicy. Wysokość przyszłych zobowiązań emerytalnych stanowi zadłużenie ukryte brutto. Nie można więc wyłącznie na ich podstawie formułować wniosków co do poziomu zadłużenia, jeśli nie bierze się jednocześnie pod uwagę tych „ukrytych należności" w postaci przyszłych strumieni podatków i składek na ubezpieczenia społeczne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ponieważ dłużnik może wprawdzie w danym momencie nie posiadać płynnych środków pieniężnych, lecz zarazem być w posiadaniu innych aktywów, których sprzedaż zapewni mu niezbędną płynność. W takiej sytuacji dłużnik na ogół jest w stanie otrzymać krótkoterminową pożyczkę, najczęściej pod zastaw części posiadanych pozostałych aktywów, aby spłacić najpierw bieżące zobowiązania, a następnie zaciągniętą pożyczkę, sprzedając posiadane aktywa w określonym czasie. Gdy dłużnik boryka się zatem jedynie z kryzysem płynności, wierzyciele (lub inne instytucje finansowe) pozostają skłonni udzielić dłużnikowi tymczasowej pomocy i dać mu czas na zdobycie płynnych środków. Można więc stwierdzić, że kryzys płynności to poważny kłopot, lecz jeszcze nie katastrofa.

Dłużnik staje się *niewypłacalny*, gdy przestaje obsługiwać zadłużenie. Najczęściej dzieje się tak, gdy bieżąca wartość posiadanych aktywów spada poniżej wartości posiadanych zobowiązań, a dłużnik traci zarazem zdolność zaciągania dalszych pożyczek. Uważa się, że dłużnik przestaje obsługiwać zadłużenie i staje się niewypłacalny, gdy jednostronnie narusza warunki zawartej umowy pożyczki lub warunki związane z emisją dłużnych papierów wartościowych. Może to przyjąć formę opóźnienia w płatnościach, dokonania spłat w kwocie niższej niż ustalona, wstrzymania lub obniżenia płatności odsetkowych bądź – w przypadku większej liczby wierzycieli – wstrzymania zapadających płatności wobec niektórych z nich. W takiej sytuacji wierzyciele mają prawo zwrócić się do sądu i rozpocząć procedurę upadłościową, która ma zmierzać do odzyskania od dłużnika całości lub części swoich należności. W przypadku długu prywatnego polega to najczęściej na przejęciu całego majątku dłużnika przez specjalną instytucję (komornik, syndyk) w celu zabezpieczenia wierzytelności. Kryzys niewypłacalności oznacza więc z reguły bankructwo, czyli upadek dłużnika.

Istnieje oczywiście związek pomiędzy kryzysem płynności i kryzysem wypłacalności. Kryzys płynności nie oznacza jeszcze niewypłacalności, ale najczęściej sygnalizuje trudności w obsłudze długu. Gdy wierzyciele uznają, że trudności te maja charakter przejściowy – na przykład są spowodowane opóźnieniem w spływie należności lub czasem potrzebnym na zbycie aktywów – będą zapewne skłonni zgodzić się na odłożenie w czasie należnych im płatności. Często trudno jest jednak ocenić, czy i na ile bieżące trudności w obsłudze długu mają charakter przejściowy, a na ile są spowodowane fundamentalnym niedostatkiem zasobów. W takiej sytuacji wierzyciele mogą nie wyrazić zgody na odłożenie spłaty w czasie i wszczynają procedurę upadłościową. Wówczas, mimo iż dłużnik może być nadal fundamentalnie wypłacalny, kryzys płynności przeistacza się w kryzys

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

niewypłacalności. Tego rodzaju sytuacja może wystąpić zwłaszcza na rynku długu suwerennego, gdzie panika rynkowa (*market panic*) i tak zwane „stadne zachowania" (*herd behaviour*) inwestorów mogą szybko doprowadzić dłużnika mającego kłopoty z płynnością do pełnego bankructwa.

3.7.3. Związek między deficytem a długiem

Związek między deficytem i tempem wzrostu zadłużenia zależy oczywiście nie tylko od wielkości sukcesywnych deficytów, ale także od dwóch innych czynników – tempa wzrostu PKB oraz stóp procentowych. Ponieważ wartość każdego zobowiązania rejestrowana jest w wartościach nominalnych, to w miarę wzrostu wartości nominalnego PKB relacja danej wielkości długu do PKB spada z okresu na okres. Zjawisko to, które czasami nosi nazwę „wyrastania z długu" (*growing out of debt*), jest korzystne w przypadku zadłużonego rządu, ponieważ pozwala zachować kontrolę nad poziomem zadłużenia nawet w warunkach występowania niewielkiego deficytu. W odwrotnym kierunku oddziałują zmiany stóp procentowych. Ponieważ ich wysokość decyduje o kosztach bieżącej obsługi długu, wzrost stóp procentowych, po jakich inwestorzy skłonni są udzielać pożyczek rządowi, zwiększa te koszty przy danym poziomie długu, co przekłada się na wzrost ryzyka niewypłacalności. Związek między deficytem a długiem w ujęciu dynamicznym wygodnie jest przedstawić w postaci tzw. arytmetyki deficyt – dług (ramka 3.8).

Ramka 3.8. Arytmetyka deficyt–dług

Przyjmijmy, że $d = D/Y$ oznacza stosunek długu do PKB, $b = B/Y$ wyraża stosunek pierwotnego deficytu budżetowego do PKB, i oznacza stopę procentową, po której rząd zaciąga pożyczki na rynku, a $y = \Delta Y/Y$ stanowi stopę wzrostu nominalnego PKB (stopa wzrostu realnego powiększona o stopę inflacji). Wówczas zmianę stosunku długu do PKB z roku na rok można zapisać następująco[36]:

[36] Zmiana poziomu zadłużenia z roku na rok wynosi $dD = D - D_{-1} = B + i \cdot D_{-1}$. Dzieląc obie strony przez Y oraz oznaczając, że $Y = Y_{-1}(1 + y)$, otrzymujemy:

$$d - d_{-1} = b + d_{-1}\{[(1 + i)/(1 + y)] - 1\},$$

a po dalszych przekształceniach:

$$d - d_{-1} = b + d_{-1}[(i - y)/(1 + y)].$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$d_0 - d_{-1} = b + d_{-1}\left[\frac{i - y}{1 + y}\right], \tag{3.25}$$

gdzie indeks (*–1*) oznacza wartość zmiennej z poprzedniego okresu. Równanie (3.25) pokazuje, że zmiana stosunku długu do PKB jest sumą deficytu pierwotnego *b* (w relacji do PKB) oraz wyrażenia $d_{-1}\left[\frac{i - y}{1 + y}\right]$, które nosi nazwę efektu „kuli śniegowej" (*snow-ball effect*). Efekt ten wyraża łączny wpływ stopy procentowej i tempa wzrostu PKB na relację długu do PKB[37]. Jest to podstawowe równanie dynamiki długu (zob. np. Burda i Wyplosz 2003).

Utrzymanie stałej relacji długu do PKB wymaga, aby $d_0 = d_{-1}$, a więc lewa strona (3.25) musi być równa zeru. Wówczas warunkiem ustabilizowania relacji długu do PKB na określonym poziomie *d* jest, aby:

$$b = d_{-1}\left[\frac{y - i}{1 + y}\right] \tag{3.26}$$

lub – w wartościach absolutnych:

$$B = D_{-1}\left(y - i\right). \tag{3.26a}$$

Równanie (3.26) pokazuje, że warunek stabilności relacji długu do PKB wymaga, aby relacja deficytu pierwotnego do PKB była równa efektowi „kuli śniegowej". Znak *b* zależy od znaku wyrażenia ($y - i$), ponieważ pozostałe elementy prawej strony równania (3.26) są zawsze dodatnie. Gdy $(y - i) > 0$, również $b > 0$, co oznacza, że jeśli tempo wzrostu PKB przewyższa stopę procentową, warunek stabilności relacji długu do PKB wymaga, aby dany kraj wykazywał deficyt pierwotny w wysokości *b*, przy czym może on być tym wyższy, im większa jest różnica ($y - i$). Również odwrotnie, jeśli $(y - i) < 0$, warunek stabilności relacji długu do PKB wymaga, aby

[37] W równaniu (3.25) zakładamy, że wpływ czynników związanych ze zmianami wyceny długu – tzw. efekt *stock-flow adjustment* – jest równy zero. Jak już wcześniej wyjaśniono, efekt *stock-flow adjustment* (*debt-deficit adjustment*) wynika z różnic w wycenie zobowiązań i innych korekt statystycznych oraz zapewnia spójność pomiędzy strumieniem nowych zobowiązań w danym czasie a stanem zadłużenia na koniec tego okresu (European Commission, *General Government Data*, Part II: tables by series, Autumn 2012, s. 11–12).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kraj realizował nadwyżkę pierwotną ($-b$), przy czym jest ona tym większa, im wyższa staje się absolutna wielkość różnicy ($y - i$).

Równanie (3.26) można przekształcić tak, aby uzyskać długookresową relację między deficytem, długiem i nominalnym tempem wzrostu PKB. Przyjmując, że $d = d_{-1}$, oraz zakładając że odsetki od długu płacone są bez opóźnień (na bieżąco), a także uwzględniając, że $y/(1 + y) \approx y$, otrzymujemy znaną relację:

$$d = b/y. \tag{3.27}$$

Formuła (3.27) wskazuje, że w długim okresie relacja długu do PKB zmierza do stałej wartości równej stosunkowi udziału deficytu pierwotnego w PKB i nominalnego tempa wzrostu PKB. Na przykład, gdy deficyt pierwotny utrzymywany jest na poziomie 3% PKB, a nominalne tempo wzrostu PKB wynosi 5% (będąc sumą potencjalnego realnego tempa wzrostu w wysokości, na przykład, 3%, i średniej inflacji w wyskości, na przykład, 2%), wówczas relacja długu do PKB zmierza w długim okresie do poziomu 3/5, tj. 60% PKB[38].

W podstawowym równaniu dynamiki długu (3.26) warunkiem stabilizacji relacji długu do PKB na określonym poziomie jest założenie, aby relacja deficytu do PKB była równa tzw. efektowi „kuli śniegowej". Nie jest on jednak stały w czasie. Kluczowym elementem decydującym o jego dynamice jest wyrażenie ($y - i$). W równaniach (3.26) i (3.26a) wyrażenie ($y - i$) jest na ogół traktowane jako stały parametr, co oznacza, że zakłada się istnienie liniowej zależności pomiędzy deficytem i długiem. Zarówno bardziej pogłębiona analiza obu równań, jak i empiryczne obserwacje zachowania się tych zmiennych (zwłaszcza w okresie kryzysu finansowego z lat 2008–2011) sugerują jednak, że rzeczywista zależność może być raczej nieliniowa, przynajmniej powyżej pewnego poziomu zadłużenia d. Wynika to stąd, że wartości parametrów y oraz i nie są stałe oraz niezależne od wartości zmiennych d i b natomiast zmieniają się wraz ze zmianami poziomu d. W szczególności można wykazać, że w przypadku wysokich wartości relacji długu do PKB, wartość y spada, podczas gdy wartość i rośnie, w rezultacie czego wyrażenie ($y - i$)

[38] Podany przykład ilustruje, w jaki sposób państwa członkowskie UE określiły w Traktacie z Maastricht w 1992 roku maksymalne limity deficytu oraz długu publicznego w ramach funkcjonowania Unii Gospodarczej i Walutowej w Europie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

staje ujemne i coraz większe w wartości bezwzględnej. Z kolei spadek wartości ($y - i$) poniżej zera wymaga, aby pierwotne saldo budżetu było dodatnie i wzrastało w relacji do PKB. Przy tym założeniu równanie (3.25) dotyczące dynamiki długu powinno zostać uzupełnione równaniem opisującym zależność między poziomem relacji długu do PKB a wielkością wyrażenia ($y - i$):

$$(y - i) = f(d_{-1}), \tag{3.28}$$

przy czym $f' < 0$, i $f'' > 0$.

Ogólny związek pomiędzy relacją długu do PKB (d) a relacją deficytu do PKB (b) w zależności od przyjęcia różnych wartości ($y - i$) został przedstawiony na rysunku 3.3 (zakładamy dla uproszczenia, że $d = d_{-1}$). Relacja deficytu pierwotnego do PKB, wymagana, aby zapewnić stabilność relacji długu do PKB, jest ukazana jako funkcja relacji długu do PKB odnośnie do różnych wartości ($y - i$). Ogólnie, w przypadku niskich wartości relacji długu do PKB różnica ($y - i$) stanie się dodatnia i dany kraj będzie mógł wykazywać pierwotny deficyt budżetowy tak, jak opisuje to równanie (3.26). Lecz wobec wysokich wartości relacji długu do PKB wyrażenie ($y - i$) staje się ujemne i dany kraj musi wykazywać coraz większą nadwyżkę budżetową, aby zapewnić stabilną relację długu do PKB (więcej na temat tej prawidłowości zob. Ostry i in. 2010). Jak pokazuje równanie (3.26), wymagana nadwyżka stanie się większa nie tylko dlatego, że musi być spójna z wyższą wartością d_{-1}, ale także dlatego, że wraz ze wzrostem d rośnie wartość bezwzględna ($y - i$).

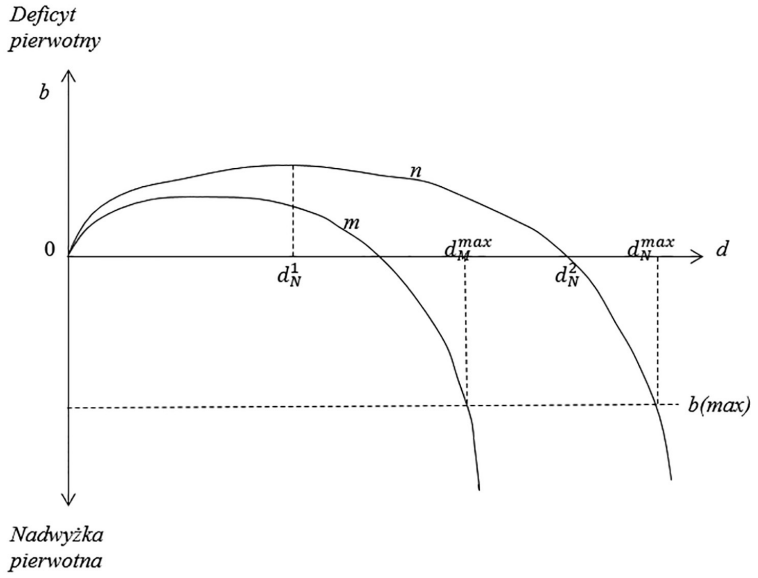
Na rysunku 3.3, linie m i n ukazują zależność d i b odnośnie do dwóch różnych hipotetycznych przypadków państw M i N, w których ścieżki zmian wyrażenia ($y - i$) mają różny przebieg. Kraj M jest mało wiarygodny i wartość ($y - i$) dla tego kraju – po początkowym wzroście – maleje szybko wraz ze wzrostem zadłużenia. Z kolei kraj N pozostaje wiarygodny i wartość ($y - i$) tego kraju najpierw długo rośnie, by następnie znacznie wolniej spadać. Oczywiście, w obu przypadkach po przekroczeniu pewnego granicznego poziomu zadłużenia wartość ($y - i$) maleje wraz ze wzrostem d. Oznacza to, że – dla danego ($y - i$) – pierwotna nadwyżka budżetu musi być tym wyższa, im wyższy jest poziom zadłużenia, tak aby zapewnić stabilność relacji długu do PKB.

Nadwyżka pierwotna nie może jednak rosnąć w nieskończoność. Oznacza ona bowiem konieczność ponoszenia coraz większych wyrzeczeń, na co rząd suwerennego kraju – politycznie odpowiedzialny przed własnym społeczeństwem – może w pewnym momencie nie wyrazić zgody. Znany francuski ekonomista Olivier

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Blanchard wprowadził pojęcie „maksymalnie politycznie dopuszczalnej nadwyżki pierwotnej", czyli pewnej granicznej relacji nadwyżki pierwotnej do PKB, powyżej której dany kraj zdecyduje się zaprzestać obsługi długu ze względu na brak politycznej akceptacji dla dalszych niezbędnych wyrzeczeń (Blanchard 1984). Linia *b*(*max*) oznacza właśnie taką maksymalną politycznie dopuszczalną nadwyżkę budżetową. W przypadku kraju M ta granica występuje przy niższym poziomie zadłużenia *d*, ponieważ wówczas wartość wyrażenia (*y* – *i*) szybko spada w miarę wzrostu zadłużenia. W przypadku kraju N granica ta występuje przy wyższym poziomie zadłużenia *d*, ponieważ wtedy wartość wyrażenia (*y* – *i*) maleje wolniej w miarę wzrostu zadłużenia[39]. Tak więc poziom zadłużenia powodujący niewypłacalność kraju N w relacji do PKB jest wyższy niż w kraju M.

Rysunek 3.3. Zależność między relacją długu do PKB (*d*) a relacją deficytu pierwotnego do PKB (*b*) spójną z warunkiem stabilizacji długu



Źródło: Rosati 2013, s. 24.

[39] Na rysunku 3.3 przyjeto, że wartość *b* (*max*) jest taka sama dla obu krajów M i N. W rzeczywistości jednak różne kraje mogą różnic się pod względem wielkości maksymalnej politycznie akceptowanej nadwyżki pierwotnej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.7.4. Mechanizm rosnącego ryzyka suwerennej niewypłacalności

Dlaczego w przypadku wyższych poziomów zadłużenia wartość wyrażenia ($y - i$) staje się jednak coraz niższa? Dzieje się tak dlatego, że wraz ze wzrostem relacji długu do PKB maleje tempo wzrostu gospodarczego y oraz rośnie rynkowa stopa procentowa i.

Stopa wzrostu gospodarczego maleje, ponieważ im wyższy staje się poziom długu w relacji do PKB, tym wyższy będzie całkowity koszt obsługi zadłużenia w relacji do PKB. To z kolei wymaga nakładania coraz wyższych podatków, co pociąga za sobą większe straty społeczne z tytułu zniekształceń rynkowych (*market distortions*) i ogranicza zasób kapitału, przeznaczanego na inwestycje[40]. W rezultacie tempo wzrostu spada. Ponadto kraje o wysokim zadłużeniu podejmują z reguły próby odzyskania równowagi w zakresie finansów publicznych, dokonując konsolidacji fiskalnej, co oznacza, że, dążąc do ograniczenia deficytu budżetowego, podnoszą podatki i ograniczają wydatki. Makroekonomiczny efekt tych działań polega na tym, że zagregowany popyt w gospodarce spada, a to przekłada się zaś na spadek tempa wzrostu gospodarczego.

Oba wspomniane efekty – wzrost podatków służący sfinansowaniu kosztów obsługi długu oraz spadek zagregowanego popytu na skutek konsolidacji fiskalnej – są uwzględniane w standardowej analizie dynamiki długu i deficytu. Mimo to pomija ona jeden ważny mechanizm, który nabrał znaczenia w okresie kryzysu zadłużenia w latach 2008–2012. Chodzi mianowicie o oddziaływanie ryzyka suwerennej niewypłacalności (*sovereign risk default*) na wzrost gospodarczy. Mechanizm ten oddziałuje na stopę wzrostu PKB za pośrednictwem dwóch kanałów (Corsetti i Mueller 2011). Po pierwsze, w miarę wzrostu relacji długu do PKB rośnie także – *ceteris paribus* – ryzyko niewypłacalności kraju, a to przekłada się na wzrost rentowności obligacji i wzrost kosztu zabezpieczenia się przez ryzykiem niewypłacalności (na przykład w postaci rosnącego kosztu kontraktów zabezpieczających CDS – *credit default swaps*)[41]. Wpływa to ujemnie na warunki zaciągania kredytów

[40] Efekt ten stanie się tym silniejszy, im silniejszy w skutkach okaże się wpływ równoważności rikardiańskiej, tzn. w im większym stopniu uczestnicy rynku będą uwzględniać w swoich bieżących decyzjach zmiany dochodu w przyszłych okresach.

[41] CDS (*credit default swap*) jest finansowym instrumentem pochodnym, który polega na tym, że posiadacz instrumentu dłużnego zawiera transakcję typu *swap* mającą na celu przeniesienie ryzyka kredytowego związanego z posiadanym instrumentem dłużnym na inny podmiot, z reguły firmę ubezpieczeniową lub bank. W ramach tej transakcji nabywca *swapa* (posiadacz instrumentu dłużnego) płaci

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i pożyczek przez sektor prywatny, ponieważ pogarszająca się sytuacja finansów publicznych zwiększa prawdopodobieństwo zwiększania podatków w przyszłości, co oznacza niższą oczekiwaną stopę przychodu z działalności gospodarczej w przypadku sektora prywatnego. W takiej sytuacji wierzyciele będą domagać się wyższych stóp procentowych także od pożyczkobiorców prywatnych, a to spowolni aktywność gospodarczą.

Drugi kanał oddziaływania ryzyka suwerennej niewypłacalności na stopę wzrostu PKB to bezpośredni wpływ rosnącej relacji długu publicznego do PKB na sektor bankowy. Wzrost zadłużenia suwerennego wywołuje nie tylko obniżenie ratingów inwestycyjnych i kredytowych w przypadku zadłużonego kraju, ale także obniżenie ratingów dotyczących jego instytucji finansowych, ponieważ żadna instytucja finansowa nie może mieć wyższego ratingu niż kraj, w którym jest ulokowana. Warunek ten wynika zarówno z formalnych lub domniemanych gwarancji rządu wobec własnych instytucji finansowych, jak i z tego, że krajowe instytucje finansowe są z reguły głównymi wierzycielami własnego rządu (jest to tzw. zjawisko *home bias*)[42], a więc pozostają bardziej narażone na straty w przypadku niewypłacalności kraju. W związku z tym obniżka ratingu państwa pociąga za sobą automatycznie spadek ratingu głównych banków, co z kolei zwiększa koszty pozyskiwania kapitału przez banki i zmusza je do ograniczania akcji kredytowej. Spowodowane tym zacieśnienie kredytowe (tzw. *credit crunch*) utrudnia dostęp sektora prywatnego do zasobów kapitału. W rezultacie mechanizm ryzyka suwerennej niewypłacalności wpływa – poprzez wyższe koszty kredytu i ograniczenie jego podaży – na spowolnienie wzrostu gospodarczego.

Wzrost poziomu zadłużenia wywołuje także wzrost rynkowych stóp procentowych. Powyżej pewnego poziomu zadłużenia dalszy jego wzrost przekłada się także na wzrost rentowności skarbowych papierów wartościowych. Rentowności rosną, ponieważ z jednej strony zwiększona zostaje premia za ryzyko związane z niewypłacalnością kraju, a z drugiej strony następuje wzrost podaży skarbowych

sprzedawcy *swapa* okresowo składkę ubezpieczeniową aż do momentu wygaśnięcia kontraktu. W zamian za składkę sprzedawca *swapa* zobowiązuje się wypłacić kwotę odpowiadającą równowartości instrumentu dłużnego w przypadku niewypłacalności dłużnika. CDS stanowi więc ubezpieczenie na wypadek niewypłacalności trzeciej ze stron.

[42] Zjawisko *home bias* polega na tym, że banki w danym kraju posiadają w swoich portfelach aktywów nieproporcjonalnie większe zasoby skarbowych papierów wartościowych emitowanych przez własny rząd. Jest to spowodowane między innymi tym, że, zgodnie z krajowymi regulacjami nadzorczymi, takie papiery nie są obciążone ryzykiem oraz tym, że są potrzebne bankom jako zastaw w operacjach otwartego rynku przeprowadzanych we współpracy z krajowym bankiem centralnym.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

papierów wartościowych, konieczny do sfinansowania rosnących kosztów długu, co obniża ich cenę i podnosi rentowność. Ogólnie można więc stwierdzić, że – powyżej pewnego poziomu zadłużenia – wyższym poziomom długu towarzyszą niższe stopy wzrostu PKB oraz wyższe stopy oprocentowania długu.

3.7.5. Zakres swobody fiskalnej

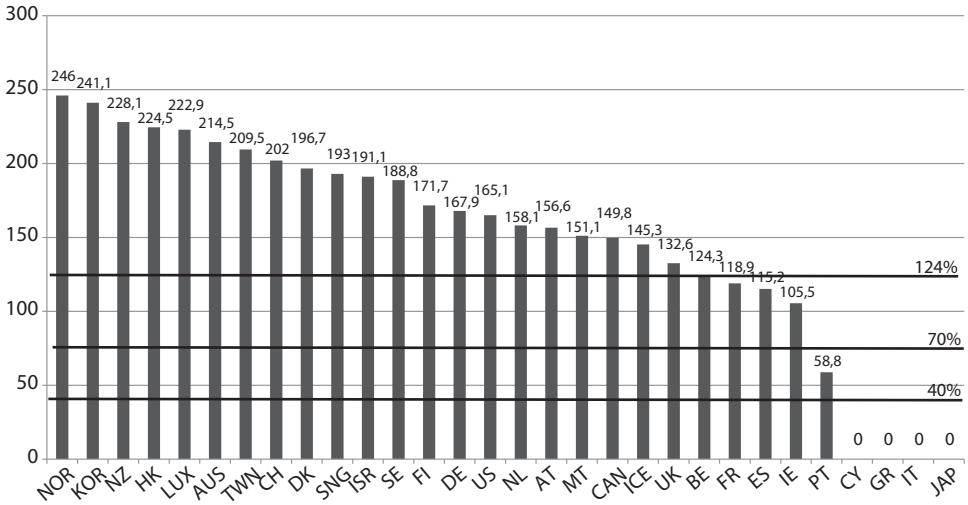
Zakres swobody fiskalnej lub inaczej przestrzeń fiskalna (*fiscal space*) oznacza możliwość zwiększenia ekspansji fiskalnej bez podważania długookresowej stabilności finansów publicznych. Pojęcie przestrzeni fiskalnej pojawiło się w literaturze ekonomicznej w pierwszej dekadzie XXI wieku w analizach Międzynarodowego Funduszu Walutowego, w kontekście prac nad dopuszczalnym poziomem zadłużenia publicznego. Heller (2005) określa ogólnie przestrzeń fiskalną jako zakres swobody manewru w budżecie, umożliwiający wygospodarowanie środków na pożądane cele bez pogorszenia finansowej sytuacji rządu lub stabilności gospodarki. Z kolei Ostry i in. (2010) definiują przestrzeń fiskalną bardziej precyzyjnie jako różnicę między aktualnym stanem zadłużenia publicznego a górnym dopuszczalnym limitem stanu zadłużenia, określonym przez dotychczasową historię dostosowań fiskalnych danego kraju. Według tej definicji, przestrzeń fiskalna oznacza zdolność państwa do dalszego bezpiecznego zadłużania się. Powracając do sytuacji przdstawionej na rysunku 3.3, przestrzeń fiskalną kraju N można określić jako różnicę między aktualnym poziomem zadłużenia a poziomem zadłużenia $d_N^2 < d_N^{max}$, powyżej którego kraj N musi realizować nadwyżkę pierwotną, aby utrzymać stabilną relacje długu do PKB.

Ostry i in. (2010) szacują maksymalny dopuszczalny limit zadłużenia odnośnie do 23 krajów rozwiniętych na podstawie historycznych danych dotyczących reakcji polityki budżetowej na wzrost zadłużenia, obejmujących lata 1970–2007. Autorzy pokazują, że przy niskich poziomach zadłużenia polityka fiskalna nie reaguje na wzrost zadłużenia obniżaniem deficytu pierwotnego (lub powiększaniem nadwyżki), ponieważ wzrost zadłużenia nie powoduje wzrostu stóp procentowych. W miarę wzrostu zadłużenia stopy procentowe na długu rosną, a polityka fiskalna ulega zacieśnieniu, ponieważ rząd stara się nie dopuścić do ich nadmiernego wzrostu i utrzymać odpowiednio duży dystans do maksymalnego dopuszczalnego poziomu zadłużenia. Powyżej tego poziomu wysiłki rządu słabną jednak ze względu na społeczny opór wobec rosnących wyrzeczeń, stopy procentowe drastycznie wzrastają i dług wymyka się spod kontroli.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Metodologia szacowania wielkości przestrzeni fiskalnej opracowana przez Ostry i in. (2010) wykorzystuje regresję między wzrostem długu w okresie (t – 1) a reakcją polityki fiskalnej w postaci obniżenia salda pierwotnego budżetu w okresie (t). Metodologia ta jest wykorzystywana m.in. przez agencję ratingową Moody's, która systematycznie oblicza przestrzeń fiskalną w przypadku wybranych krajów. Wykres 3.9 przedstawia oszacowane w ten sposób wielkości przestrzeni fiskalnej wśród 30 krajów rozwiniętych w 2014 roku.

Wykres 3.9. Przestrzeń fiskalna w wybranych krajach rozwiniętych – stan na maj 2014 (dane w % PKB)[a]



a) Różnica między maksymalnym dopuszczalnym limitem zadłużenia a aktualnym poziomem zadłużenia publicznego.

Źródło: Moody's Analytics, https://www.economy.com/dismal/tools/global-fiscal-space-tracker (27.03.2016).

Na wykresie zaznaczono trzy wartości progowe. Powyżej linii 124% PKB przestrzeń fiskalna jest bardzo duża i ryzyko przekroczenia limitu zadłużenia staje się minimalne. W przypadku wartości przestrzeni fiskalnej mieszczącej się między 70% a 124% PKB konieczna jest już bardziej ostrożna polityka fiskalna. Odnośnie do wartości przestrzeni fiskalnej mieszczącej się między 40% a 70% PKB ryzyko przekroczenia limitu zadłużenia okazuje się już poważne, w związku z czym polityka fiskalna powinna być restrykcyjna. Wreszcie w przypadku wartości przestrzeni fiskalnej kształtującej się poniżej 40% PKB ryzyko przekroczenia limitu zadłużenia jest bardzo duże i kraj powinien realizować pierwotna nadwyżkę budżetową, aby

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

uniknąć kryzysu. Wyniki uzyskane dla poszczególnych krajów są zasadniczo zgodne z intuicją. Największą swobodą fiskalna cieszą się kraje skandynawskie (Norwegia, Dania, Szwecja), kraje Azji Południowo-Wschodniej (Korea Południowa, Hong Kong, Taiwan i Singapur) oraz Nowa Zelandia, Luksemburg, Australia i Szwajcaria. Najbardziej stabilne kraje strefy euro charakteryzują się już wyraźnie niższą przestrzenią fiskalną (Finlandia, Niemcy, Holandia, Austria). Stany Zjednoczone uplasowały się mniej więcej pośrodku tabeli wyników, a Wielka Brytania w dolnej części, lecz w obu wymienionych krajach przestrzeń fiskalna jest większa niż wartość progowa 124% PKB, co oznacza, że w praktyce nie występuje tam ryzyko przekroczenia limitu zadłużenia. Znacznie gorzej wygląda sytuacja we Francji, Belgii i Hiszpanii, które są na granicy lub poniżej progu 124% PKB, oraz w Portugalii, gdzie przestrzeń fiskalna znajduje się poniżej 70% PKB. Najgorsza sytuacja panuje zaś w czterech krajach o najwyższym poziomie długu publicznego (Cypr, Grecja, Włochy i Japonia), gdzie przestrzeń fiskalna jest zerowa, a państwa te są zasadniczo pozbawione swobody działań w zakresie polityki fiskalnej.

3.8. Relacje między długiem publicznym a wzrostem gospodarczym

3.8.1. Obserwacje empiryczne

Ekonomiści od dawna zdawali sobie sprawę z tego, że wysokie zadłużenie publiczne może negatywnie wpływać na tempo wzrostu gospodarczego. Zależność ta była przedmiotem wielu badań empirycznych. Wykorzystując dane dotyczące krajów OECD z lat 1960–2008, eksperci Międzynarodowego Funduszu Walutowego oszacowali, że wzrostowi relacji długu do PKB o 10% PKB towarzyszy spadek tempa wzrostu gospodarczego średnio o 0,46 punktu procentowego. Przeprowadzone symulacje wskazują, że podwojenie stanu zadłużenia (z 50% do 100% PKB) powoduje trwałe obniżenie produkcji, konsumpcji i poziomu majątku produkcyjnego o 2 punkty procentowe w stosunku do sytuacji, w której zadłużenie nie ulega zmianie (Ostry i in. 2015).

W obliczeniach tych zakłada się, że analizowana zależność ma charakter liniowy. Tymczasem wiadomo już, że wpływ zadłużenia na wzrost jest nieliniowy – przy niskich poziomach zadłużenia dodatkowy dług może przyspieszać wzrost gospodarczy, a przy wysokich wartościach tego czynnika może go hamować. Powstaje więc

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pytanie o to, od jakiego poziomu zadłużenia dalsze zadłużanie się ma negatywny wpływ na wzrost ekonomiczny. Reinhart, Rogoff i Savastano (2003) badali historię zadłużenia zagranicznego w odniesieniu do ponad 100 krajów o gospodarkach wschodzących (*emerging economies*) i rozwijających się w okresie ponad 175 lat. Przyjmując za punkt wyjścia hipotezę mówiącą o tym, że nadmierne obciążenie gospodarki kosztami obsługi długu wpływa negatywnie na tempo wzrostu gospodarczego, starali się znaleźć odpowiedź na pytanie o to, przy jakim poziomie relacji długu do PKB rosnące koszty obsługi długu zaczynają negatywnie oddziaływać na dynamikę wzrostu gospodarczego. Otrzymane wyniki wskazują, że taki poziom graniczny rzeczywiście istnieje i że w krajach rozwijających się jest on znacznie niższy niż w krajach rozwiniętych, kształtując się ogólnie na poziomie poniżej 50%, a w wielu przypadkach nawet na poziomie około 15–20%. W krajach tych kłopoty z obsługą zadłużenia wywoływały silne spowolnienie tempa wzrostu (zob. również Ciżkowicz, Michalska i Rzońca 2011). Najniższy poziom krytyczny zadłużenia był obserwowany w krajach, które miały w swej niedawnej historii epizody zawieszenia obsługi długu, restrukturyzacji lub suwerennego bankructwa – kraje te określono mianem krajów o niskiej tolerancji długu (*debt-intolerant countries*). Z kolei Reinhart i Rogoff (2010) na podstawie danych przekrojowych dotyczących krajów rozwiniętych wyliczają, że wartość krytyczna relacji długu do PKB, powyżej której dług wyraźnie obniża tempo wzrostu, wynosi około 90% PKB. Niemal identyczne wyniki uzyskali Cecchetti i in. (2011), którzy oszacowali taki krytyczny poziom zadłużenia w przypadku krajów OECD na 80–100% PKB, oraz Checherita i Rother (2010), którzy wykazali na podstawie danych panelowych zebranych dla 12 krajów strefy euro, że krytyczny poziom zadłużenia mieści się w przedziale 90%–100%.

Wyniki otrzymane przez Reinhart i Rogoff (2010) zostały częściowo zakwestionowane przez Herndon, Ash i Pollin (2013), którzy dopatrzyli się licznych błędów w obliczeniach. Ich zdaniem, o ile tempo wzrostu PKB jest rzeczywiście negatywnie skorelowane z relacją długu do PKB, o tyle siła tej korelacji jest słabsza niż szacowali to Reinhart i Rogoff i w efekcie krytyczny poziom zadłużenia wśród krajów rozwiniętych jest wyższy niż 90% PKB.

Oczywiście krytyczny poziom zadłużenia dotyczący poszczególnych krajów zależy od ich indywidualnej sytuacji, w tym od takich czynników jak udział długu zagranicznego w całości zadłużenia, średni okres zapadalności długu, wartość aktywów w posiadaniu sektora rządowego, struktura demograficzna kraju, a także od fazy cyklu koniunkturalnego i awersji do ryzyka inwestorów (Iron i Bevins 2010, Eichengreen i in. 2011). Niezależnie od tych dodatkowych uwarunkowań, wyniki

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

otrzymane przez Reinhart i Rogoffa oraz Cecchettiego i in. mogą stanowić punkt odniesienia przy określaniu wartości progowej zadłużenia wśród krajów rozwiniętych.

3.8.2. Wpływ zadłużenia publicznego na tempo wzrostu gospodarczego i poziom stóp procentowych w krajach Unii Europejskiej i strefy euro

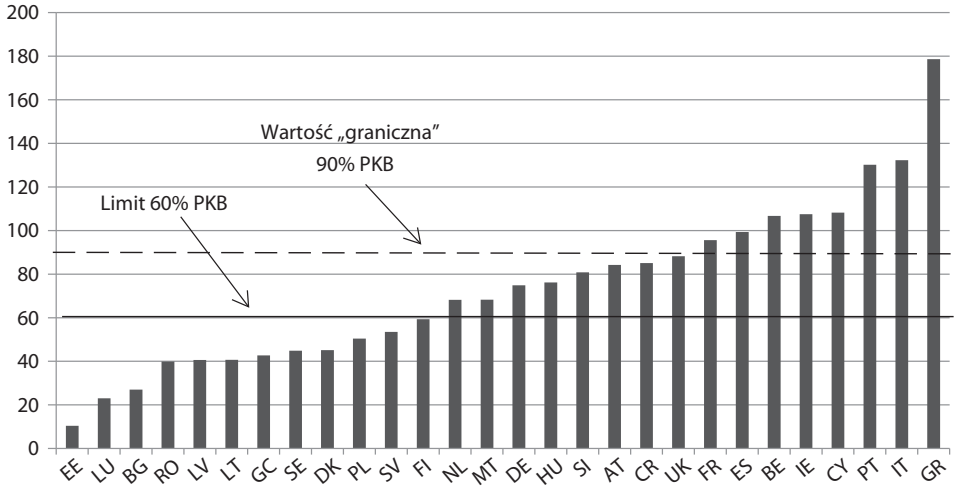
Przegląd sytuacji w krajach członkowskich UE z lat 2010–2014 w zasadzie potwierdza omówione wyżej wyniki. Wykres 3.10 pokazuje, że na koniec 2014 roku w ośmiu krajach członkowskich UE poziom zadłużenia publicznego przekraczał wartość progową 90% (linia przerywana). Cztery spośród tych krajów zostały objęte programem pomocy finansowej (Grecja, Irlandia, Portugalia, Cypr), a dwa doświadczyły przejściowo poważnych trudności w utrzymaniu bieżącej płynności (Włochy, Hiszpania). Co więcej, gospodarki tych krajów w latach 2010–2014 znajdowały się albo w bardzo głębokim kryzysie (Grecja), albo w stanie przewlekłej recesji (Portugalia, Włochy, Cypr). Istnieje zresztą jeszcze trzecia możliwość: obserwowane tempo wzrostu okazało się bardzo niskie i spadło poniżej średniej dotyczącej całej UE 28. Potwierdza to, że wysoki poziom zadłużenia jest skorelowany ujemnie z tempem wzrostu gospodarczego. Wykres 3.11 pokazuje, że istnieje statystycznie istotna, ujemna korelacja pomiędzy poziomem zadłużenia a tempem wzrostu w krajach UE 27 w 2011 roku: wzrostowi stosunku długu do PKB o 10 punktów procentowych towarzyszy – średnio – spadek rocznego tempa wzrostu PKB o 0,5 punktu procentowego. Z kolei krytyczny poziom zadłużenia, powyżej którego tempo wzrostu PKB staje się ujemne, wynosi w przypadku tej próby niemal dokładnie 100% PKB[43].

Jeśli chodzi o wpływ wysokich poziomów zadłużenia na rynkowe stopy oprocentowania długu, to ujawnił się on na początku 2008 roku, kiedy inwestorzy finansowi zorientowali się, że ryzyko związane z papierami skarbowymi poszczególnych państw strefy euro jest bardziej zróżnicowane, niż zakładano wcześniej, z uwagi na brak wzajemnych gwarancji długu i brak automatycznego mechanizmu łączenia ryzyka (*pooled risk*) w ramach unii walutowej. Od tego momentu stopień rentowności różnicuje się i widać wyraźny związek między poziomem zadłużenia publicznego a wysokością stóp rentowności obligacji rządowych.

[43] Z równania regresji oszacowanego w przypadku danych na wykresie 3.11 otrzymujemy: $0{,}0545x - 5{,}4366 > 0$, i stąd $x > 99{,}75\%$.

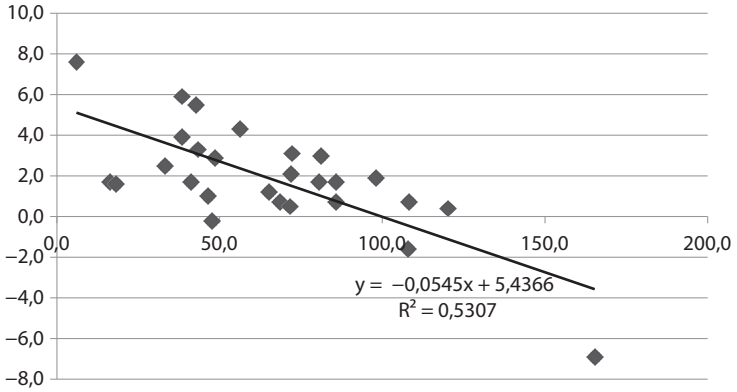
Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 3.10. Dług publiczny w krajach UE 28 – stan na dzień 31.12.2014 roku (dane w % PKB)



Źródło: na podstawie danych Eurostatu.

Wykres 3.11. Dług publiczny i tempo wzrostu PKB w 2011 roku w krajach UE 27 (dane w % PKB)



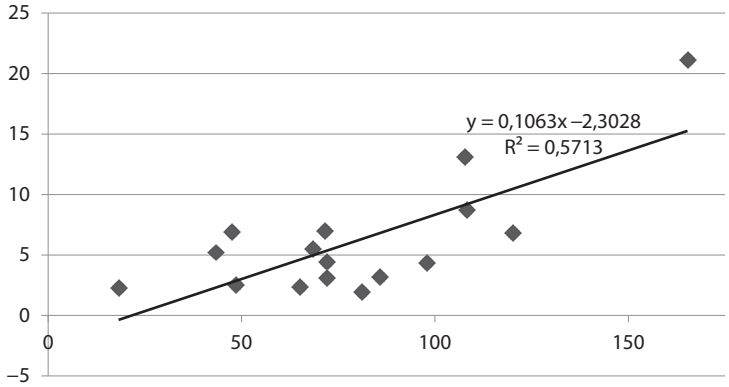
Źródło: obliczenia własne na podstawie danych Eurostatu.

Wykres 3.12 przedstawia zależnośćmiędzy relacją długu do PKB i rentownością obligacji 10-letnich w krajach strefy euro na koniec 2011 roku. Analiza celowo ograniczona jest do krajów strefy euro, aby wyeliminować wpływ takich czynników jak odrębna polityka pieniężna i kursowa na wysokość stóp rentowności. Dane

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przedstawione na wykresie dowodzą, że pomiędzy analizowanymi zmiennymi zachodzi statystycznie istotna dodatnia zależność: zgodnie z oszacowanym równaniem regresji wzrostowi stosunku długu do PKB o 1 punkt procentowy towarzyszy – średnio – wzrost rentowności obligacji o około 11 punktów bazowych.

Wykres 3.12. Dług publiczny (% PKB) i rentowności obligacji (%) w krajach strefy euro – stan na grudzień 2011 roku



Uwaga: wielkość relacji długu do PKB na koniec 2011 roku. Poziom rentowności jest obliczony jako średnia z dziennych notowań rentowności 10-letnich obligacji skarbowych w grudniu 2011 roku.

Źródło: obliczenia własne na podstawie danych Eurostatu.

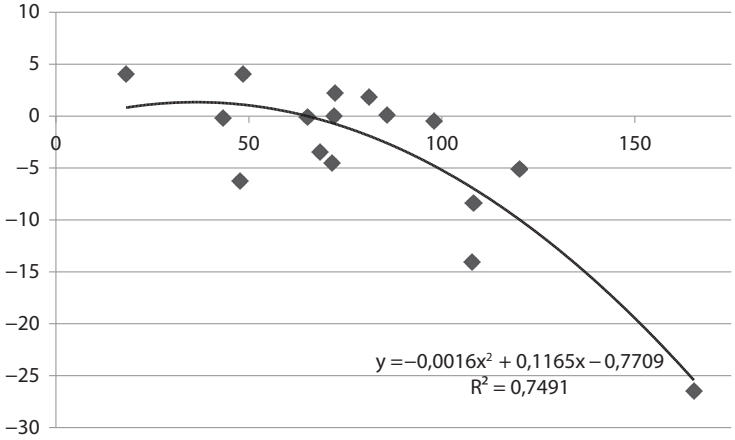
3.8.3. Maksymalna politycznie możliwa nadwyżka pierwotna i pułapka długu

Można teraz uwzględnić wpływ poziomu zadłużenia zarówno na wyrażenie $(y - i)$, jak i na poziom deficytu pierwotnego, zapewniającego stabilizację długu na danym poziomie. Duża – w przypadku modułu – ujemna wartość $(y - i)$ w równaniu (3.26) wymaga, aby dany kraj realizował dużą nadwyżkę pierwotną w celu stabilizacji relacji długu do PKB. Uzyskanie dużej nadwyżki pierwotnej wymaga jednak drastycznych ograniczeń w zakresie wydatków publicznych, w tym również tych przeznaczanych na cele społeczne. Budzi to powszechny sprzeciw. Utrzymanie dużej nadwyżki pierwotnej powyżej danego krytycznego poziomu staje się niewykonalne ze względów politycznych, a relacja długu do PKB wymyka się spod kontroli. O ile w krótkim okresie jednorazowe dostosowania fiskalne mogą być dość wysokie, o tyle utrzymanie dużej nadwyżki pierwotnej przez wiele lat jest na ogół niemożliwe. Dzieje się tak, gdy nadwyżka fiskalna, konieczna w przypadku obsługi

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zadłużenia w wysokości zapewniającej niezwiększanie się relacji długu do PKB, przekracza opisaną przez Blancharda (1984) „maksymalną politycznie możliwą nadwyżkę pierwotną". Można zatem zakładać, że istnieje górna granica dla zmiennej *b*, określona przez uwarunkowania polityczne. Powyżej tej granicy bodźce do zawieszenia obsługi zadłużenia przez dany kraj stają się silniejsze niż bodźce do dalszego honorowania zobowiązań płatniczych.

Wykres 3.13. Relacja długu do PKB (oś pozioma) i wartość wyrażenia (*y* – *i*) (oś pionowa) odnotowane wśród 16 krajów strefy euro 2011 roku



Uwaga: relacja długu publicznego do PKB na koniec 2011 roku. Wartości (*y* – *i*) obliczone zostały jako różnica pomiędzy stopą nominalnego wzrostu PKB w 2011 roku a średnią stopą rentowności rządowych obligacji 10-letnich w grudniu 2011 roku.

Źródło: obliczenia własne na podstawie danych Eurostatu.

Wykres 3.13 ukazuje związek pomiędzy poziomem zadłużenia a wielkością wyrażenia (*y* – *i*), przy czym – z powodów podanych wyżej – analiza ograniczona jest tylko do krajów strefy euro. Dane przedstawione na wykresie potwierdzają ogólną zależność zaprezentowaną na rysunku 3.3. Jak można zauważyć, pomiędzy analizowanymi zmiennymi zachodzi statystycznie istotna zależność ujemna: po przekroczeniu pewnego poziomu zadłużenia, dalszemu wzrostowi stosunku długu do PKB o 1 punkt procentowy towarzyszy coraz szybszy spadek wyrażenia (*y* – *i*). Porównanie parametrów dotyczących dopasowania różnych typów specyfikacji równania regresji do danych sugeruje, że badana relacja ma charakter nieliniowy[44].

[44] Wartość *R (kwadrat)* w przypadku równania regresji liniowej wyniosła 0,61.

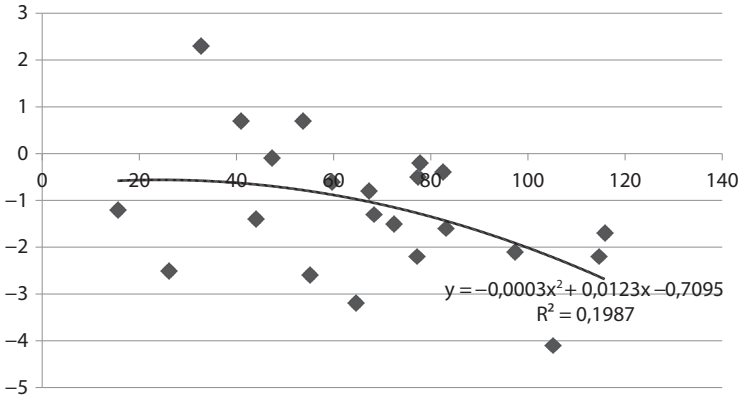
Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wzrost zadłużenia powoduje najpierw wzrost wartości wyrażenia ($y - i$), a gdy poziom zadłużenia przekroczy wielkośc 36% PKB, wartość wyrażenia ($y - i$) zaczyna spadać. W przypadku relacji długu do PKB mieszczących się w przedziale 50–100% wartość wrażenia ($y - i$) spada z 1,05 do –5,12, gdy dług osiąga natomiast 150% PKB, wartość ($y - i$) spada już do –19,30.

Przyspieszony spadek wartości ($y - i$) w przypadku wysokich wartości zadłużenia potwierdza istnienie granicy, poza którą obsługa zadłużenia staje się niewykonalna. Wymagana nadwyżka pierwotna staje się tak duża, że rząd zadłużonego kraju nie jest w stanie – respektując zasady systemu demokratycznego – narzucić społeczeństwu tak dużych wyrzeczeń i może zdecydować się zawiesić obsługę zadłużenia. Własny interes polityczny bierze górę nad interesem wierzycieli[45].

Analogiczną prawidłowość można zaobserwować w przypadku większej grupy europejskich i pozaeuropejskich krajów rozwiniętych. Wykorzystując dane zapożyczone z opracowania Ostry i in. (2010), udaje się wykazać, że w krajach tych wyższym poziomom zadłużenia odpowiadają coraz niższe (co do modułu) wartości wyrażenia ($y - i$). Wykres 3.14 pokazuje, że podobnie jak w krajach strefy euro zależność ta ma charakter nieliniowy, choć siła związku jest słabsza niż w poprzednim przypadku[46].

Wykres 3.14. Relacja długu do PKB (oś pozioma) i wartość wyrażenia ($y - i$) (oś pionowa) odnotowane wśród 23 krajów rozwiniętych[a] (dane w %)



a) Relacja długu do PKB dla 2009 roku. Wyrażenie ($y - i$) wyliczone w przypadku średnich prognozowanych przez MFW stóp rentowności obligacji rządowych i nominalnych stóp wzrostu PKB na lata 2010–2015.

Źródło: obliczenia własne na podstawie Ostry i in. 2010 (zob. tabela 1, s. 10).

[45] Wydaje się, że w maju 2012 roku Grecja znalazła się w takiej właśnie sytuacji.

[46] Wartość *R (kwadrat)* w przypadku równania regresji liniowej wyniosła 0,17.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Jeszcze inne podejście do problemu spłaty zadłużenia przyjmuje za punkt wyjścia ekonomię polityczną kosztów obsługi długu, czyli ich podział w obrębie danego społeczeństwa. Podkreśla się mianowicie, że zaciąganie pożyczek na koszt przyszłych pokoleń jest możliwe tylko wówczas, kiedy nierówności dochodowe pomiędzy grupami społecznymi nie są zbyt duże. Z kolei gdy rozpiętości dochodowe są znaczne, a bogactwo – w tym zasób obligacji rządowych – pozostaje w gestii niewielkiej mniejszości, pozostała większość ma silną motywację aby przestać spłacać podatkami wcześniej zaciągnięte pożyczki (Tabellini 1991, Bulow i Rogoff 1989). Oznacza to, że w krajach, w których przy danym poziomie zadłużenia utrzymują się wysokie rozpiętości dochodowe,, pojawi się silniejszy opór wobec polityki „zaciskania pasa", i w efekcie większe będzie prawdopodobieństwo bankructwa suwerennego.

3.8.4. Czy dług trzeba spłacać?

Pytanie nie jest bynajmniej pozbawione sensu i tylko na pozór tchnie prowokacją. W rzeczywistości zaciągnięty dług trzeba bowiem przede wszystkim skrupulatnie obsługiwać, niekoniecznie jednak spłacać – w sensie zmniejszenia poziomu zadłużenia do zera. Wierzyciele, w tym zwłaszcza profesjonalni pośrednicy finansowi, tacy jak banki oraz inwestorzy instytucjonalni, są z reguły zainteresowani dobrym ulokowaniem posiadanych środków, w sposób przynoszący pożądany dochód i gwarantujący bezpieczeństwo. Jeśli udzielenie pożyczki rządowi zapewnia im oczekiwane korzyści, są gotowi finansować potrzeby pożyczkowe tego rządu praktycznie w nieskończoność. O prawdziwości tego zjawiska świadczy to, że ogromna większość rządów na świecie pozostaje permanentnie zadłużona i żaden z nich dobrowolnie nie spłaca zadłużenia do zera.

Przychodzi również zauważyć, że powyższe pytanie stało się szczególnie aktualne w czasie ostatniego kryzysu finansowego i gospodarczego, kiedy wiele zadłużonych państw podjęło wysiłki w celu skonsolidowania swoich finansów publicznych, dążąc do obniżenia poziomu zadłużenia publicznego. Wysiłki te oraz towarzyszące im spowolnienie wzrostu gospodarczego, spotkały się z krytycznymi opiniami co do celowości i ekonomicznego sensu tych działań. Więcej na temat wpływu konsolidacji fiskalnej na tempo wzrostu piszemy w dalszej części rozdziału (zob. punkt 3.11.2), obecnie omówione zostaną zaś tylko zasadnicze argumenty teoretyczne.

Pomińmy trywialny przypadek kraju wysoko zadłużonego, który ma ograniczoną przestrzeń fiskalną i napotyka na trudności z obsługą długu. W takiej sytuacji rośnie ryzyko niewypłacalności (*sovereign default*) i wierzyciele będą domagać

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

się jak najszybszej spłaty zadłużenia, aby przenieść swoje środki do bezpiecznych „przystani". Lecz co dzieje się z krajem, który ma reputację solidnego dłużnika i bez problemów obsługuje zadłużenie? Czy kraj taki też powinien spłacić swoje długi? Na to pytanie próbowali odpowiedzieć ekonomiści Międzynarodowego Funduszu Walutowego. W ich opinii, optymalnym rozwiązaniem w takiej sytuacji jest „żyć z długiem", pozwalając, aby wzrost gospodarczy „sam z siebie" stopniowo redukował relację długu do PKB (Ostry i in. 2015).

Jakie argumenty przemawiają za przyjęciem takiego stanowiska? Z jednej strony rządowi opłaca się obniżać poziom zadłużenia publicznego, ponieważ wysokie zadłużenie zwiększa ryzyko wystąpienia kryzysu zadłużenia w przypadku nagłego pogorszenia się koniunktury gospodarczej, załamania rynkowego czy kryzysu w bezpośrednim sąsiedztwie kraju. W związku z tym obniżenie poziomu zadłużenia niweluje to ryzyko i dzięki temu zmniejsza stopy procentowe, po jakich rząd zaciąga pożyczki. Ponadto, wysoki poziom zadłużenia wymaga przeznaczania dużych środków na obsługę długu. Pociąga to za sobą konieczność utrzymywania wysokiego opodatkowania i związanych z tym zniekształceń rynkowych, co ma spowalniający wpływ na wzrost gospodarczy. Obniżenie zadłużenia, zwłaszcza z wysokiego poziomu, zmniejsza długofalowo koszty obsługi długu, a to pozwala obniżyć podatki i przeznaczyć więcej środków na inwestycje oraz przyspieszenie wzrostu gospodarczego. Z drugiej jednak strony zmniejszenie zadłużenia wymaga doraźnego zwiększenia obciążeń podatkowych. W takiej sytuacji zniekształcający wpływ doraźnego wzrostu podatków może przewyższyć potencjalne korzyści płynące z obniżenia ryzyka niewypłacalności i potencjalnego przyspieszenia wzrostu gospodarczego.

Próbując rozstrzygnąć ten dylemat, Ostry i in. (2015) dokonują rozróżnienia miedzy krajami wysoko zadłużonymi, w których przestrzeń fiskalna jest mała lub zerowa, oraz krajami o niższym poziomie zadłużenia, gdzie przestrzeń fiskalna pozostaje duża. Pierwsze z wymienionych państw praktycznie nie mają wyboru – ograniczony dostęp (lub brak dostępu) do rynków finansowych zmusza je do obniżania długu i akceptowania związanych z tym kosztów makroekonomicznych. Natomiast kraje o umiarkowanym zadłużeniu nie powinny spłacać długu za wszelką cenę – w ich przypadku wskazane jest raczej stopniowe „wyrastanie" z długu dzięki wzrostowi PKB.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.9. Wpływ wydatków publicznych i deficytu na popyt krajowy. Spór wokół polityki stabilizacyjnej

3.9.1. Mnożnik wydatków w modelu IS-LM

Jak wspomniano wcześniej, stabilizacyjna funkcja polityki fiskalnej polega na uzupełnianiu niedoboru popytu prywatnego dodatkowym popytem publicznym, tak aby utrzymać wysoki poziom produkcji, dochodów i zatrudnienia. Za twórcę koncepcji wykorzystania polityki fiskalnej do kształtowania poziomu zagregowanego popytu uważa się Johna Maynarda Keynesa, który w *Ogólnej teorii…* opisał zjawisko chronicznej nierównowagi makroekonomicznej w gospodarce rynkowej i wykazał, że popyt rządowy może skutecznie uzupełniać popyt prywatny w sytuacji niedoboru zagregowanego popytu. W przeciwieństwie do ekonomistów klasycznych Keynes twierdził, że w normalnych warunkach w gospodarce rynkowej ceny są sztywne w krótkim okresie oraz istnieją wolne moce produkcyjne, dzięki czemu wzrost wydatków rządowych przekłada się na wzrost produkcji, a nie na wzrost cen. Sposób, w jaki wydatki rządowe wpływają na produkcję, jest określany przez mechanizm mnożnika wydatków. Mechanizm ten, omawiany w podręcznikach makroekonomii, stanowi jeden z głównych elementów składowych teorii Keynesa[47].

Wygodnie jest zacząć omawianie mechanizmu mnożnika wydatków od jego wersji najbardziej elementarnej. Na wstępie należy przypomnieć, że w ramach modelu IS-LM konsumpcja indywidualna C jest funkcją dochodu rozporządzalnego Y (zob. rozdział 2, równanie (2.4)):

$$C = a + cY, \quad a, c > 0, \tag{3.29}$$

gdzie współczynnik c oznacza skłonność do konsumpcji (*propensity to consume*). Wskazuje on, ile groszy z każdej złotówki rozporządzalnego dochodu gospodarstwa domowe przeznaczają na konsumpcję. Pozostała część dochodu, $(1 - c)$, jest wówczas oszczędzana, a współczynnik $s = (1 - c)$ nazywamy skłonnością do oszczędzania (*propensity to save*). Gdy podaż staje się doskonale elastyczna, produkcja dostosowuje się do popytu i zostaje spełniony warunek równowagi podaży i popytu na rynku dóbr:

[47] Teorię mnożnika rozwijali też polscy ekonomiści – zob. m.in. Kalecki (1936, 1939) i Sulmicki (1945).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$Y = C(Y) + I(r) + G, \tag{3.30}$$

gdzie inwestycje I są funkcją stopy procentowej r, a symbolem G oznaczone są wydatki rządowe. Różniczkując całkowicie równanie (3.30) i dzieląc obustronnie przez dG, otrzymujemy:

$$dY / dG = \partial C / \partial Y \left(\frac{dY}{dG} \right) + \partial I / \partial r \left(\frac{dr}{dG} \right) + 1. \tag{3.31}$$

Załóżmy, że $dr = 0$, co oznacza w pełni akomodacyjną politykę pieniężną (stopy procentowe nie rosną, bo bank centralny dostosowuje podaż pieniądza do zwiększonego popytu na pieniądz w związku ze wzrostem wydatków rządowych). Ponieważ $\partial C / \partial Y = c$, otrzymujemy wzór na mnożnik wydatków:

$$\frac{dY}{dG} = c\left(\frac{dY}{dG} \right) + 1,$$

czyli po przekształceniu:

$$\frac{dY}{dG} = \frac{1}{1 - c} = \frac{1}{s}. \tag{3.32}$$

Równanie (3.32) informuje o tym, o ile jednostek wzrośnie produkcja Y, gdy wydatki rządowe G wzrosną o jednostkę (jednostkę pieniężną lub o 1% PKB). Ponieważ krańcowa skłonność do konsumpcji c przyjmuje przeciętne wartości między 0,8 a 0,9, to skłonność do oszczędzania s wynosi między 0,1 a 0,2. Dla tych parametrów wielkość mnożnika waha się więc między 5 a 10. W opisanych warunkach polityka pobudzania gospodarki za pomocą wzrostu wydatków rządowych okazuje się więc bardzo skuteczna.

Badania empiryczne dowodzą jednak, że w praktyce wielkość mnożnika we współczesnych gospodarkach jest o wiele niższa i waha się przeciętnie miedzy 0,5 a 1,0 (zob. np. Alesina i Ardagna 2010), co znacząco ogranicza skuteczność polityki fiskalnej w pobudzaniu produkcji, dochodu i zatrudnienia. Wynika to z oddziaływania kilku czynników. Po pierwsze, gospodarstwa domowe płacą podatki, które zmniejszają ich dochód rozporządzalny i ograniczają wydatki konsumpcyjne. Uwzględniając stopę podatkową t, równanie konsumpcji przyjmuje postać $C = a + c(1 - t)Y$. Nałożenie podatków powoduje, że poszczególne przyrosty produkcji ulegają obniżeniu, a wzór ogólny na mnożnik przyjmuje następującą postać:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$dY/dG = 1/\left[1 - c(1-t)\right]. \tag{3.33}$$

Dla każdego $t > 0$, wartość mnożnika (3.33) jest mniejsza niż wartość mnożnika (3.32).

Po drugie, w gospodarce otwartej część produkcji jest eksportowana, a część rozporządzalnego dochodu wydawana na zakup towarów importowanych. O ile eksport stanowi zmienną egzogeniczną (zależy od zmiennych nie uwzględnionych w modelu), o tyle import jest funkcją poziomu dochodu, $M = mY$, gdzie $0 < m < 1$, wyraża skłonność do importu (*propensity to import*). Wówczas warunek równowagi podaży i popytu w gospodarce otwartej uwzględnia saldo wymiany ($E - mY$), czyli eksport netto:

$$Y = C + I + G + (E - mY).$$

Wydatki na import powodują „wyciekanie" części dochodu za granicę i zmniejszają mnożnikowy wzrost dochodów w kraju. Jeśli krańcowa skłonność do importu wynosi $m = M/Y$, postać ogólna mnożnika staje się następująca:

$$dY/dG = \frac{1}{[(1 - c(1 - t) + m]}. \tag{3.34}$$

Ponieważ $m > 0$, wielkość mnożnika staje się jeszcze niższa.

Po trzecie, jeśli polityka pieniężna banku centralnego nie jest w pełni akomodująca (co wydaje się bardzo prawdopodobne), wzrost wydatków rządowych wywołuje pewien wzrost stóp procentowych, a to z kolei przekłada się na spadek inwestycji prywatnych wskutek efektu wypychania. W takiej sytuacji mnożnikowy wzrost dochodów i konsumpcji jest częściowo neutralizowany spadkiem inwestycji i wartość mnożnika jest ostatecznie niższa. W skrajnym przypadku, na skutek zaostrzenia polityki pieniężnej z powodu wzrostu wydatków rządowych, wartość mnożnika może nawet spaść do zera. W takiej sytuacji fiskalna stymulacja gospodarki jest całkowicie nieskuteczna[48].

Aby zbadać wpływ polityki pieniężnej na wielkość mnożnika wydatków rządowych, uchylimy obecnie założenie o stałości stopy procentowej r. Oznacza to, że nie ma pełnej akomodacji pieniężnej, a więc podaż pieniądza rośnie w mniejszym

[48] Taki pogląd reprezentują niektórzy monetaryści.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

stopniu niż popyt na pieniądz lub nie rośnie wcale. Wówczas $\partial I / \partial r < 0$, a wzór na mnożnik wydatków rządowych przybiera postać[49]:

$$\frac{dY}{dG} = c\left(\frac{dY}{dG}\right) + \frac{\partial I}{\partial r}\left(\frac{dr}{dG}\right) + 1.$$

Dzieląc i mnożąc jednocześnie drugi wyraz prawej strony równania przez (rI) i uwzględniając wzór na elastyczność zmiennych, otrzymujemy ostatecznie:

$$\frac{dY}{dG} = \frac{\varepsilon_r^I\left(\frac{dr}{r}\right)\left(\frac{I}{dG}\right) + 1}{(1 - c)}, \tag{3.35}$$

gdzie ε_r^I oznacza elastyczność inwestycji względem stopy procentowej. Ponieważ $\varepsilon_r^I < 0$, licznik ułamka po prawej stronie równania (3.35) jest zawsze mniejszy od jedności, co powoduje, że mnożnik opisany równaniem (3.35) pozostaje zawsze mniejszy niż mnożnik opisany równaniem (3.32). Wynika to stąd, że stymulujący efekt wzrostu wydatków rządowych jest słabszy ze względu na spadek inwestycji spowodowany wzrostem stóp procentowych. Zauważmy, że w przypadku dostatecznie dużych wartości elastyczności ε_r^I lub dostatecznie dużych wielkości wyrażenia $\left[\left(\frac{dr}{r}\right)\left(\frac{I}{dG}\right)\right]$ licznik prawej strony równania (3.35) może być bliski zeru lub nawet ujemny.

Tak właśnie wygląda mechanizm dostosowawczy w modelu IS-LM. Wzrost wydatków publicznych przekłada się na przesunięcie krzywej *IS* w prawo, a nowy punkt równowagi cechuje się wyższym poziomem produkcji i dochodu oraz wyższym poziomem stóp procentowych (brak pełnej akomodacji pieniężnej). Gdy inwestycje są bardzo wrażliwe na zmiany stopy procentowej, a bank centralny utrzymuje stałą podaż pieniądza, spadek inwestycji może całkowicie zrównoważyć wzrost wydatków rządowych i krzywa *IS* powraca do początkowego położenia (mnożnik = 0), a nawet może okazać się większy niż wzrost wydatków rządowych (wówczas krzywa *IS* przesuwa się jeszcze bardziej w lewo, a mnożnik staje się ujemny).

Wreszcie należy też pamiętać, że kluczowe keynesowskie założenie o całkowicie sztywnych cenach na ogół nie jest w pełni spełnione w rzeczywistości. Wzrost wydatków rządowych może więc tylko częściowo wywołać wzrost produkcji,

[49] Pamiętajmy, że $I'(r) < 0$, bo inwestycje są malejącą funkcją stopy procentowej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

a częściowo wzrost cen. Wartość mnożnika wydatków w takiej sytuacji będzie więc odpowiednio niższa. Przykład liczbowy ilustrujący działanie mnożnika wydatków przedstawiony jest w ramce 3.9.

Ramka 3.9. Działanie mnożnika wydatków – przykład liczbowy

Rozpatrzmy prosty przykład liczbowy. Załóżmy, że rząd zwiększa wydatki publiczne o jednostkę, np. o 1000 PLN. Spowoduje to wzrost produkcji i wzrost dochodów także o 1000 PLN. Jeśli skłonność do konsumpcji wynosi na przykład $c = 0{,}8$, to z kwoty dochodu brutto 1000 PLN gospodarstwa domowe 800 PLN przeznaczą na konsumpcję, a 200 PLN na oszczędności. Wzrost wydatków konsumpcyjnych o 800 PLN wywoła ponowny wzrost produkcji i dochodów o tę samą kwotę. Z tej kwoty dochodu brutto gospodarstwa domowe 640 PLN przeznaczą na konsumpcję, a 160 – na oszczędności. Ten cykl wydatków i dochodów będzie się powtarzał, przy czym przyrosty produkcji, dochodu i wydatków zaczną stawać się coraz mniejsze, zmierzając w długim okresie do zera. Całkowity przyrost produkcji równa się sumie ciągu o postaci:

$$\Sigma(Y_t) = 1\,000(1 + c + c^2 + \ldots) = 1/(1-c) = 1/(1-0{,}8) = 5\,000 \text{ PLN}. \qquad (3.36)$$

W postaci ogólnej wzór na mnożnik wydatków rządowych przyjmuje postać określoną równaniem (3.33). Ponieważ $(1 - c) = s = 0{,}2$, mnożnik wynosi w naszym przykładzie 5. Jak wspomniano wyżej, warunkiem wystąpienia efektów mnożnikowych są wolne moce produkcyjne i sztywność cen w krótkim okresie.

Wykorzystamy teraz równanie (3.33) do obliczenia mnożnika w warunkach funkcjonowania podatków, przyjmując, że stopa podatkowa wynosi $t = 0{,}1$. Wówczas całkowity przyrost produkcji – na podstawie równania (3.36) – wynosi $\Sigma(Y_t) = 3571$ PLN, a mnożnik zmniejszy się z 5 do około 3,6. Z kolei w warunkach gospodarki otwartej właściwym równaniem umożliwiającym obliczenie mnożnika jest równanie (3.34). Przykładowo, dla $c = 0{,}8$, $t = 0{,}1$ i $m = 0{,}2$, całkowity przyrost produkcji wynosi $\Sigma(Y_t) = 2083$ PLN, co oznacza, że wartość mnożnika spada do 2,08. Jak widać, urealnienie rachunku poprzez uwzględnienie podatków i importu powoduje znaczące przybliżenie teoretycznej wartości mnożnika do empirycznych obserwacji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.9.2. Mnożnik wydatków w modelu neoklasycznym*

Jak już wiadomo, w ekonomii neoklasycznej – w przeciwieństwie do modeli keynesowskich – obowiązują założenia o doskonale giętkich cenach i płacach oraz o pełnym wykorzystaniu zdolności produkcyjnych. W standardowym modelu neoklasycznym zakłada się ponadto, że uczestnicy rynku kierują się oczekiwaniami w nieskończonym horyzoncie czasu (*infinitely-lived forward-looking agents*), rynki aktywów są kompletne, a podatki ryczałtowe (zob. np. Baxter i King 1993). W takim modelu produkcja pozostaje funkcją kapitału i pracy. W krótkim okresie jedynym źródłem wzrostu produkcji może być tylko wzrost zatrudnienia, lecz jednocześnie przy założeniu pełnego zatrudnienia taki wzrost staje się niemożliwy. Wzrost wydatków rządowych – przeznaczanych, na przykład, na zakup sprzętu wojskowego – spowoduje w tych warunkach jedynie przesunięcie części pracowników z innych sektorów do sektora wytwarzającego sprzęt wojskowy. Ogólna wartość produkcji krajowej zwiększy się nieznacznie lub w ogóle nie ulegnie zmianie. Dlatego też w modelu neoklasycznym wartość mnożnika jest o wiele niższa niż w modelu IS-LM. Co więcej, konsumpcja spada, ponieważ zwiększenie podatków w celu sfinansowania wydatków pociąga za sobą ujemny efekt majątkowy (*wealth effect*), który obniża bieżące wydatki gospodarstw domowych (Barro i Redlick 1992, Ramey i Shapiro 1998, Ramey 2008).

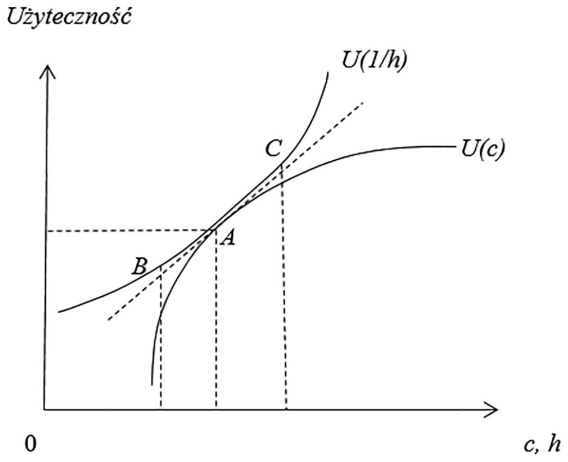
Rozpatrzmy teraz prosty neoklasyczny model równowagi ogólnej (Hall 2009, Woodford 2010). Załóżmy, że ludzie maksymalizują użyteczność, która jest zdefiniowana jako różnica pomiędzy wartością konsumpcji c a nakładem pracy (mierzonym w godzinach) h w kolejnych okresach t. Wówczas funkcję indywidualnych preferencji reprezentatywnego gospodarstwa domowego można przedstawić w postaci:

$$\left[\frac{c_t^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}}\right]-\gamma\left[\frac{h_t^{1+\frac{1}{\psi}}}{\left(1+\frac{1}{\psi}\right)}\right]=max, \tag{3.37}$$

gdzie σ mierzy międzyokresową elastyczność substytucji konsumpcji, ψ wyraża elastyczność podaży pracy względem użyteczności, a t jest indeksem okresu. Problem optymalizacji przedstawiono na rysunku 3.4.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 3.4. Równowaga na rynku pracy z perspektywy użyteczności konsumpcji



Źródło: opracowanie własne.

Ludzie zwiększają podaż pracy aż do momentu, w którym krańcowa użyteczność konsumpcji wyrówna się z krańcowym „kosztem" (*disutility*) z tytułu wzrostu nakładu pracy. Jest to punkt *A*, gdzie funkcje $U(c)$ i $U(h)$ mają jednakowe nachylenie. Gdy podaż pracy staje się niższa, na przykład odpowiada punktowi *B*, ludziom opłaca się więcej pracować, ponieważ krańcowa użyteczność z dodatkowej konsumpcji jest wyższa niż krańcowy „koszt" pracy (w punkcie *B* nachylenie krzywej $U(c)$ jest większe niż nachylenie krzywej $U(h)$). Gdy z kolei podaż pracy jest wyższa niż w punkcie *A*, na przykład odpowiada punktowi *C*, ludziom opłaca się obniżyć podaż pracy ze względu na to, że krańcowa użyteczność z dodatkowej konsumpcji pozostaje niższa niż dodatkowy „koszt" pracy (w punkcie *C* nachylenie krzywej $U(c)$ jest mniejsze niż nachylenie krzywej $U(h)$).

Jeśli przyjmiemy, że wartość konsumpcji mierzona jest ceną *p*, a wartość pracy płacą jednostkową *w*, to optymalna podaż pracy będzie określona przez punkt zrównania się relacji krańcowych użyteczności konsumpcji i „kosztu" pracy z relacją ceny i płacy:

$$\frac{\frac{\partial U}{\partial c}}{\frac{\partial U}{\partial h}} = \frac{p}{w}. \tag{3.38}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Normalizując – dla uproszczenia – wielkość p na poziomie równym jedności, i uwzględniając fakt, że parametry σ i ψ mierzą nachylenie funkcji użyteczności konsumpcji $U(c)$ oraz funkcji użyteczności pracy $U(1/h)$, możemy uznać, że warunkiem optymalności jest (dla każdego t):

$$wc^{-\frac{1}{\sigma}} = \gamma h^{\frac{1}{\psi}}. \tag{3.39}$$

Załóżmy obecnie, że proces produkcji opisuje funkcja Cobba–Douglasa o postaci:

$$y = Ak^{(1-\alpha)}h^{\alpha} \tag{3.40}$$

i przyjmijmy, że kapitał k jest w krótkim okresie stały i równy jedności, podczas gdy A pozostaje pewną stałą. Produkcja jest więc wyłącznie funkcją podaży pracy h. W warunkach pełnego zatrudnieniu wielkość h może wzrosnąć tylko w przypadku dobrowolnego zwiększenia liczby godzin pracy przez poszczególnych pracowników. Zauważmy jednak, że wzrost h przy danym zasobie kapitału sprawia – zgodnie z równaniem (3.40) – że jednostkowa płaca realna w musi spadać. Jednak wzrost h przy spadającej płacy w wymaga – na podstawie (3.39) – aby $c^{-\frac{1}{\sigma}}$ wzrosło proporcjonalnie bardziej niż wynosi spadek w, co z kolei oznacza, że c musi spadać. Wynika z tego, że egzogeniczny wzrost zagregowanego popytu i zatrudnienia musi ograniczać konsumpcję. Ten paradoksalny rezultat wynika oczywiście wprost z neoklasycznego założenia, że kanałem transmisji impulsu w postaci wzrostu wydatków rządowych na poziom konsumpcji jest efekt majątkowy. Wzrost wydatków rządowych oznacza wzrost podatków w przyszłości, czyli spadek zaktualizowanej wartości przyszłych dochodów, co przekłada się na spadek bieżącej konsumpcji (Galí, Lopez-Salida i Vallés 2007; Monacelli i Perotti 2008)[50].

Płaca realna w jest równa krańcowemu produktowi pracy:

$$w = \alpha h^{\alpha-1}. \tag{3.41}$$

Produkcja dzieli się na konsumpcję c i zakupy rządowe g:

[50] Alternatywnym wyjaśnieniem paradoksu spadającej konsumpcji i płacy realnej jest teoria realnego cyklu koniunkturalnego, w ramach której wzrost pracy realnej może być rezultatem stosowania innowacji technologicznych wpływających na wzrost wydajności pracy.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$y = c + g. \tag{3.42}$$

Łącząc warunek optymalności (3.33) oraz równania (3.41) i (3.42), otrzymujemy zmodyfikowaną postać równania równowagi ogólnej:

$$(y-g)^{-\frac{1}{\sigma}} = \frac{\gamma}{\alpha} y^{\frac{1+\frac{1}{\psi}}{\alpha}-1}. \tag{3.43}$$

Normalizując wielkość y na poziomie równym jedności dla określonego g, upraszczamy równanie (3.43) do postaci:

$$(y-g)^{-\frac{1}{\sigma}} = \frac{\gamma}{\alpha}. \tag{3.44}$$

Różniczkując całkowicie równanie (3.44) oraz uwzględniając, że $(1-g)^{-1/\sigma} \cong$ $\cong \left[1-\left(-\frac{1}{\sigma}\right)g\right]$, i przekształcając je, otrzymujemy wzór na mnożnik wydatków rządowych:

$$\frac{dy}{dg} = \frac{\alpha}{\alpha + \sigma(1-g)\left(1-\alpha+\frac{1}{\psi}\right)}. \tag{3.45}$$

Zauważmy, że mianownik w równaniu (3.45) jest zawsze większy od α, ponieważ wyrażenie $\left[\sigma(1-g)\left(1-\alpha+\frac{1}{\psi}\right)\right]$ pozostaje zawsze większe od zera dla $0 < g < 1$ (gdyż zgodnie z założeniami $\psi > 0$, a $\alpha < 1$). Oznacza to, że mnożnik (3.45) jest zawsze mniejszy od jedności. Ten rezultat – zasadniczo odmienny niż w modelu keynesowskim – wynika stąd, że w modelu neoklasycznym wydatki rządowe w znacznym stopniu wypierają wydatki prywatne, co zasadniczo ogranicza wielkość mnożnika. Równanie (3.45) pokazuje, że mnożnik wydatków będzie generalnie tym większy, im wyższa jest elastyczność podaży pracy ψ oraz im mniejsza elastyczność produkcji względem podaży pracy α.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.9.3. Mnożnik w modelach neokeynesowskich

W modelach neokeynesowskich przyjmuje się, że ceny i płace dostosowują się do zmian popytu, ale nie w pełni i z opóźnieniem – zakłada się więc częściową sztywność cen i płac. Z drugiej strony – w porównaniu z wersją czysto keynesowską – uwzględnia się oczekiwania, optymalizację wielookresową konsumpcji (funkcje preferencji) oraz mikroekonomiczne podstawy kształtowania się zachowań firm (zob. np. Galí, Lopez-Salida i Vallés 2007; Hall 2009).

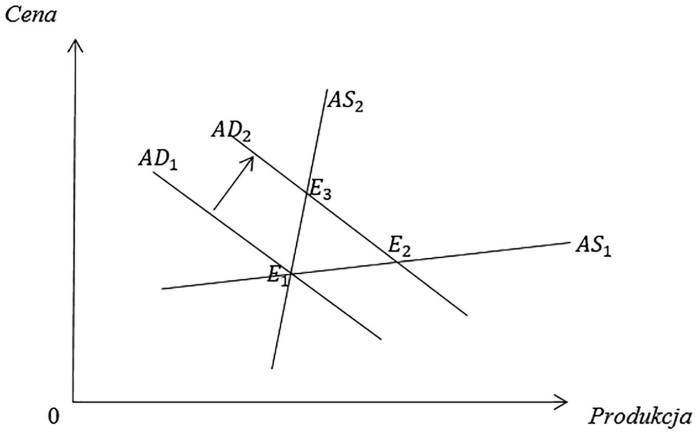
Działanie mechanizmu mnożnika w warunkach ograniczonej sztywności cen można zilustrować, wykorzystując prosty model AS-AD, opisany w rozdziale 2 i przywołany ponownie w celu opracowania rysunku 3.5. Założenie o względnej sztywności cen oznacza, że krzywa podaży *AS* jest stosunkowo płaska, choć niekonieczne pozioma (AS_1). Wzrost wydatków rządowych powoduje przesunięcie krzywej popytu AD_1 w prawo i w górę, do położenia AD_2. W takiej sytuacji podaż silnie reaguje na zmiany popytu, przyrosty produkcji są duże (równowaga przesuwa się z E_1 do E_2), a wielkości mnożnika pozostają odpowiednio wysokie. W nowym punkcie równowagi E_2 poziom dochodu *Y* jest wyższy, natomiast związany z tym wzrost cen pozostaje niewielki. Jak łatwo zauważyć, skala zmian produkcji i cen zależy od względnego nachylenia krzywych *AD* i *AS* względem siebie. Obserwacje empiryczne sugerują, że krzywa *AS* w krótkim okresie jest stosunkowo płaska, co oznacza, że w krótkim okresie ceny istotnie są względnie sztywne. W dłuższym okresie ceny zaczynają się jednak dostosowywać do popytu i nachylenie krzywej *AS* staje się bardziej strome.

Z kolei wzrost wydatków rządowych przy braku akomodacyjnej polityki banku centralnego przekłada się na „wypychanie" inwestycji prywatnych, co oznacza, że położenie krzywej popytu *AD* zasadniczo nie ulega zmianie. Popyt łączny nie rośnie, a zatem gospodarka pozostaje w punkcie E_1. Jak widać, czynnikiem zwiększającym skuteczność polityki fiskalnej staje się łagodzenie polityki pieniężnej, czyli akomodacja pieniężna. Oznacza to, że efekty ekspansywnej polityki fiskalnej okazują się silniejsze, gdy są wsparte swobodniejszą polityką pieniężną banku centralnego. Przy braku krajowej akomodacji pieniężnej jej substytutem może być transgraniczna swoboda przepływu kapitału – wzrost stóp procentowych wywołany wzrostem wydatków rządowych jest wówczas ograniczony przez napływ kapitału zagranicznego. Wreszcie w przypadku giętkich cen i sztywnej podaży (założenia odwrotne do keynesowskich) krzywa podaży *AS* jest niemal

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pionowa (AS_2), a zmiany popytu globalnego wywołują przede wszystkim zmiany cen (przejście z E_1 do E_3).

Rysunek 3.5. Polityka fiskalna w warunkach elastycznej i nieelastycznej podaży



Źródło: opracowanie własne.

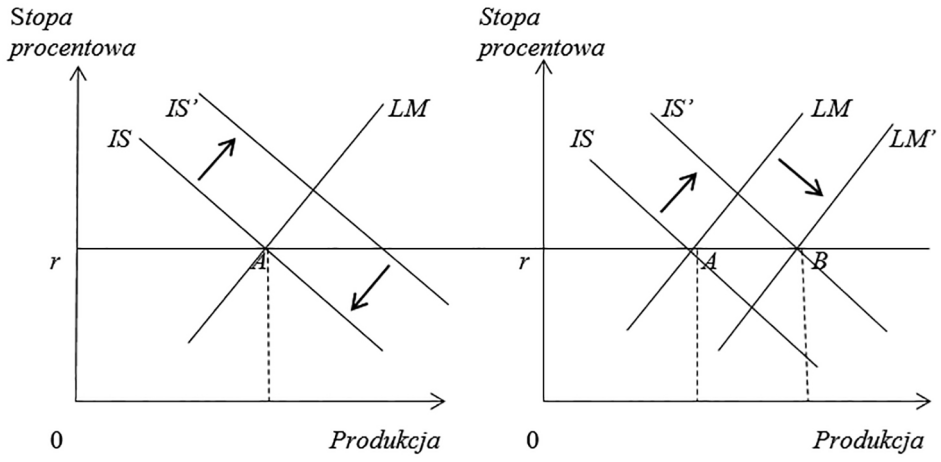
3.9.4. Polityka fiskalna w gospodarce otwartej (na podstawie modelu Mundella-Fleminga)

W podstawowej wersji model IS-LM opisuje gospodarkę zamkniętą. W warunkach gospodarki otwartej skuteczność polityki fiskalnej w pobudzaniu gospodarki zależy dodatkowo od przyjętego systemu kursowego i od stopnia transgranicznej mobilności kapitałów. Taką otwartą gospodarkę opisuje model Mundella-Fleminga (zob. też rozdział 2). Początkowo wzrost wydatków rządowych w gospodarce otwartej przekłada się na wzrost zagregowanego popytu, w tym popytu na import, co wpływa na pogorszenie salda bilansu obrotów bieżących i tendencję do deprecjacji waluty krajowej. Gdy obowiązuje system kursu płynnego i kapitał jest mobilny, wzrost stóp procentowych wywołany wzrostem wydatków rządowych przyciąga kapitał z zagranicy, co przeciwdziała deprecjacji waluty krajowej, a przy dużej skali napływu kapitału prowadzi do aprecjacji kursu. Związany z tym spadek eksportu netto zmniejsza wielkość mnożnika wydatków i ogranicza skuteczność polityki fiskalnej. W skrajnym przypadku (gdy kapitał jest doskonale mobilny), aprecjacja kursu staje się na tyle duża, że jej efekt całkowicie neutralizuje skutki początkowego impulsu fiskalnego. Mnożnik wydatków w takich warunkach równa się zeru.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Z kolei gdy obowiązuje system stałego kursu walutowego i kapitał jest mobilny w skali międzynarodowej, wzrost stóp procentowych wywołany wzrostem wydatków rządowych również przekłada się na napływ kapitału z zagranicy. Tym razem bank centralny skupuje jednak waluty obce w zamian za walutę krajową, broniąc ustalonego poziomu kursu. Spowodowany tym wzrost podaży pieniądza krajowego zwiększa popyt krajowy. Mnożnik wydatków rządowych jest w tej sytuacji dodatni.

Rysunek 3.6. Wzrost wydatków rządowych i wpływ na produkcję, zarówno przy płynnym, jak i stałym kursie walutowym, w warunkach pełnej mobilności kapitału



Źródło: Bénassy-Quéré i in. (2011).

Efekty ekspansywnej polityki fiskalnej w różnych systemach kursowych przedstawia rysunek 3.6. Lewa część rysunku odnosi się do systemu kursu płynnego. Załóżmy, że bank centralny utrzymuje stałą podaż pieniądza, a doskonała mobilność kapitału sprawia, że krajowa stopa procentowa nie może trwale odchylać się od światowej stopy procentowej *r*. Wzrost wydatków rządowych przesuwa krzywą *IS* w prawo, ponieważ rośnie zagregowany popyt przy danej stopie procentowej. Równolegle wzrasta popyt transakcyjny na pieniądz, co wywołuje presję na wzrost stóp procentowych i, w konsekwencji, napływ kapitału zagranicznego. Kurs waluty krajowej ulega aprecjacji, popyt krajowy przenosi się na import, a produkcja krajowa w rezultacie spada. W efekcie krzywa *IS* przesuwa się z powrotem w lewo i gospodarka powraca do początkowego punktu *A*. Ekspansja jest w tych warunkach nieskuteczna.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Prawa część rysunku odnosi się do systemu stałego kursu. Przy tych samych założeniach, wzrost wydatków rządowych przesuwa krzywą *IS* w prawo, równolegle rośnie popyt transakcyjny na pieniądz, co powoduje presję na wzrost stóp procentowych i napływ kapitału zagranicznego. W warunkach stałego kursu bank centralny skupuje waluty obce w zamian za walutę krajową, a to przekłada się na wzrost podaży pieniądza (akomodacja pieniężna). W rezultacie krzywa *LM* przesuwa się w prawo i gospodarka przemieszcza się do nowego punktu równowagi *B*. Ekspansja fiskalna okazuje się w tych warunkach skuteczna.

Uchylenie założenia o transgranicznej mobilności kapitału powoduje, że wpływ systemu kursowego na wielkość mnożnika wydatków ulega odwróceniu. W warunkach braku mobilności kapitału to system płynnego kursu zapewnia większą skuteczność poliki fiskalnej, ponieważ napływ kapitału nie przeciwdziała deprecjacji kursu na skutek pogorszenia bilansu obrotów bieżących. Z kolei w systemie kursu stałego pogorszenie salda bilansu bieżącego wywołuje odpływ rezerw i zmniejszenie podaży pieniądza, co ogranicza popyt krajowy.

3.9.5. Zwiększać wydatki czy zmniejszać podatki?

Gdy rząd chce stymulować gospodarkę w okresach słabszej koniunktury, ma generalnie dwie opcje do wyboru – może zwiększyć wydatki lub obniżyć podatki. W obu przypadkach bezpośrednim skutkiem polityki stabilizacyjnej dla budżetu państwa będzie spadek salda budżetowego – najczęściej wzrost deficytu – który trzeba będzie dodatkowo sfinansować. Skutki, jakie przyniesie to dla gospodarki, będą jednak różne. Panuje przekonanie, że wzrost wydatków rządowych ma generalnie silniejszy efekt na zagregowany popyt niż obniżka podatków dla firm i gospodarstw domowych, ponieważ w przypadku tych ostatnich występuje dodatnia skłonność do oszczędzania. Oznacza to, że z każdego złotego pozostawionego w kieszeniach podatników tylko część zostanie przeznaczona na bieżące zakupy, a pozostała część nie zostanie wydana, a więc będzie zaoszczędzona[51]. Oszczędności te będą generalnie wyższe w okresach zwiększonej niepewności, związanej na przykład z wysokim poziomem bezrobocia lub wysokim poziomem zadłużenia. Z kolei w przypadku wydatków rządowych na zakup dóbr i usług łączna wartość

[51] Założenie o istnieniu dodatniej skłonności do oszczędzania nie odnosi się do najuboższych gospodarstw domowych, które wydają 100% dochodu (ich wydatki są ograniczone przez dostępność płynnych środków – *liquidity-constrained households*).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

tych kosztów zwiększa popyt globalny w gospodarce (o ile nie występuje efekt wypychania). W konsekwencji wartości fiskalnych mnożników wzrostu wydatków są generalnie wyższe niż wartości fiskalnych mnożników w przypadku obniżek podatków (zob. Blanchard i Perotti 2002).

Powyższe rozumowanie opiera się na kluczowym keynesowskim założeniu mówiącym o tym, że konsumpcja jest funkcją bieżącego dochodu, a krańcowa skłonność do konsumpcji jest mniejsza od jedności. W pewnych sytuacjach zależności te mogą jednak ulec zaburzeniu. W szczególności, gdy uwzględni się efekty po stronie podaży (*supply side effects*), wpływ obniżki podatków może być silniejszy. Będzie tak wówczas, gdy skutkiem obniżenia obciążeń podatkowych stanie się zwiększenie potencjalnego tempa wzrostu PKB, co może spowodować dodatkowy wzrost konsumpcji. Ponadto, wzrost potencjalnego poziomu PKB zmniejsza presję na wzrost stóp procentowych i przeciwdziała efektowi wypychania, co ogranicza lub eliminuje ewentualne negatywne skutki deficytu budżetowego (Alesina i Ardagna 2010).

W świetle powyższych uwag trudno jest jednoznacznie wyrokować o wyższości jednej opcji polityki nad drugą. To, czy w danej gospodarce skuteczniejsze okaże się zwiększanie wydatków rządowych czy obniżanie podatków, zależy od wielu czynników, w tym od fazy cyklu koniunkturalnego, stopnia otwartości na wymianę zagraniczną oraz polityki pieniężnej banku centralnego[52].

3.9.6. Równoważność rikardiańska – jak finansować deficyt: z długu czy z podatków?

Skuteczność działania mechanizmu mnożnika jest także kwestionowana na gruncie tzw. równoważności rikardiańskiej (*Ricardian equivalence, Barro-Ricardo equivalence*)[53]. Jej istotę stanowi następujące twierdzenie. Gdy rząd zwiększa wydatki publiczne finansowane z długu, aby stymulować gospodarkę, racjonalne

[52] Krótki przegląd rozmaitych czynników wpływających na efektywność różnych instrumentów polityki fiskalnej znaleźć można w pracy Spilimbergo i in. 2008.

[53] Nazwa pochodzi od nazwiska angielskiego ekonomisty Davida Ricardo, który jako pierwszy (w 1817 roku) sformułował hipotezę, zgodnie z którą wzrost wydatków publicznych finansowany długiem wywołuje pojawienie się oczekiwań wzrostu podatków w przyszłości (w celu spłaty długu), co skłoni racjonalnie zachowujące się gospodarstwa domowe do ograniczenia bieżących wydatków i zwiększenia oszczędności. W takiej sytuacji stymulacja fiskalna jest całkowicie nieskuteczna. Gdy zachodzi równoważność rikardiańska, nie ma różnicy między finansowaniem wzrostu wydatków z długu i z podatków. W późniejszych pracach Ricardo wycofał się z tej propozycji. Autorem współczesnej wersji twierdzenia jest Robert Barro (1974).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gospodarstwa domowe z dostatecznie długim (nieskończonym) horyzontem planowania reagują na to ograniczeniem swoich wydatków (konsumpcyjnych i inwestycyjnych), ponieważ przewidują, że zadłużenie powstałe w wyniku deficytu będzie w przyszłości musiało być spłacone z nowych podatków. Oczekiwany wzrost podatków nakazuje im zatem ograniczyć bieżące wydatki i zwiększyć oszczędności. Oznacza to, że wzrost wydatków rządowych jest kompensowany równoległym spadkiem wydatków prywatnych. Polityka fiskalna staje się w tych warunkach nieskuteczna, ponieważ popyt globalny w gospodarce nie ulega zmianie. Równoważność rikardiańska stawia więc znak równości pomiędzy długiem a odłożonymi w czasie podatkami. Oznacza to, że odpowiedź na pytanie o to, czy rząd powinien finansować dodatkowe wydatki z długu czy ze zwiększonych podatków, jest nieistotna, ponieważ w obu przypadkach nastąpi spadek popytu prywatnego. W skrócie – finansowanie wydatków z długu jest równoznaczne z finansowaniem ich z podatków – uczestnicy rynku zdają sobie sprawę, że spłata długu wymaga podniesienia podatków w przyszłości i z tego powodu ograniczają swoje bieżące wydatki[54] (zob. ramka 3.10).

Ramka 3.10. Równoważność rikardiańska

Załóżmy, że typowe gospodarstwo domowe otrzymuje dochody Y, może zaciągać i udzielać pożyczki po stopie procentowej r, a wydatki konsumpcyjne wynoszą C. Ponadto, gospodarstwo domowe optymalizuje poziom konsumpcji w nieskończonym horyzoncie czasu. Po przyjęciu tych założeń międzyokresowe ograniczenie budżetowe wymaga, aby zdyskontowana suma wydatków konsumpcyjnych danego gospodarstwa w kolejnych okresach równała się zdyskontowanej sumie dochodów[55]:

$$\sum_{0}^{\infty} \frac{C_{t+i}}{(1+r)^{i}} = \sum_{0}^{\infty} \frac{Y_{t+i}}{(1+r)^{i}}. \tag{3.46}$$

Gospodarstwa domowe maksymalizują międzyokresową funkcję użyteczności pod warunkiem spełnienia ograniczenia budżetowego:

[54] Innymi słowy, wpływ na poziom aktywności gospodarczej ma tylko ścieżka wydatków państwowych, a nie ścieżka finansujących je podatków (Romer 2000, s. 86).

[55] Oszczędności są w tym ujęciu jedynie wielkością bilansującą międzyokresowe różnice w poziomach konsumpcji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$U(t) = \sum_{0}^{\infty} \frac{u(C_{t+i})}{(1+\rho)^i}, \tag{3.47}$$

gdzie ρ jest indywidualną stopą preferencji czasowej, a $U(t)$ pozostaje funkcją wklęsłą o malejącej krańcowej użyteczności konsumpcji.

W celu znalezienia optymalnych wartości $C(t)$ budujemy funkcję Lagrange'a o postaci:

$$Z = \sum_{0}^{\infty} \frac{u(C_{t+i})}{(1+\rho)^i} - \lambda[\sum_{0}^{\infty} \frac{C_{t+i}}{(1+r)^i} - \sum_{0}^{\infty} \frac{Y_{t+i}}{(1+r)^i}], \tag{3.48}$$

gdzie λ jest mnożnikiem Lagrange'a. Różniczkując funkcję Z i przyrównując pochodną do zera, znajdujemy warunek pierwszego rzędu zapewniający maksymalizację funkcji U:

$$\frac{\partial Z}{\partial C_{t+i}} = \frac{u'(C_{t+i})}{(1+\rho)^i} - \lambda \frac{1}{(1+r)^i} = 0,$$

czyli:

$$u'(C_{t+i}) = \lambda \left(\frac{1+\rho}{1+r} \right)^i. \tag{3.49}$$

Powyższe równanie pozwala określić optymalną ścieżkę konsumpcji w czasie. Pamiętając, że $u'(\cdot)$ jest malejącą funkcją poziomu konsumpcji, gdy $r = \rho$, konsumpcja jest stała w czasie, gdy $r > \rho$, lub $r < \rho$, konsumpcja odpowiednio rośnie lub maleje. Przy tej specyfikacji, a jednocześnie w przeciwieństwie do założeń modelu IS-LM, konsumpcja jest niezależna od bieżącego dochodu Y, a zatem przejściowe zmiany dochodu nie wpływają na poziom konsumpcji.

Wprowadźmy obecnie do modelu rząd, który pobiera podatki T i przeznacza je na wydatki G, przy czym $G = T$. Jeśli wydatki rządowe nie mają wpływu na funkcję użyteczności gospodarstw domowych (tak jest, gdy rząd nie dostarcza dóbr publicznych), międzyokresowe ograniczenie budżetowe typowego gospodarstwa domowego przyjmuje następującą postać:

$$\sum_{0}^{\infty} \frac{Y_{t+i} - T_{t+i}}{(1+r)^i} = \sum_{0}^{\infty} \frac{C_{t+i}}{(1+r)^i}. \tag{3.50}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Po podstawieniu tego równania do funkcji Lagrange'a widzimy, że warunek pierwszego rzędu nie zmienia się, a więc, że wprowadzenie do modelu wydatków rządowych finansowanych podatkami nie wpływa na ścieżkę konsumpcji w czasie, choć poziom konsumpcji w każdym okresie jest niższy.

Załóżmy obecnie, że rząd, utrzymując pierwotny poziom wydatków, obniża podatki w okresie $t = 0$ o kwotę B i w zamian emituje obligacje na tę samą kwotę, oprocentowane według stopy r, na okres N lat. Obligacje te skupują gospodarstwa domowe, przy czym zakłada się, że zarówno kapitał, jak i odsetki od długu są spłacane z podwyższonych podatków. W funkcji użyteczności gospodarstw domowych nic się nie zmienia. Przeznaczają one kwotę B z bieżącego dochodu na zakup obligacji, ale zarazem o tę samą kwotę spadają ich obciążenia podatkowe. Ponadto, w okresie aż do $t = N$ otrzymują corocznie dochód w kwocie rB z posiadanych obligacji, lecz zarazem płacą podatki tej samej wysokości. Wreszcie w roku $t = N$ gospodarstwa domowe otrzymują zwrot kapitału B i jednocześnie płacą podatki zwiększone o kwotę B.

Na tym właśnie polega równoważność rikardiańska – dług zaciągnięty obecnie równy jest przyszłym podatkom, a to, czy zwiększone wydatki rządowe są finansowane z długu, czy z podatków nie wpływa w żaden sposób na ścieżkę konsumpcji w czasie.

Równoważność rikardiańska – jeśli istotnie jest prawidłowym opisem rzeczywistości – miałaby daleko idące konsekwencje dla polityki gospodarczej. Oznaczałaby mianowicie, że jakiekolwiek próby pobudzania gospodarki poprzez wydatki finansowane z deficytu są z góry skazane na niepowodzenie. Polityka stabilizowania koniunktury traci zatem sens, podczas gdy polityka gospodarcza staje się zupełnie bezradna wobec fluktuacji wywołanych przebiegiem cyklu koniunkturalnego. W praktyce – i na szczęście dla zwolenników polityki stabilizacyjnej – badania empiryczne generalnie nie potwierdzają jednak istnienia równoważności rikardiańskiej, a przynajmniej w statystycznie istotnej postaci. Gdyby było bowiem inaczej, występowałaby pełna ujemna korelacja pomiędzy oszczędnościami publicznymi a oszczędnościami prywatnymi – wzrostowi deficytu publicznego towarzyszyłby wzrost oszczędności prywatnych. Tymczasem, choć dane empiryczne potwierdzają istnienie ujemnej korelacji między tymi dwoma zmiennymi, wartość współczynnika korelacji jest wyraźnie niższa (co do modułu) od jedności.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 3.7 prezentuje wartości współczynnika korelacji pomiędzy zmianami udziału oszczędności w dochodzie rozporządzalnym gospodarstw domowych a saldem budżetowym sektora publicznego w wybranych krajach w okresie 1995–2011. Z przedstawionych danych wynika, że między tymi dwoma zmiennymi istotnie występuje ujemna korelacja, lecz brakuje jednocześnie pełnej symetrii i między danymi z poszczególnych państw występują duże róznice. Ogólnie rzecz biorąc, korelacja jest silniejsza w większych krajach o mniejszym uzależnieniu od wymiany międzynarodowej (Hiszpania, Francja, Włochy), a znacznie słabsza w przypadku małych otwartych gospodarek (Belgia, Holandia, a także Polska). Można więc stwierdzić, że efekt rikardiański występuje w ograniczonym zakresie[56].

Tabela 3.7. Czy równoważność rikardiańska istnieje? Korelacja między zmianami udziału oszczędności w dochodzie rozporządzalnym gospodarstw domowych a saldem budżetowym sektora publicznego (*general government*) w wybranych krajach w latach 1995–2011

Kraj	Współczynnik korelacji
Austria	–0,3588
Belgia	–0,1450
Francja	–0,5803
Hiszpania	–0,6885
Holandia	–0,1450
Niemcy	–0,2700
Polska[a]	–0,0110
Szwecja	–0,1655
Wielka Brytania[a]	–0,1917
Włochy	–0,4783
Japonia[a]	–0,2703
Stany Zjednoczone	–0,3661

a) Lata 1995–2010.

Źródło: obliczenia własne na podstawie bazy danych AMECO.

Istnieje kilka ważnych przyczyn, które sprawiają, że twierdzenie o równoważności finansowania wydatków rządowych z podatków i z długu nie sprawdza się

[56] Również badanie przeprowadzone przez de Melo i in. (2004) na próbie 21 krajów członkowskich OECD wykazało, że wzrost prywatnych oszczędności kompensuje jedynie między jedną trzecią a połową spadku oszczędności publicznych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w praktyce, a odchylenia rzeczywistych zmian aktywności gospodarczej od teoretycznych wskazań mogą być znaczne.

Przede wszystkim ludność danego kraju podlega rotacji pokoleniowej, a horyzont wybranego pokolenia jest skończony. Oznacza to, że gospodarstwa domowe, kupując obligacje, nie w pełni uwzględniają w ramach swojego horyzontu życia perspektywę podniesienia podatków w celu spłaty obligacji. Przyjmują mianowicie założenie (być może nie do końca świadomie), że część kosztów związanych ze spłatą długu poniosą następne pokolenia. W takiej sytuacji obligacje znajdujące się w posiadaniu tych gospodarstw domowych stanowią ich majątek netto i w takim zakresie, w jakim zwiększają cały majątek, nie stanowią czynnika ograniczającego wydatki gospodarstw domowych. To zastrzeżenie odnoszące się do równoważności rikardiańskiej byłoby jednak bezpodstawne, gdyby gospodarstwa domowe troszczyły się o dobrobyt własnych potomków i uwzględniały ich obciążenia podatkowe w swoich bieżących decyzjach, na przykład, przeznaczając większą część bieżącego dochodu na zapisy spadkowe. Zwolennikiem takiego właśnie poglądu był m.in. amerykański ekonomista Robert Barro, który przedstawił sformalizowaną wersję równoważności rikardiańskiej (Barro 1974)[57]. Wyznawcą odmiennego poglądu, mówiącego o tym, że gospodarstwa domowe troszczą się głownie o własny dobrobyt, a w znacznie mniejszym stopniu o przyszłe pokolenia, jest m.in. Blanchard (1985).

Druga okoliczność, którą należy wziąć pod uwagę, to ograniczona zdolność gospodarstw domowych do powiększania wydatków wobec funkcjonowania ograniczeń kredytowych (zob. np. Tobin 1980). Gdy rząd emituje obligacje i sprzedaje je gospodarstwom domowym, zaciąga w istocie dług w ich imieniu, bo to one będą musiały ostatecznie spłacić go z własnych zwiększonych podatków. Gospodarstwa domowe zareagują na to wzrostem bieżących wydatków tylko wówczas, gdy w wyniku tej operacji uznają, że rząd zaciągnął pożyczki na lepszych warunkach, niż one same mogłyby to uczynić. Innymi słowy, jeśli gospodarstwa domowe nie są w stanie obecnie zaciągać pożyczek na korzystnych warunkach, bieżąca obniżka podatków i kupno obligacji rządowych zwiększa ich majątek i skłania do podnoszenia wydatków. Lecz, podobnie jak w przypadku rotacji pokoleniowej, odpowiedź na pytanie o rzeczywistą skalę wpływu ograniczeń płynnościowych

[57] Rozstrzygnięcie tego, czy rotacja pokoleniowa wpływa na osłabienie solidarności pokoleniowej i podważa równowartość rikardiańską, pozostaje w znacznej mierze kwestią empiryczną. Poterba i Summers (1987) dowodzą, że nawet jeśli przyrost majątku gospodarstw domowych wsktek nabycia obligacji ma wpływ wyłącznie na (cały) okres życia danego pokolenia, zwiększają one swoje wydatki konsumpcyjne zaledwie o 3% kwoty przyrostu majątku.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gospodarstw domowych na zmianę stuktury ich wydatków zależy od wyników badań empirycznych.

Trzecia okoliczność, która może zaburzać równoważność rikardiańską, dotyczy tego, że obecne wydatki rządowe finansowane pożyczkami oddziałują na ogół na bieżącą i przyszłą produkcję, choćby poprzez wpływ na wydajność pracy i przyspieszenie potencjalnego tempa wzrostu PKB. Oznacza to, że obecne wydatki publiczne mogą wpływać na wzrost oczekiwanych dochodów indywidualnych w przyszłości, co jest w stanie skłaniać ludzi do bieżącego zwiększania wydatków.

Wreszcie hipoteza o równoważności rikardiańskiej zbudowana jest na założeniu doskonałej racjonalności gospodarstw domowych, które są w stanie przewidywać obecne skutki polityki wydatków rządowych i przyszłą ścieżkę własnych dochodów. Jest to założenie bardzo daleko idące, i dlatego wątpliwe. W praktycznych zachowaniach gospodarstw domowych dominuje raczej poczucie niepewności co do nadchodzących lat oraz skłonność do skracania horyzontu decyzji (Romer 2000).

3.9.7. Efekty podażowe i spór wokół polityki stabilizacyjnej

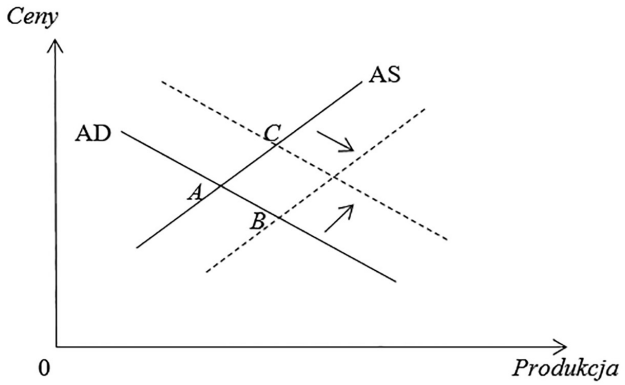
Ekonomiści toczą spór co do skuteczności polityki fiskalnej jako instrumentu pobudzania wzrostu gospodarczego. Jak wykazaliśmy wyżej, mnożnik wydatków publicznych podlega licznym ograniczeniom i nie zawsze zwiększanie wydatków publicznych przynosi pożądane skutki. Dotyczy to sytuacji, w której brakuje wolnych mocy wytwórczych albo gdy duża część wydatków przeznaczana jest na import dóbr i usług zamiast na produkty krajowe, albo gdy bank centralny podnosi stopy procentowe w reakcji na wzrost wydatków (brak akomodacji pieniężnej), a także gdy wzrost wydatków jest finansowany z długu i występuje równoważność rikardiańska. Wszystkie te okoliczności sprawiają, że ekonomiści neoklasyczni kwestionują keynesistowski pogląd o skuteczności polityki fiskalnej w stabilizowaniu gospodarki i pobudzaniu wzrostu produkcji oraz zatrudnienia.

Jak wspomniano wcześniej, chcąc stymulować gospodarkę przy pomocy polityki fiskalnej, rząd ma do wyboru dwie opcje – zwiększać wydatki lub obniżać podatki. Zarówno keynesiści, jak i neoklasycy zgadzają się, że obniżka podatków przyczynia się do przyspieszenia wzrostu gospodarczego, lecz podają różne wyjaśnienia tego związku. Neoklasycy twierdzą, że obniżenie podatków wpływa na wzrost podaży, ponieważ produkcja rośnie dzięki spadkowi kosztów, z kolei w opinii keynesistów obniżenie podatków oddziałuje na wzrost popytu ze względu na to, że zwiększa dochód rozporządzalny. Obie interpretacje zostały przedstawione na rysunku 3.7.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W ramach prostego modelu AS-AD obniżka podatków może albo powodować przesunięcie krzywej podaży *AS* w prawo i wzrost produkcji z punktu *A* do punktu *B* (pogląd neoklasyczny), albo wpływać na przesunięcie krzywej popytu *AD* w prawo i wzrost produkcji z punktu *A* do punktu *C* (pogląd neokeynesowski). W pierwszym przypadku wzrostowi podaży przy niezmienionym popycie musi towarzyszyć spadek cen, z kolei w drugim przypadku wzrostowi popytu przy niezmienionej podaży musi towarzyszyć wzrost cen.

Rysunek 3.7. Obniżka podatków: pogląd neoklasyczny i pogląd keynesistowski



Źródło: opracowanie własne.

O ile reprezentanci obu kierunków myślenia zgadzają się, że obniżka podatków jest czymś generalnie pożądanym z perspektywy pobudzania koniunktury (choć z różnych powodów), o tyle mają zasadniczo odmienne stanowisko co do skutków wzrostu wydatków publicznych. Ekonomiści neoklasyczni podkreślają, że wzrost wydatków – zwłaszcza permanentny – prowadzi do podniesienia podatków w przyszłości, a to przekłada się na obniżenie wydatków bieżących ze strony racjonalnych uczestników rynku. W efekcie wzrost wydatków rządowych jest częściowo lub całkowicie neutralizowany spadkiem wydatków prywatnych, a polityka fiskalna staje się nieskuteczna (Blanchard i in. 1991, Alesina i Perotti 1995). Kluczową rolę w funkcjoniwaniu tego mechanizmu odgrywają oczekiwania uczestników rynku, którzy spodziewają się podwyżki podatków w przyszłości, i, dążąc do wyrównywania konsumpcji w czasie, obniżają swoje bieżące wydatki.

Istotną rolę w tym sposobie myślenia odgrywa kwestia długu publicznego. Gdy wzrost wydatków rządowych jest finansowany zaciąganiem pożyczek, rośnie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zadłużenie publiczne. Wzrost zadłużenia do pewnego poziomu nie wywołuje negatywnych skutków w postaci ograniczania bieżącej konsumpcji, a polityka fiskalna osiąga w tej fazie skutki typowo keynesowskie. Lecz po przekroczeniu danego poziomu zadłużenia uczestnicy rynku zaczynają zachowywać się w sposób „rikardiański" – spada konsumpcja i tempo wzrostu gospodarczego ulega obniżeniu (Bertola i Drazen 1993, Sutherland 1997). Ekspansywna polityka fiskalna wywołuje w tych warunkach skutki niekeynesowskie (*non-keynesian*), a nawet antykeynesowskie (*anti-keynesian*).

Jak widać, spór o rolę polityki fiskalnej w stabilizowaniu gospodarki jest trudny do rozstrzygnięcia w kategoriach ogólnych. W zależności od przyjętych założeń, prócz wielu innych czynników określających stan gospodarki, polityka fiskalna może być bardziej lub mniej skuteczna. W tabeli 3.8 ukazano, jak przedstawiają się efekty ekspansywnej polityki fiskalnej w ramach różnych koncepcji analitycznych.

Tabela 3.8. Skutki ekspansywnej polityki fiskalnej w ramach różnych modeli makroekonomicznych

Model	Główne założenia	Mechanizm	Mnożnik	Skutki
Neokeynesowski	krótki okres, ceny sztywne, podaż elastyczna (wolne moce wytwórcze)	zwiększenie popytu powoduje wzrost produkcji i zatrudnienia	M > 0	polityka fiskalna jest skuteczna (skutki keynesowskie)
Równoważność rikardiańska	międzyokresowe ograniczenie budżetowe, oczekiwania racjonalne	„wypieranie" wydatków prywatnych przez wydatki publiczne	M = 0	polityka fiskalna jest nieskuteczna (skutki niekeynesowskie)
Neoklasyczny	podatki przekładają się na spadek efektywności (efekty podażowe), skutki prowadzonej polityki pozostają zależne od tego, czy ekspansja fiskalna polega na wzroście wydatków, czy na obniżaniu podatków	wzrost wydatków publicznych obniża wydatki prywatne (efekt wypierania) i obniża potencjalny poziom PKB	M ≤ 0	polityka fiskalna jest nieskuteczna (skutki niekeynesowskie lub antykeynesowskie)
		spadek podatków zwiększa potencjalny poziom PKB	M > 0	polityka fiskalna jest skuteczna (skutki keynesowskie)
Mieszany (z efektami progowymi)	krótki okres, ceny stałe, podaż elastyczna (wolne moce wytwórcze), wydatki prywatne zależą od poziomu zadłużenia	poniżej progu zadłużenia następuje wzrost produkcji i zatrudnienia	M > 0	polityka fiskalna jest skuteczna (skutki keynesowskie)
		powyżej progu zadłużenia dochodzi do spadku wydatków prywatnych	M ≤ 0	polityka fiskalna jest nieskuteczna (skutki niekeynesowskie lub antykeynesowskie)

Źródło: opracowanie własne na podstawie Bénassy-Quéré i in. 2011, s. 213.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.9.8. Mnożnik wydatków rządowych – wyniki badań empirycznych

Rozbieżność poglądów na temat teoretycznych podstaw działania mnożnika wydatków sprawia, że pytanie o skuteczność polityki fiskalnej w pobudzaniu gospodarki staje się w dużym stopniu pytaniem empirycznym. Niestety, liczne badania nie dostarczyły w tej kwestii jednoznacznej odpowiedzi. Szczególnie starannie badano pod tym względem gospodarkę amerykańską w okresie od II wojny światowej do czasu Wielkiej Recesji lat 2008–2010.

Christiano i Eichenbaum (1992), Aiyagari, Christiano i Eichenbaum (1992) oraz Baxter i King (1993), wykorzystując modele neoklasyczne, wykazali, że wstrząs w postaci wzrostu wydatków rządowych finansowany w całości zwiększeniem podatków ryczałtowych wywołuje umiarkowany wzrost produkcji (mnożniki fiskalne w przypadku dochodu są dodatnie, lecz mniejsze od jedności), zarazem powoduje jednak spadek konsumpcji i płac realnych (mnożniki fiskalne w przypadku konsumpcji są ujemne), co jest zgodne z neoklasycznymi założeniami. Burnside, Eichenbaum i Fisher (2003) zastosowali dwie wersje modelu neoklasycznego – standardową i rozbudowaną, uwzględniającą koszty dostosowań inwestycji i tworzenie przyzwyczajeń (*habit formation*), wykazując, że wzrost wydatków rządowych o jednostkę zwiększa PKB w obu wersjach modelu średnio o 0,3–0,8 jednostki, lecz jednocześnie obniża płacę realną i konsumpcję. Podobne wyniki uzyskali Ramey i Shapiro (1998), Ramey (2009) i Barro i Redlick (2009), wykorzystując różne warianty neoklasycznego modelu makroekonomicznego. Ci ostatni wykazali również, że podwyżki podatków mają wyraźnie negatywny wpływ na wzrost PKB i konsumpcję.

Blanchard i Perotti (2002) za punkt wyjścia przyjęli neokeynesowski model równowagi ogólnej z częściowymi sztywnościami cen i płac, i zastosowali technikę VAR (autoregresji wektorowej – *vector autoregression*) do szacowania mnożników fiskalnych[58]. Otrzymane wyniki wskazują na istnienie wyraźnie dodatniej korelacji między wzrostem wydatków rządowych a wzrostem PKB i konsumpcji indywidualnej oraz ujemnej korelacji między wzrostem wydatków a inwestycjami. Podobnie jak Barro i Redlick, Blanchard i Perotti wykazali, że podniesienie podatków ma

[58] Modele VAR stały się bardzo popularnym narzędziem analizy wpływu polityki fiskalnej na zmiany produkcji, konsumpcji i inwestycji. Na ograniczenia informacyjne związane z wykorzystaniem technik VAR wskazują m.in. Ramey i Shapiro (1997) oraz Edelberg, Eichenbaum i Fisher (1999).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ujemny wpływ na wzrost gospodarczy, konsumpcję i inwestycje. Porównywalne wyniki uzyskali Galí i in. (2007), stosując technikę VAR do danych z okresu 1960–2003. Z kolei Monacelli i Perotti (2008) dowiedli, że wyniki zastosowań modeli neoklasycznych mogą być spójne z modelami neokeynesowskimi wykorzystującymi techniki VAR i SVAR pod warunkiem, że uwzględni się zmienność marż w cenach produktów. W ramach tego podejścia wzrost wydatków rządowych nie wywołuje spadku płacy realnej, a jedynie spadek marży, rozumianej jako różnica między ceną i kosztami produkcji, co pozwala na wzrost konsumpcji indywidualnej. Oszacowane przez nich wartości mnożnika są bliskie jedności.

Robert Hall (2009) wykorzystał zwykłą regresję w celu oszacowania wielkości reakcji PKB i konsumpcji indywidualnej na zmian wydatków rządu USA, przeznaczonych na zakupy dóbr i usług w latach 1930–2008. Otrzymana wielkość mnożnika fiskalnego w przypadku PKB wynosi 0,55 dla całego okresu, przy czym wielkości mnożnika w odniesieniu do okresów zwiększonych wydatków zbrojeniowych są znacznie wyższe. Wyniki uzyskane przez Halla pozostają zgodnie z wnioskami Monacellego i Perottiego (2008). Uwzględnienie w modelach neoklasycznych zmiennych marż oraz wysokiej elastyczności podaży pracy pozwala na lepsze oszacowanie mnożników, zbliżone do tych wyliczanych na podstawie modeli neokeynesowskich oraz przy pomocy technik VAR. Jednocześnie oszacowane przez Halla wartości mnożników w przypadku konsumpcji indywidualnej są ujemne i w zasadzie statystycznie nieistotne. Podobne podejście zastosowali Alesina i Ardagna (2010), szacując mnożniki fiskalne na podstawie danych panelowych dla krajów OECD za okres 1970–2007. Lecz uzyskane przez nich wyniki wskazują raczej na negatywny wpływ stymulacji fiskalnej na wzrost gospodarczy i konsumpcję. Z kolei Christiano, Eichenbaum i Rebelo (2009) wykorzystują neokeynesowski model równowagi ogólnej i techniki VAR w celu wykazania, że mnożniki fiskalne są szczególnie wysokie – znacznie wyższe od jedności – gdy stopy procentowe pozostają na poziomie bliskim zeru (co jest też zgodne z wynikami badań przeprowadzonych przez innych autorów – zob. m.in. Hall 2009 oraz Cogan i in. 2009).

Podsumowując wyniki tych analiz, można stwierdzić, że większość z nich potwierdza umiarkowany dodatni wpływ wzrostu wydatków na produkcję (mnożniki poniżej jedności, średnio w granicach 0,3–0,7), choć skala tego wpływu pozostaje różna w rozmaitych okolicznościach i przy odmiennych założeniach. Ogólnie rzecz biorąc, mnożniki fiskalne są wyższe (bliskie jedności lub większe od jedności) w przypadku zmian wydatków niż podatków, a także gdy dotyczą wydatków

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

na obronę narodową oraz przy niskich stopach procentowych. Z kolei wpływ impulsów fiskalnych na konsumpcję i płace realne nie jest jednoznaczny. Wyniki uzyskane przez część ekonomistów, takich jak Fatás i Mihov (2001), Blanchard i Perotti (2002), Galí, López-Salido i Vallés (2007) oraz Perotti (2007), wskazują, że wzrost wydatków rządowych przekłada się na wzrost konsumpcji i płacy realnej, podczas gdy według innych autorów skutek oddziaływania tego zjawiska jest odwrotny (zob. np. Ramey i Shapiro 1998; Edelberg, Eichenbaum i Fisher 1999; Burnside, Eichenbaum i Fisher 2004 oraz Ramey 2008). Wśród przyczyn tego typu rozbieżności wymienia się najczęściej różne sposoby identyfikacji szoków fiskalnych (na podstawie statystycznie obserwowanych, rzeczywistych zmian wydatków lub dochodów bądź publicznych zapowiedzi tych zmian), brak precyzyjnego rozróżnienia szoków mających na celu pobudzenie gospodarki lub nakierowanych na inne cele (np. względy społeczne) oraz różnice zachodzące na poziomie pomiaru długotrwałości szoków i opóźnień czasowych po stronie zmiennych makroekonomicznych (zob. ramka 3.11).

Ramka 3.11. Spór wokół programu stymulacji fiskalnej w USA w latach 2008–2010

Gdy w lecie 2008 roku kryzys na rynku kredytów hipotecznych uderzył w gospodarkę Stanów Zjednoczonych, rząd federalny w szybkim tempie przygotował Plan ożywienia i reinwestycji (*American Recovery and Reinvestment Plan* – ARRP), który przewidywał fiskalną stymulację gospodarki w postaci zwiększenia wydatków publicznych o kwotę 787 mld USD, z czego ponad 90% (717 mld USD) miało przypadać na lata 2009–2011. Stosowna legislacja została przyjęta przez Kongres w lutym 2009 roku i niezwłocznie weszła w życie.

Program ARRP wzbudził wielkie kontrowersje. Jego autorzy – przyszła przewodnicząca zespołu doradców ekonomicznych Prezydenta USA Christina Romer i główny ekonomista w biurze Wiceprezydenta USA Jared Bernstein – szacowali, że w ówczesnej sytuacji gospodarczej, przy niskich stopach procentowych i akomodacyjnej polityce pieniężnej, mnożnik wydatków fiskalnych wynosił 1,6 w perspektywie 2 lat, co oczywiście uzasadniało wzrost wydatków publicznych w celu wydźwignięcia gospodarki z recesji. Romer i Bernstein (2009) przyjęli tradycyjne keynesowskie założenie o niezmienności stóp procentowych w okresie trwania programu, a także założyli, że większość środków objętych pakietem zostanie przeznaczona na wzrost zakupu dóbr i usług. Krytycy takiego podejścia (m.in. Cogan

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i in. 2009) zarzucili autorom programu przyjęcie nierealistycznych, nadmiernie optymistycznych założeń, w tym przede wszystkim zignorowanie tego, że uczestnicy rynku kierują się racjonalnymi oczekiwaniami co do skutków polityki rządu oraz że część wydatków ma charakter transferowy, co osłabia ich pobudzający charakter. Według ich szacunków, przeprowadzonych przy zastosowaniu neokeynesowskiego modelu makroekonomicznego wykorzystującego oczekiwania uczestników rynku (Smets i Wouters 2007) oraz przy prognozowaniu rychłej podwyżki stóp procentowych, wartość mnożnika powinna spaść po dwóch latach do 0,6.

Jakie okazały się zaś rzeczywiste skutki wprowadzenia programu ARRP? W opinii autorów, pakiet miał doprowadzić do zwiększenia amerykańskiego PKB o 3,7% w okresie dwóch lat. Z kolei krytycy uznali, że uzyskany dzięki pakietowi przyrost PKB wyniesie tylko 0,6%, oczywiście w stosunku do scenariusza nieuwzględniającego programu. W rzeczywistości gospodarka amerykańska po spadku PKB o 2,8% w 2009 roku, w następnym roku osiągnęła już wzrost o 2,5%, w 2011 roku o 1,6%, a w 2012 roku o 2,6%. Rzeczywisty przebieg wydarzeń wydaje się więc potwierdzać zasadność wprowadzenia pakietu stymulującego. Cogan i in. (2009) pomylili się także, prognozując, że ultra niskie stopy procentowe zostaną podniesione najpóźniej w 2011 roku. W istocie stopy procentowe pozostały na poziomie bliskim zeru aż do końca 2015 roku.

3.10. Reguły fiskalne

Jak już podkreślano, w państwach demokratycznych występuje systematyczna skłonność do deficytu budżetowego, spowodowana politycznym cyklem wyborczym oraz istnieniem problemu wspólnego zasobu. Z kolei nadmierne, powtarzające się deficyty niosą szereg negatywnych następstw dla wzrostu gospodarczego: zwiększają dług publiczny, obniżają tempo wzrostu gospodarczego, a po pewnym czasie mogą doprowadzić do kryzysu zadłużenia. Dlatego też od wielu lat ekonomiści zastanawiają się, w jaki sposób przeciwdziałać tendencji do nadmiernego deficytu, zapewnić dyscyplinę budżetową i stabilność finansów publicznych w długim horyzoncie czasowym. Jedną z możliwości jest ograniczenie dyskrecjonalności władzy gospodarczej i podporządkowanie polityki fiskalnej określonym, publicznie ogłoszonym regułom. Innym rozwiązaniem jest delegowanie uprawnień dotyczących podejmowania decyzji dotyczących deficytu i długu do niezależnych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

instytucji. Zanim oba te sposoby zostaną jednak szerzej omówione, należy wyjaśnić, dlaczego polityka dyskrecjonalna niesie zawsze duże prawdopodobieństwo utraty kontroli nad sytuacją budżetową kraju.

3.10.1. Polityka dyskrecjonalna a polityka oparta na regułach

Polityka fiskalna, jak każda polityka, może być dyskrecjonalna lub podporządkowana pewnym regułom. Dyskrecjonalność, zwana inaczej uznaniowością, oznacza, że władza gospodarcza nie ogłasza żadnych wiążących reguł postępowania w zakresie polityki fiskalnej – poza, być może, jej ogólnym kierunkiem. Główną zaletą polityki dyskrecjonalnej jest to, że zapewnia ona dużą swobodę działania – rząd może szybko korygować kształt i kierunek polityki w odpowiedzi na zmieniające się warunki. Poza tym brak przyjętych wcześniej reguł i obietnic zwalnia rząd z obowiązku wyjaśniania przyczyn ewentualnych odstępstw od tych zasad. Ta swoboda działania jest niewątpliwie bardzo atrakcyjna z punktu widzenia każdej władzy. Lecz polityka dyskrecjonalna posiada też bardzo poważne mankamenty. Jej największą wadą jest wpływ na spadek przewidywalności działań rządu i wzrost niepewności. Uczestnicy rynku nie wiedzą, jak rząd będzie postępował w przyszłości, co zwiększa ryzyko podjęcia z ich strony błędnych decyzji, a w konsekwencji poniesienia strat. W takiej sytuacji rosną premie za ryzyko w stopach procentowych i spada aktywność gospodarcza. Jednocześnie brak reguł ułatwia podporządkowanie prowadzonej polityki doraźnym korzyściom, kosztem długookresowej stabilizacji.

Istotną wadą polityki dyskrecjonalnej jest również to, że cechuje ją tzw. niespójność czasowa lub dynamiczna (*time inconsistency, dynamic inconsistency*). Oznacza to, że działania rządu, które pozostają optymalne w danym okresie i w danych warunkach, wywołują zmianę tych warunków, w tym zwłaszcza zmianę zachowań uczestników rynku, co powoduje, że w następnym okresie dana polityka przestaje być optymalna. Wynika to z funkcjonowania racjonalnych oczekiwań. W rezultacie polityka dyskrecjonalna staje się źródłem szoków, które destabilizują gospodarkę[59]. Taka polityka jest więc czynnikiem hamującym wzrost gospodarczy, a niekiedy może nawet doprowadzić do kryzysu.

W przeciwieństwie do tego przyjęcie określonych, racjonalnych i wiarygodnych reguł postępowania ogranicza niepewność dotyczącą przyszłości, zwiększa zaufanie

[59] Problem niespójności czasowej polityki jest szerzej omówiony w rozdziale 5.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i obniża koszty działalności gospodarczej, a także minimalizuje skutki szoków i zwiększa stabilność gospodarki. Dlatego też stosowanie w polityce makroekonomicznej przejrzystych i stabilnych zasad jest w opinii większości ekonomistów pożądane.

3.10.2. Rodzaje reguł fiskalnych

Reguła fiskalna to trwałe ograniczenie nałożone na politykę fiskalną, wyrażone w postaci określonego syntetycznego miernika sytuacji fiskalnej, takiego jak deficyt finansów publicznych, dług publiczny, poziom wydatków budżetowych lub potrzeby pożyczkowe rządu (Kopits i Symansky 1998, Marchewka-Bartkowiak 2010). Reguła ma na ogół charakter prawnie wiążący, co nakłada na rząd obowiązek jej przestrzegania i grozi sankcjami prawnymi za jej naruszanie. Aby zostało uznane za wiarygodne, zobowiązanie takie musi odnosić się do dostatecznie długiego okresu, obejmującego horyzont czasowy decyzji podejmowanych przez większość prywatnych uczestników rynku. W ten sposób reguły fiskalne oddziałują na spadek niepewności związanej z polityką fiskalną rządu w przyszłości. W ramce 3.12 przedstawiono krótką historię reguł fiskalnych, poczynając od XIX-wiecznej „złotej reguły" wydatków, a kończąc na regułach fiskalnych obowiązujących obecnie w Unii Europejskiej.

Należy również pamiętać o tym, że skuteczność reguł fiskalnych zależy od ich rzeczywistego kształtu prawno-instytucjonalnego oraz od politycznej gotowości do ich bezwarunkowego przestrzegania. Choć głównym celem stosowania reguł fiskalnych jest wzmocnienie ogólnej dyscypliny budżetowej, mogą one także przyczyniać się do poprawy koordynacji polityki fiskalnej między różnymi szczeblami władzy publicznej.

Ramka 3.12. Najkrótsza historia reguł fiskalnych[60]

Pierwsze próby zaprowadzenia dyscypliny w zakresi finansów publicznych przy pomocy formalno-prawnych reguł były podejmowane przez władze publiczne już w połowie XIX wieku. Początkowo działania te dotyczyły budżetów władzy publicznej szczebla niższego niż rządowy i przyjmowały postać tzw. złotej reguły wydatków (*golden rule*). Stosowała ją na przykład większość rządów stanowych w USA,

[60] Na podstawie Kopits (2001).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

a także niektóre kantony w Szwajcarii. Istota tej reguły polegała na utrzymaniu bieżących wydatków na poziomie bieżących dochodów oraz finansowaniu ewentualnych inwestycji z pożyczek pozyskanych na rynku kapitałowym.

Druga fala stosowania reguł fiskalnych miała miejsce po II wojnie światowej w niektórych krajach uprzemysłowionych (Włochy, Japonia, Holandia, Republika Federalna Niemiec), gdzie w ramach odbudowy gospodarki wprowadzano programy stabilizacyjne i reformy walutowe, wymagające utrzymania ścisłej dyscypliny budżetowej. Większość tych obostrzeń przyjmowała nadal postać „złotej reguły", ale pojawiły się także nowe przepisy, wprowadzające m.in. zakaz finansowania deficytu budżetowego przez banki centralne. Ograniczenie takie wprowadziły – poza kilkoma krajami uprzemysłowionymi – także niektóre kraje rozwijające się (Indonezja, kraje Afryki Zachodniej). Skuteczność ich działań pozostawiała jednak wiele do życzenia, głównie ze względu na małą przejrzystość finansów publicznych i duży zakres dyskrecjonalności w zakresie księgowości budżetowej.

Najnowsza moda na stosowanie reguł fiskalnych została zapoczątkowana wprowadzeniem w 1994 roku w Nowej Zelandii głośnej ustawy o odpowiedzialności fiskalnej (*Fiscal Responsibility Act*), która nakładała na władze publiczne obowiązek pełnego informowania opinii publicznej o sytuacji budżetowej i zaciąganych pożyczkach. W ślad za Nową Zelandią poszły inne kraje rozwinięte. Współcześnie większość państw stosuje wybraną odmianę reguł fiskalnych, ograniczających wzrost wydatków, deficytu lub długu. Ich wspólnym mianownikiem pozostaje nacisk na jak największą przejrzystość polityki fiskalnej, w tym na przestrzeganie powszechnie akceptowanych standardów w zakresie rachunkowości, sprawozdawczości i demokratycznej odpowiedzialności (*democratic accountability*) za podejmowane decyzje.

Unia walutowa utworzona w Europie w 1999 roku nałożyła na państwa członkowskie ograniczenia w postaci reguł fiskalnych odnoszących się do deficytu budżetowego i długu publicznego. Ich celem była ochrona wspólnej waluty przed ryzykiem wynikającym z nadmiernie ekspansywnej polityki fiskalnej w warunkach samodzielności budżetowej poszczególnych państw członkowskich.

Polityka fiskalna prowadzona jest w każdym kraju w ramach określonych struktur prawno-instytucjonalnych. Ich istotnym składnikiem są reguły fiskalne, w tym zwłaszcza reguły numeryczne. Mogą one przyjmować różną postać. Część z nich ma charakter jakościowy i ogranicza się do zapisów prawnych zapewniających jak

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

największą przejrzystość (*transparency*) finansów publicznych oraz poddających je demokratycznej kontroli. Takie reguły fiskalne spotyka się w krajach anglosaskich. Inne przyjmują zaś postać reguł numerycznych w formie ilościowo zdefiniowanych limitów lub celów odnoszących się do wybranych zmiennych opisujących kształtowanie się sytuacji budżetowej. Do zmiennych tych zalicza się przede wszystkim wielkość deficytu budżetowego, wydatków budżetowych i długu publicznego. W przypadku reguł numerycznych można mieć do czynienia z celem lub limitem sformułowanym w postaci wartości bezwzględnej danej zmiennej (na przykład bezwzględna wielkość deficytu lub tempo wzrostu wydatków) lub w postaci określonej relacji dwóch zmiennych (na przykład relacja długu do PKB, deficytu do PKB lub wydatków do dochodów). Tego typu reguły stosowane są w wielu krajach Europy kontynentalnej, w tym zwłaszcza w państwach Unii Europejskiej.

Z punktu widzenia kategorii budżetowych, których dotyczą, numeryczne reguły fiskalne można podzielić na cztery zasadnicze grupy:

a) reguły dotyczące salda budżetu,
b) reguły dotyczące długu,
c) reguły dotyczące wydatków,
d) reguły dotyczące dochodów.

Reguły salda budżetu określają najczęściej maksymalny dopuszczalny poziom deficytu budżetowego, wyrażonego albo w wielkościach absolutnych, albo w relacji do PKB lub do dochodów. Na przykład, zgodnie z zapisami Protokołu nr 12 dołączonego do Traktatu o Funkcjonowaniu UE, dopuszczalna maksymalna wielkość deficytu całego sektora finansów publicznych (*general government*) w krajach Unii Europejskiej wynosi 3% PKB wyrażonego w cenach rynkowych. Niektóre kraje narzucają sobie jednak bardziej wymagające przepisy. Przykładowo, w Niemczech deficyt strukturalny (liczony w kategoriach potrzeb pożyczkowych netto) nie może przekraczać 0,35% PKB w przypadku budżetu federalnego i 0% w przypadku budżetów poszczególnych krajów związkowych. W Szwecji od 2006 roku saldo strukturalne sektora finansów publicznych musi wykazywać nadwyżkę w wysokości 1% PKB średnio w cyklu koniunkturalnym. Z kolei w Polsce w latach 2006–2007 przez krótki czas stosowana była reguła maksymalnego limitu deficytu budżetu państwa na poziomie 30 mld PLN.

Reguły długu publicznego odnoszą się do maksymalnego dopuszczalnego poziomu zadłużenia publicznego brutto, wyrażonego albo w wielkościach absolutnych, albo w relacji do PKB. Na przykład, zgodnie z zapisami wspomnianego Protokołu nr 12 dołączonego do Traktatu o Funkcjonowaniu UE, dopuszczalna

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wielkość zadłużenia publicznego nie może przekraczać 60% PKB wyrażonego w cenach rynkowych. Z kolei w Stanach Zjednoczonych obowiązuje federalny limit wartościowy zadłużenia publicznego, ustalany każdorazowo przez Kongres. W sierpniu 2011 roku limit ten został wyznaczony na poziomie 16 394 mld USD, w 2013 roku zwiększono go do 16 699 mld USD[61], a od października 2013 roku limit został zawieszony decyzją Kongresu, dając rządowi USA *de facto* możliwość dalszego zadłużania się.

Reguły dotyczące sposobu wydatkowania określają albo maksymalny poziom, albo maksymalne tempo wzrostu ogółu wydatków publicznych lub ich części, często w powiązaniu z dochodami lub PKB. Na przykład w Polsce w latach 2011–2013 obowiązywała tzw. tymczasowa dyscyplinująca reguła wydatkowa, zgodnie z którą tzw. elastyczne wydatki budżetu państwa (czyli wydatki, które nie są zdeterminowane ustawowo) nie mogą z roku na rok rosnąć szybciej niż stopa rocznej inflacji CPI powiększona o 1 punkt procentowy. We Francji w latach 2011–2013 ogólne wydatki budżetu państwa zostały zamrożone w wyrażeniu wartościowym, z wyłączeniem kosztów obsługi długu i wydatków emerytalnych. Z kolei w Niemczech średnie tempo wzrostu wydatków publicznych rządu federalnego i rządów poszczególnych krajów związkowych w latach 2008–2009 musiało być niższe niż średnie tempo wzrostu dochodów budżetowych.

Reguły odnoszące się do dochodów są stosowane najrzadziej i dotyczą poziomu lub tempa wzrostu ogółu dochodów budżetowych bądź ich części. Przykładem reguły dochodowej jest obowiązująca w Danii w latach 2001–2010 decyzja o niepodnoszeniu stawek podatków pośrednich i bezpośrednich. Inną jeszcze realizację tego obostrzenia może stanowić przyjęta we Francji w 2009 roku decyzja mówiąca o tym, że wszelkie nadwyżki dochodów budżetowych mają być przeznaczone wyłącznie na obniżenie deficytu. Podobne rozwiązanie zastosowano na Litwie.

Skuteczne reguły fiskalne muszą spełniać dwa często sprzeczne ze sobą wymogi. Po pierwsze, chodzi o umocnienie wśród uczestników rynku przekonania, że przepisy te okazują się skuteczne w utrzymaniu dyscypliny budżetowej i że politycy nie będą mogli zmieniać reguł stosownie do swoich potrzeb. Chodzi więc o to, aby przepisy fiskalne były trwałe i wiarygodne, bo tylko takie reguły obniżają niepewność i zwiększają przewidywalność polityki budżetowej. Z drugiej strony, wadą wszelkich sztywnych reguł jest brak elastyczności i procykliczny charakter. Uwzględnienie trwałości i niezbędnej elastyczności w ramach wyłącznie jednej normy wymaga nie

[61] Levitt Brass, Nicola i Nuschler (2013).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

lada wysiłku. Osiąga się to poprzez wprowadzenie do reguł fiskalnych tzw. klauzul ucieczki (*escape clauses*), czyli procedur łagodzących lub zawieszających działanie reguły fiskalnej w określonych sytuacjach, zwłaszcza gdy wystąpią nieprzewidziane i niekorzystne okoliczności zewnętrzne. Niezbędny zakres elastyczności uzyskuje się także poprzez wydłużenie okresu, w jakim mają obowiązywać reguły, uwzględnianie dopuszczalnych marginesów odchyleń lub uzupełnianie reguły mechanizmem samokorygującym. Przykładem wydłużenia okresu obowiązywania reguły jest zastąpienie reguły zrównoważonego budżetu w każdym kolejnym roku budżetowym (saldo faktyczne równe zeru) regułą zrównoważonego budżetu w okresie całego cyklu koniunkturalnego (saldo strukturalne równe zeru), co pozwala na uwzględnienie działania automatycznych stabilizatorów koniunktury i eliminuje w ten sposób procykliczny charakter reguły[62]. Przykładem zasstosowania dopuszczalnych odstępstw od reguły są zapisy Paktu Stabilności i Wzrostu w Unii Europejskiej, które określają, w jakich okolicznościach dopuszczalne są odchylenia od reguły deficytu (3% PKB) i reguły długu (60% PKB). Wreszcie, przykładem mechanizmu samokorygującego mogą być również przepisy polskiej ustawy o finansach publicznych, nakładające na rząd bezwzględny obowiązek podjęcia określonych działań naprawczych, gdy przekroczone zostaną dane wartości progowe relacji długu publicznego do PKB. Instrumentem zwiększającym elastyczność reguł jest też tworzenie rezerw w okresach dobrej koniunktury, służących do ewentualnego wykorzystania w okresach recesji (tzw. fundusze na „złą pogodę" – *rainy day funds*). Rozwiązanie takie jest stosowane w wielu stanach USA.

Popularną odmianą reguły fiskalnej jest wspomniana już wcześniej „złota reguła" finansów publicznych, która zakłada zbilansowanie bieżących wydatków i bieżących dochodów natomiast dopuszcza zaciąganie pozyczek wyłącznie na finansowanie inwestycji. Jej zaleta polega na tym, że, zapewniając określony poziom dyscypliny fiskalnej w zakresie wydatków bieżących, reguła ta zachowuje jednocześnie charakter prowzrostowy, ponieważ wyłącza inwestycje publiczne spod działania mechanizmu dyscyplinującego. Dodatkowym uzasadnieniem w przypadku zaciągania pożyczek na finansowanie inwestycji – zwłaszcza infrastrukturalnych – staje się to, że choć spłata tych pożyczek obciąża przyszłe pokolenia, to będą jednak one korzystały ze zbudowanej wcześniej infrastruktury. W praktyce, „złota reguła" nie

[62] Na przykład w Szwajcarii obowiązujące reguły wymagają, aby rząd realizował nadwyżkę budżetową, gdy faktyczne tempo wzrostu PKB przekracza 1,8% rocznie, i dopuszczają deficyt, gdy tempo wzrostu spada poniżej 0,5%.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zawsze funkcjonuje zgodnie z oczekiwaniami. Jej główną wadą pozostaje możliwość manipulowania klasyfikacją wydatków, tak aby włączać niektóre wydatki bieżące do kategorii wydatków inwestycyjnych, co oczywiście osłabia mechanizm dyscyplinujący wydatki.

Konstruując i stosując reguły fiskalne, należy pamiętać o ryzyku przeregulowania (*overidentification*). Na przykład, zapisy Paktu Stabilności i Wzrostu narzucają państwom członkowskim UE obowiązek jednoczesnego przestrzegania reguły zrównoważonego budżetu w cyklu koniunkturalnym oraz reguły górnego limitu długu publicznego na poziomie 60% PKB. W istocie, gdy pierwsza z tych reguł jest konsekwentnie przestrzegana, relacja długu do PKB zmierza w długim okresie do zera, a więc druga reguła staje się zbędna. Jednocześnie – w krótszym okresie – w krajach, w których relacja długu do PKB przekracza 60%, stosowanie reguły zrównoważonego budżetu w cyklu koniukturalnym pozwala obniżać zadłużenie do dopuszczalnego poziomu. Przykład ten pokazuje również, jak duże znaczenie ma osadzenie reguł fiskalnych w ramach wielookresowego planowania budżetowego.

Skuteczność stosowania reguł fiskalnych zależy w dużym stopniu od rangi ich umocowania prawnego, systemu nadzoru i kontroli przestrzegania, oraz sankcji za ich naruszenie (zob. ramka 3.13). Najbardziej wiarygodne pozostają reguły zapisane w konstytucji danego państwa lub co najmniej na poziomie ustawy parlamentarnej. Akty niższej rangi mają słabszą moc dyscyplinującą, ponieważ łatwiej mogą być zmieniane przez władzę polityczną. System nadzoru i kontroli respektowania reguł fiskalnych musi być niezależny, ciągły i przejrzysty dla uczestników rynku i opinii publicznej. W szczególności system monitorowania sytuacji na poziomie finansów publicznych powinien zapewniać stałą kontrolę nad procesem budżetowym i przestrzeganiem reguł fiskalnych. W większości krajów funkcję kontroli i nadzoru pełnią specjalnie do tego powołane urzędy prowadzące audyt finansów publicznych[63]. W niektórych krajach, niezależnie od kontroli wykonania budżetu, zadanie nadzoru nad utrzymaniem dyscypliny fiskalnej powierza się niezależnej od rządu instytucji kontrolnej (rada polityki fiskalnej – zob. dalej). Wreszcie formalne sankcje za złamanie obowiązujących przepisów są niezbędnym elementem skuteczności działania reguł fiskalnych. W prawie europejskim w ramach Unii Europejskiej istnieje możliwość nałożenia sankcji prawnych na państwa naruszające dyscyplinę budżetową. Stosuje się w tym przypadku m.in. zablokowanie

[63] W Polsce rolę tę pełni Najwyższa Izba Kontroli w odniesieniu do budżetu państwa oraz Regionalne Izby Obrachunkowe w odniesieniu do finansów jednostek samorządu terytorialnego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

dostępu do środków z funduszu spójności, obowiązkowe nieoprocentowane depozyty oraz kary finansowe[64].

Ramka 3.13. Elementy składowe reguł fiskalnych

Siła oraz skuteczność reguł fiskalnych w zakresie utrzymania dyscypliny finansów publicznych zależy od spełnienia szeregu warunków instytucjonalnych i prawnych w ramach istniejącego w danym kraju systemu planowania i wykonywania budżetu. Do tych niezbędnych warunków należą:

- **podstawa prawna** – reguła fiskalna powinna mieć silne umocowanie prawne w konstytucji lub na poziomie ustawowym w celu zapewnienia jej wiarygodności, nienaruszalności i trwałości. Wymagane jest także, aby stosowny akt prawny zawierał przepisy określające mechanizm monitorowania i kontroli przestrzegania reguły, warunki i procedury jej zmiany, oraz zasady postępowania w przypadku jej naruszenia;
- **wieloletnie ramy obowiązywania** – reguła fiskalna powinna obejmować co najmniej średni okres czasu, aby zapewnić odpowiedni stopień elastyczności w danym czasie oraz zapobiec krótkookresowym manipulacjom na poziomie dochodów i wydatków publicznych;
- **system rachunkowości** – Metodologia ESA 2010 jest spójna z unijnym systemem nadzoru i kontroli dyscypliny budżetowej, lecz metoda kasowa może okazać się skuteczniejsza w sygnalizowaniu odstępstw od reguły. Dlatego też zaleca się równoległe stosowanie obu metod: reguła fiskalna może zostać skonstruowana zgodnie z metodą kasową, a następnie jej wykonanie daje się okresowo przetransponować na system ESA 2010;
- **kontrola przestrzegania reguły** – skuteczne monitorowanie przestrzegania reguły fiskalnej wymaga powołania do życia niezależnej instytucji kontrolnej oraz zapewnienia jej nieograniczonego dostępu do pełnej sprawozdawczości budżetowej;
- **mechanizm korekcyjny** – przepisy prawa regulujące funkcjonowanie reguły fiskalnej muszą zapewniać automatyczne uruchamianie prawnego mechanizmu korygującego w sytuacji naruszenia reguły. Podjęcie działań korekcyjnych nie może być uzależnione od dyskrecjonalnej decyzji władz;

[64] Należy jednak zaznaczyć, że sankcje te nigdy jeszcze nie były stosowane.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

– **sankcje** – mechanizm wymuszonego stosowania reguły fiskalnej można uzupełnić sankcjami za jej naruszanie. Są to z reguły kary finansowe, nakładane na instytucje budżetowe odpowiedzialne za utrzymanie dyscypliny budżetowej, oraz kary dyscyplinarne, finansowe i prawne nakładane na osoby odpowiedzialne za naruszenie reguły;
– **klauzule ucieczki** – klauzule ucieczki (*escape clauses*) określają precyzyjnie warunki i sytuacje, w których możliwe jest zawieszenie lub opóźnienie działania reguły fiskalnej. Ich poprawne i precyzyjne zdefiniowanie ma istotne znaczenie w przypadku utrzymania wysokiej wiarygodności reguł.

Źródło: *Public finances in EMU – 2010*, European Commission, część II, rozdział 3, Brussels 2011, s. 101.

3.10.3. Reguły fiskalne w Unii Europejskiej

W 2011 roku w krajach UE obowiązywało łącznie 77 różnych reguł fiskalnych, przy czym najwięcej z nich dotyczyło salda budżetu. Jak pokazuje tabela 3.9, liczba reguł fiskalnych systematycznie rosła na przestrzeni ostatnich 20 lat. W 2011 roku 24 państwa członkowskie stosowały jedną lub więcej reguł[65]. Świadczy to o tym, że w coraz większym stopniu traktuje się je jako skuteczne instrumenty umacniania dyscypliny finansów publicznych.

Tabela 3.9. Liczba numerycznych reguł fiskalnych obowiązujących w krajach Unii Europejskiej (wybrane lata)

Rodzaj reguły fiskalnej	1990	1995	2000	2006	2010	2011
Reguły salda budżetu	9	14	19	27	28	31
Reguły długu	3	5	13	19	21	21
Reguły wydatków	1	5	11	16	17	19
Reguły dochodów	0	3	5	5	6	6
Razem	13	27	46	67	72	77
Liczba reguł przypadająca średnio na jeden kraj członkowski	1,08	1,80	3,07	2,68	2,67	2,85

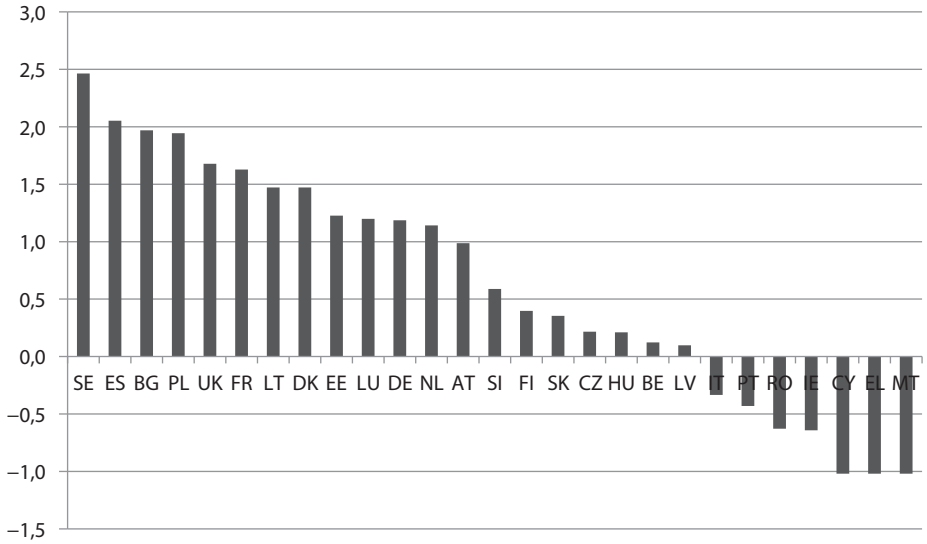
Źródło: http://ec.europa.eu/economy_finance/db_indicators/fiscal_governance/fiscal_rules/index_en.htm (10.05.2014).

[65] Trzy kraje UE, w których nie stosowano reguł fiskalnych, to Grecja, Cypr i Malta.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Komisja Europejska zastosowała specjalny wskaźnik mierzący siłę wpływu danej reguły fiskalnej na sytuację budżetową kraju (*the fiscal rule strength index* – FRSI). Uwzględnia on informacje na temat (1) podstawy prawnej reguły fiskalnej, (2) prawnych i faktycznych możliwości dokonywania zmian w regule, (3) stopnia niezależności instytucji odpowiedzialnej za monitorowanie i kontrolę przestrzegania reguły oraz za działania korekcyjne, (4) mechanizmu korekcyjnego uruchamianego w przypadku naruszenia reguły oraz (5) przejrzystości mechanizmu reguły oraz stopnia, w jakim opinia publiczna jest w stanie oceniać jej stosowanie.

Wykres 3.15. Wskaźnik siły reguł fiskalnych FRI w krajach członkowskich UE w 2011 roku



Źródło: opracowanie własne na podstawie http://ec.europa.eu/economy_finance/db_indicators/fiscal_governance/fiscal_rules/index_en.htm (18.12.2014).

Wykorzystując tak oszacowane indywidualne wskaźniki siły poszczególnych reguł, Komisja Europejska oblicza całościowy wskaźnik siły reguł fiskalnych (*fiscal rule index* – FRI) w odniesieniu do danego kraju. Procedura polega na sumowaniu wartości poszczególnych wskaźników FRI ważonych zakresem pokrycia kosztów sektora finansów publicznych objętych działaniem reguły[66]. W sytuacji, gdy istnieje

[66] Szczegółowy opis metody liczenia wskaźnika FRI można znaleźć na stronach Komisji Eurropejskiej pod adresem: http://ec.europa.eu/economy_finance/db_indicators/fiscal_governance/fiscal_rules/index_en.htm (18.12.2014).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

więcej niż jedna reguła fiskalna odnosząca się do tego samego obszaru finansów publicznych, kolejne reguły otrzymują malejące wagi, tak aby uwzględnić malejącą użyteczność reguł odnoszących się do danego zakresu. Wartości wskaźnika FRI w przypadku krajów Unii Europejskiej w 2011 roku zostały przedstawione na wykresie 3.15.

3.10.4. Reguły fiskalne w Polsce. Hamulec długu i progi ostrożnościowe

Polska już w latach 90. XX wieku jako jedyny kraj w Europie wprowadziła regułę długu publicznego na najwyższym, konstytucyjnym poziomie. Przyjmując w 1997 roku nową Konstytucję, zawarto w niej zapis o tzw. hamulcu długu, służącym ograniczaniu poziomu zadłużenia publicznego. Zgodnie z art. 216 ust. 1 Konstytucji RP, „nie wolno zaciągać pożyczek lub udzielać gwarancji i poręczeń finansowych, w następstwie których państwowy dług publiczny przekroczy 3/5 wartości rocznego produktu krajowego brutto. Sposób obliczania wartości rocznego produktu krajowego brutto oraz państwowego długu publicznego określa ustawa".

Zakaz zwiększania zadłużenia ponad poziom 60% PKB znalazł oczywiście ekonomicznie uzasadnienie w dążeniu do wzmocnienia wiarygodności finansowej kraju i uniknięcia ryzyka niekontrolowanego wzrostu zadłużenia[67]. Zapis ten wynikał także z chęci dostosowania się do istniejących już od 1992 roku przepisów unijnych, nakazujących państwom członkowskim ograniczenie relacji długu do PKB do 60%. Ponieważ Polska złożyła w 1994 roku wniosek o akcesję do Unii Europejskiej i liczyła na uzyskanie członkostwa w ciągu następnych kilku lat, zaakcentowała również w ten sposób to, że poważnie traktuje swoje przyszłe zobowiązania wynikające z wejścia do Unii.

Ze względu na swą wyjątkowość i brak podobnych rozwiązań w innych krajach, polski „hamulec długu" został początkowo uznany za przejaw nadmiernej i niekoniecznie uzasadnionej gorliwości w przestrzeganiu dyscypliny fiskalnej. O jego zaletach przekonano się jednak podczas kryzysu zadłużenia w latach 2008–2012, i to do tego stopnia, że – poczynając od 2013 roku – inne kraje członkowskie UE zaczęły wdrażać podobne rozwiązania u siebie w ramach wzmacniania prawno-instytucjonalnych ram polityki fiskalnej.

[67] Polska w latach 1980–1994 jednostronnie zawiesiła obsługę zadłużenia zagranicznego (z wyjątkiem odsetek od długu bankowego). W sensie technicznym była więc państwem niewypłacalnym, choć ani polski rząd, ani polscy wierzyciele nie ogłosili formalnie stanu niewypłacalności (*default*).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Konstytucyjnemu „hamulcowi długu" towarzyszą w polskim prawie dodatkowe reguły fiskalne, mające wymuszać konieczne działania naprawcze z odpowiednim wyprzedzeniem. W ustawie o finansach publicznych wprowadzono tzw. progi ostrożnościowe, nakazujące podejmowanie określonych działań naprawczych i dyscyplinujących w sytuacji, gdy zadłużenie zbliża się do progu 60% PKB. W latach 1999–2013 działania naprawcze miały być sukcesywnie podejmowane, gdy stosunek długu do PKB przekroczy kolejno progi 50%, 55% i 60%.

Gdy relacja długu do PKB w danym roku jest wyższa niż 50%, ale nie przekracza 55%, na kolejny rok Rada Ministrów uchwala projekt ustawy budżetowej, w którym relacja deficytu budżetu państwa do dochodów budżetu państwa nie może być wyższa niż relacja deficytu budżetu państwa do dochodów budżetu państwa z roku bieżącego (jak wynika z obowiązującej w danym roku ustawy budżetowej). Gdy relacja długu do PKB przekracza 55%, ale pozostaje mniejsza niż 60%, wówczas na kolejny rok Rada Ministrów uchwala projekt ustawy budżetowej, w ramach którego nie przewiduje się w ogóle deficytu budżetu państwa oraz wzrostu wynagrodzeń wśród pracowników państwowej sfery budżetowej, a waloryzacja rent i emerytur nie może przekroczyć poziomu odpowiadającego wzrostowi cen towarów i usług konsumpcyjnych. Dodatkowo wprowadza się zakaz udzielania pożyczek i kredytów z budżetu państwa z wyjątkiem rat kredytów i pożyczek udzielonych w latach poprzednich. Ponadto, Rada Ministrów przedstawia Sejmowi program sanacyjny, mający na celu obniżenie relacji długu do PKB. Podobnym ograniczeniom podlegają także budżety jednostek samorządu terytorialnego.

Gdy relacja długu do PKB przekroczy 60%, Rada Ministrów, najpóźniej w terminie miesiąca od dnia publicznego ogłoszenia tej relacji, przedstawia Sejmowi program sanacyjny, mający na celu ograniczenie tej relacji do poziomu niższego niż 60%. Ponadto Rada Ministrów uchwala na kolejny rok projekt ustawy budżetowej, w którym nie przewiduje się deficytu budżetu państwa, a wydatki budżetów jednostek samorządu terytorialnego określone w uchwale budżetowej na kolejny rok nie mogą być wyższe niż dochody tych budżetów. Jednocześnie, poczynając od siódmego dnia od chwili publicznego ogłoszenia relacji długu do PKB, jednostki sektora finansów publicznych nie mogą udzielać nowych poręczeń i gwarancji[68]. Ograniczenia budżetowe związane z progami ostrożnościowymi zostały przedstawione w tabeli 3.10.

[68] Ustawa o finansach publicznych z 27 sierpnia 2009 roku, art. 86–88.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 3.10. Budżetowe procedury ostrożnościowe w Polsce obowiązujące w latach 1998–2013

Próg ostrożnościowy (dług publiczny/PKB)	Ograniczenia budżetu centralnego	Ograniczenia budżetu samorządu lokalnego
50% < X ≤ 55%	DEF (n + 1) DEF (n) ---------- ≤ ---------- REV (n + 1) REV (n)	DEF (n + 1) DEF (n) --------- ≤ ---------- REV (n + 1) REV (n)
55% < X < 60%	DEF ≥ 0 Lub: DSP (n + 1) DSP (n – 1) ---------- < ---------- PKB (n + 1) PKB (n – 1)	DEF (n + 1) ≥ ŚZNZ
X ≥ 60%	DEF ≥ 0 ΔPG = 0	DEF ≥ 0 ΔPG = 0

Oznaczenia:
X – relacja państwowego długu publicznego w relacji do PKB (dane w %),
DEF – deficyt budżetu państwa,
REV – dochody budżetu państwa,
DSP – dług Skarbu Państwa,
ΔPG – nowe gwarancje i poręczenia udzielane przez jednostki sektora finansów publicznych,
ŚZNZ – środki zagraniczne niepodlegające zwrotowi, wykorzystywane przez samorządy lokalne (z UE),
n – rok bieżący.

Źródło: opracowanie własne na podstawie ustawy o finansach publicznych z 27 sierpnia 2009 roku.

Poza regułą długu, w latach 2011–2013 w polskim systemie finansów publicznych funkcjonowała także reguła wydatków. Określana mianem „tymczasowej reguły wydatkowej", ograniczała tempo wzrostu tzw. wydatków elastycznych budżetu państwa do poziomu stopy inflacji CPI powiększonej o 1 punkt procentowy[69]. Ponieważ w normalnych warunkach realny PKB rośnie szybciej niż w tempie 1%, zastosowanie tej reguły oznaczałoby spadek udziału wydatków budżetowych w PKB. Dopiero gdy tempo wzrostu spadnie poniżej 1% PKB, reguła wydatkowa umożliwi antycykliczny wzrost udziału wydatków w PKB. Mimo tych oczywistych zalet, rozwiązanie to miało też kilka wad. Po pierwsze, nie dało się go zastosować do wydatków prawnie zdeterminowanych, stanowiących około 70% wszystkich wydatków budżetu państwa. Po drugie, nie obejmowało pozostałych części sektora finansów publicznych, co umożliwiało przesuwanie niektórych kategorii wydatków

[69] Reguła ta została po raz pierwszy zaproponowana przez ministra finansów Marka Belkę w 2001 roku, lecz ostatecznie nie została przyjęta przez rząd Leszka Millera i nie weszła w życie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

publicznych poza budżet państwa i wyłączanie ich w ten sposób z dyscyplinującego działania reguły wydatkowej[70].

3.10.5. Stabilizująca reguła wydatkowa

Słabą stroną systemu progów ostrożnościowych dyscyplinujących finanse publiczne w Polsce była ich procykliczność oraz brak niezbędnej elastyczności. Progi ostrożnościowe zmuszają wprawdzie do podjęcia działań oszczędnościowych w zakresie finansów publicznych, ale dopiero w momencie powstania poważnej nierównowagi, z którą z reguły ma się do czynienia w okresach złej koniunktury. Z kolei w okresie wysokiego tempa wzrostu gospodarczego, a, co za tym idzie, także dobrej sytuacji dochodowej budżetu państwa nie funkcjonują żadne automatyczne mechanizmy służące stworzeniu buforów finansowych i przygotowaniu finansów publicznych na gorsze czasy. Dlatego też system progów ostrożnościowych został zreformowany w 2013 roku. Reguła długu zawarta w Konstytucji RP pozostała bez zmian. Lecz w miejsce tymczasowej reguły wydatkowej wprowadzono tzw. trwałą stabilizującą regułę wydatkową.

Stabilizująca reguła wydatkowa (SRW) usuwa mankamenty poprzedniego systemu reguł. Ma charakter antycykliczny, jest znacznie elastyczniejsza i obejmuje swoim zakresem niemal cały sektor finansów publicznych. Jej przyjęcie stanowi jednocześnie wypełnienie postanowień unijnych regulacji, wprowadzonych Dyrektywą Rady UE 2011/85/UE z dniem 8 listopada 2011 roku[71]. Generalnie, działanie SRW polega na tym, że w sposób elastyczny dostosowuje ona pułap dopuszczalnych wydatków publicznych do aktualnego stanu zadłużenia państwa oraz do bieżącego stanu koniunktury gospodarczej. W szczególności, gdy dług publiczny nie przekracza 50% PKB, a stan finansów publicznych nie wymaga korekty, limit wydatków rośnie w tempie równym średniookresowemu tempu wzrostu PKB, powiększonemu o prognozowana stopę inflacji. Z kolei gdy relacja długu publicznego przekracza 50% lub 55% lub jeśli deficyt sektora finansów publicznych staje się wyższy niż 3% PKB, uruchamiany jest automatyczny mechanizm korygujący,

[70] W latach 2006–2007 obowiązywała także wprowadzona przez rząd Jarosława Kaczyńskiego reguła salda budżetowego w postaci ograniczenia deficytu budżetu państwa do poziomu 30 mld PLN. Ponieważ jej konstrukcja budziła jednak wiele wątpliwości (głównie ze względu na swój procykliczny charakter i arbitralnie ustalaną wielkość dopuszczalnego deficytu), została ona szybko zarzucona.

[71] Council Directive 2011/85/EU of 8 November 2011 on requirements for budgetary frameworks of the Member States.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

który ogranicza tempo wzrostu wydatków. Ponadto, skala korekty wydatków dostosowywana jest do aktualnego tempa wzrostu PKB – im wolniejsze tempo wzrostu gospodarczego, tym wyższy limit wydatków. W przypadku wystąpienia głębokiej recesji nie dokonuje się korekty wydatków.

Wreszcie, SRW uwzględnia konieczność realizowania średniookresowego celu budżetowego (*medium-term objective* – MTO). W ramach mechanizmu korygującego funkcjonuje tzw. konto kontrolne, na którym zapisywane są skumulowane odchylenia faktycznego salda sektora finansów publicznych od celu średniookresowego. Jeśli w danym roku wynik sektora będzie niższy od celu MTO (obecnie –1% PKB), to na koncie kontrolnym odłoży się odchylenie ujemne, jeśli zaś wyższy – dodatnie. W sytuacji, gdy poziom odchyleń od celu (ujemnych lub dodatnich) przekroczy ustalony limit (+/–6% PKB), zostanie zastosowana dodatkowa korekta dynamiki limitu wydatków. Aby uniknąć procykliczności polityki fiskalnej, mechanizm korygujący nie będzie wymuszał korekty wydatków w warunkach tzw. „złych" lub „dobrych" czasów, zdefiniowanych jako odchylenie prognozowanego wzrostu gospodarczego (w dół lub w górę) o więcej niż 2 punkty procentowe od poziomu średniookresowego. Ogólny schemat funkcjonowania SRW przedstawiony został w ramce 3.14.

Ramka 3.14. Mechanizm korygowania nierównowagi w ramach stabilizującej reguły wydatkowej

Ogólny mechanizm funkcjonowania SRW można przedstawić w postaci następującej sekwencji działań[72]. Punktem wyjścia przy ustaleniu nieprzekraczalnego limitu wydatków na rok następny t jest kwota wydatków przyjęta na rok bieżący (t – 1) w ramach ustawy budżetowej. Następnie kwota ta jest korygowana zgodnie z następującym algorytmem:

– Krok 1: kwota wydatków przyjęta na rok (t – 1) jest korygowana o wskaźnik inflacji równy celowi inflacyjnemu NBP ustalonemu przez Radę Polityki Pieniężnej na rok t[73].

[72] Zob. art. 112aa ustawy o finansach publicznych z dnia 27 sierpnia 2009 roku, z późn. zm.

[73] W pierwszej wersji reguły wydatkowej (obowiązującej w latach 2014–2015) kwota wydatków z roku (t – 1) była dodatkowo korygowana o różnicę między inflacją CPI założoną w ustawie budżetowej na rok (t – 1) a prognozowaną inflacją CPI na rok (t) (jeśli inflacja rzeczywista okazała się wyższa od przyjętej w ustawie budżetowej, planowana kwota wydatków ulegała zwiększeniu o tę różnicę).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

- Krok 2: tak skorygowana kwota wydatków jest następnie mnożona przez wskaźnik średniookresowej dynamiki wzrostu PKB (w cenach stałych), wyliczony jako średnia z ośmiu obserwacji, obejmujących rzeczywiste stopy wzrostu PKB w ciągu sześciu lat poprzedzających rok ($t-1$) oraz prognozowane stopy wzrostu na lata ($t-1$) i (t).
- Krok 3: wielkość planowanych wydatków na okres (t), obliczona w kroku 2, podlega następnie korekcie, zgodnie z mechanizmem korygowania nierównowagi opisanym na poniższym schemacie (rysunek 3.11). Korekta (obniżenie wydatków) może wynosić 1,5 lub 2 punkty procentowe i ma miejsce w sytuacji, gdy:
  a) dług publiczny przekracza 50% (brak korekty w tzw. „złych czasach", obniżenie wzrostu wydatków o 1,5 punkt procentowy w pozostałych sytuacjach) lub 55% PKB (obniżenie wzrostu wydatków o 2 punkty procentowe);
  b) skumulowany deficyt sektora finansów publicznych przekracza 6% PKB (brak korekty w tzw. „złych czasach", obniżenie wzrostu wydatków o 1,5 punktu procentowego w pozostałych sytuacjach). Wielkość korekty waha się w zależności od odchylenia prognozowanej dynamiki PKB w roku (t) od średniookresowej dynamiki PKB. W tzw. „złych czasach", gdy prognozowana dynamika jest niższa od średniookresowej wartości o co najmniej 2 punkty procentowe, korekta wynosi zero.
- Krok 4: tak wyliczona wartość wydatków zostaje następnie skorygowana o wartość ewentualnych działań dyskrecjonalnych podjętych przez rząd.

Formuła wykorzystywana do wyliczenia nieprzekraczalnego limitu wydatków na rok t ma postać następującą[74]:

$$WYD_t = WYD_{t-1} \cdot E_{t-1}(CCPI) \cdot \left[WPKB_t + K_t\right] + E_t\left(\Delta DD_t\right), \tag{3.51}$$

przy czym:

$$WPKB_t = \sqrt[8]{\frac{PKB_{t-2}}{PKB_{t-8}} \cdot E_t\left(\frac{PKB_{t-1}}{PKB_{t-2}}\right) \cdot E_t\left(\frac{PKB_t}{PKB_{t-1}}\right)}, \tag{3.52}$$

[74] Zgodnie z nowelizacją ustawy o finansach publicznych z dnia 10 grudnia 2015 roku. Celem dokonanej nowelizacji było złagodzenie rygorów stabilizującej reguły wydatkowej w trzech punktach: ustalenie wyższej początkowej kwoty wydatków na rok 2015, wprowadzenie zasady korygowania wydatków nominalnych o (stały) cel inflacyjny NBP, zamiast o rzeczywisty wskaźnik inflacji w roku ($t-1$), oraz podniesienie limitu wydatkowego na rok (t) o dochody jednorazowe i tymczasowe (bez wpływu na kwotę wydatków obliczaną na kolejny rok).

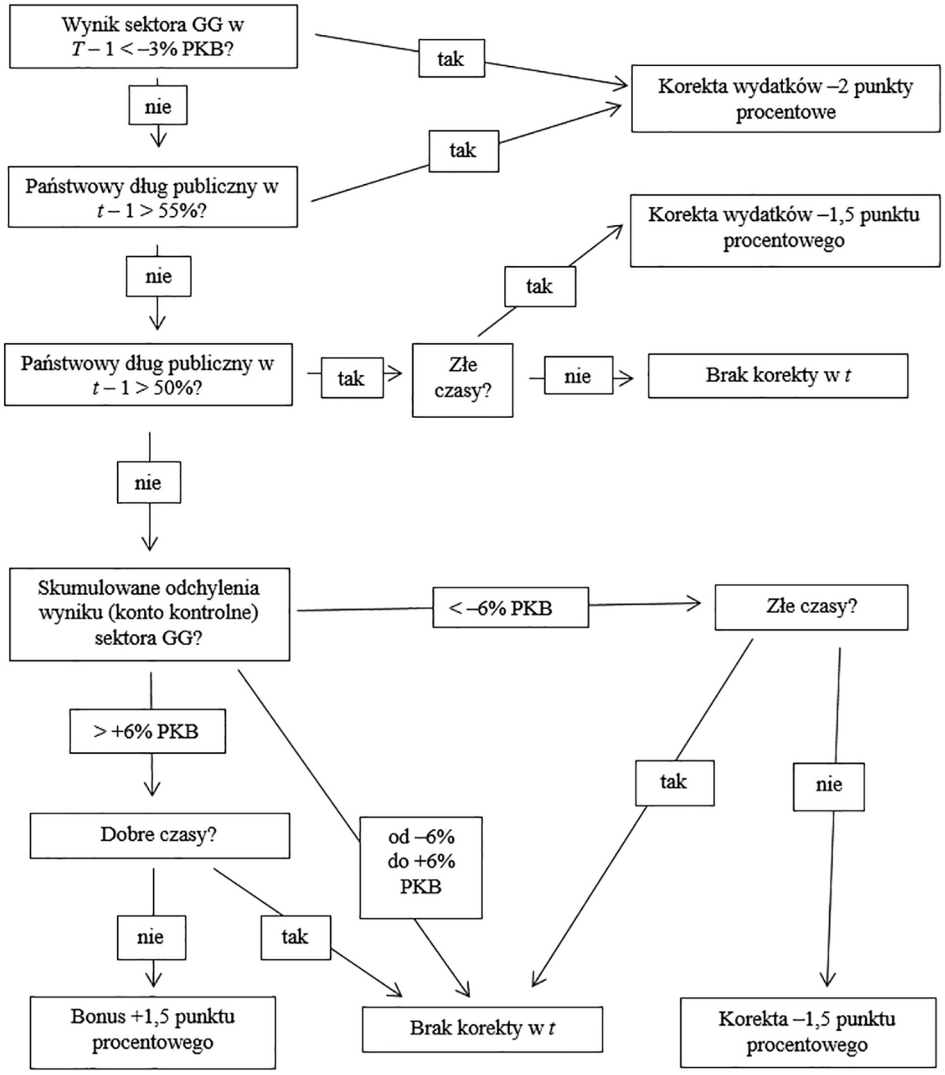
Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gdzie poszczególne symbole oznaczają:
t – rok, na który jest obliczana kwota wydatków,
WYD_t – kwota wydatków określona w przedłożonym Sejmowi projekcie ustawy budżetowej na rok t,
WYD_{t-1} – kwota wydatków określona w przedłożonym Sejmowi projekcie ustawy budżetowej na rok $t-1$,
$E_{t-1}(CCPI)$ – poziom średniookresowego celu inflacyjnego określonego na rok $(t-1)$ przez Radę Polityki Pieniężnej w ramach założeń polityki pieniężnej,
$WPKB_t$ – wskaźnik średniookresowej dynamiki wartości PKB w cenach stałych określony w przedłożonym Sejmowi projekcie ustawy budżetowej na rok t z dokładnością do setnych części procentu,
K_t – wielkość korekty kwoty wydatków określonej w przedłożonym Sejmowi projekcie ustawy budżetowej na rok t,
PKB_t – wartość produktu krajowego brutto w roku t,
$E_t\left(\frac{PKB_{t-1}}{PKB_{t-2}}\right)$ – prognozowana w przedłożonym Sejmowi projekcie ustawy budżetowej na rok t roczna dynamika wartości PKB w cenach stałych w roku $t-1$,
$E_t\left(\frac{PKB_t}{PKB_{t-1}}\right)$ – prognozowana w przedłożonym Sejmowi projekcie ustawy budżetowej na rok t roczna dynamika wartości PKB w cenach stałych w roku t,
$E_t(\Delta DD_t)$ – prognozowana wartość działań dyskrecjonalnych ogółem, planowanych na rok t,
jeżeli wartość każdego z nich przekracza 0,03% wartości PKB prognozowanej w przedłożonym Sejmowi projekcie ustawy budżetowej na rok t[75]; w przypadku planowanej w roku t rezygnacji z działań dyskrecjonalnych, dokonywanych w latach poprzednich, w prognozowanej wartości ogółem działań dyskrecjonalnych w roku n uwzględnia się wpływ na kwotę WYD_t dokonanych działań dyskrecjonalnych, jaki nastąpiłby w przypadku braku rezygnacji z tych działań w roku t.

[75] Chodzi o działania dyskrecjonalne, o których mowa w rozporządzeniu wymienionym w art. 104 ust. 1, pkt 1 w zakresie podatków i składek na ubezpieczenia społeczne, o których mowa w rozporządzeniu Parlamentu Europejskiego i Rady (UE) nr 549/2013 z dnia 21 maja 2013 roku w sprawie europejskiego systemu rachunków narodowych i regionalnych w Unii Europejskiej (DzU UE L 174 z dnia 26 czerwca 2013 roku, z późn. zm., s. 1).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 3.8. Schemat mechanizmu korygowania nierównowagi w ramach SRW



Źródło: na podstawie www.mf.gov.pl oraz www.finanse.mf.gov

Równanie (3.51) pokazuje, że limit wydatków na rok t jest równy kwocie wydatków zaplanowanych na rok bieżący ($t-1$), powiększoną o wielkość inflacji równą celowi inflacyjnemu, a także zwiększoną o średnie tempo wzrostu realnego PKB za okres 8 lat i skorygowaną o wielkość ewentualnej korekty wydatków (określonej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przez wielkość luki podażowej, wysokość deficytu i poziom zadłużenia). Równanie (3.52) pokazuje, że średnie tempo wzrostu PKB za okres 8-letni obliczone jest jako średnia geometryczna stóp wzrostu za 6 lat, w przypadku których dostępne są definitywne dane statystyczne, oraz prognozowanych stóp wzrostu za rok bieżący i rok następny.

Mechanizm SRW w znacznym stopniu eliminuje mankamenty tradycyjnych reguł fiskalnych. Po pierwsze albo ma on charakter antycykliczny, ponieważ korekta wydatków jest odpowiednio mniejsza przy niższym tempie wzrostu PKB, albo w ogóle nie funkcjonuje, gdy odchylenie (w dół) od średniookresowej dynamiki PKB wynosi co najmniej 2 punkty procentowe. Po drugie, SRW zapewnia elastyczność polityki fiskalnej poprzez uwzględnienie działań dyskrecjonalnych oraz uzależnienie rozmiaru korekty od skumulowanego salda sektora finansów publicznych – gdy suma odchyleń na koncie kontrolnym jest niższa niż 6 punktów procentowych, korekta może być mniejsza. Ponieważ SRW funkcjonuje dopiero od 2014 roku, z oceną skuteczności nowej reguły należy zaczekać jeszcze kilka lat[76].

3.10.6. Reguły fiskalne jednostek samorządu terytorialnego w Polsce

Do końca 2013 roku procedury budżetowe jednostek samorządu terytorialnego w Polsce były podporządkowane trzem regułom fiskalnym, których celem było utrzymanie dyscypliny finansowej na szczeblu władz lokalnych[77]. Po pierwsze, obowiązywała reguła zrównoważonego budżetu jednostek samorządu terytorialnego w części bieżącej. W myśl tej zasady organ stanowiący jednostki samorządu terytorialnego nie może uchwalić budżetu, w którym planowane wydatki bieżące są wyższe niż planowane dochody bieżące, powiększone ewentualnie o nadwyżkę budżetową z lat ubiegłych i wolne środki pochodzące z rozliczeń wyemitowanych papierów wartościowych oraz kredytów i pożyczek z lat ubiegłych[78]. Po drugie, wprowadzono limit ogólnego zadłużenia JST, którego nie można przekroczyć

[76] Już w drugim roku obowiązywania SRW Minister Finansów zetknął się z niemiłą niespodzianką w postaci deflacji, która miała miejsce w polskiej gospodarce w latach 2014–2016. Spowodowała ona obniżenie skorygowanej kwoty wydatków za 2014 rok (krok 1) i w ten sposób obniżyła również limit wydatków na 2015 rok. Deflacja stała się jedną z przyczyn nowelizacji ustawy o finansach publicznych w części dotyczącej mechanizmu SRW, przeprowadzonej w grudniu 2015 roku.

[77] Na podstawie *Gospodarka finansowa jednostek samorządu terytorialnego*, GUS, Warszawa, listopad 2014, s. 16–17.

[78] W przypadku realizacji zadań bieżących ze środków UE i EFTA zrealizowane wydatki bieżące mogą być wyższe niż uzyskane dochody bieżące, powiększone o nadwyżkę budżetową z lat ubiegłych i wolne środki związane z realizacją wydatków bieżących z udziałem środków UE i EFTA z lat ubiegłych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w poszczególnych kwartałach– wówczas wynosił on 60% dochodów[79]. Ustawowa granica nie obejmowała jednak zobowiązań powstałych w związku z realizacją programów i projektów finansowych z udziałem środków unijnych. Po trzecie, obowiązywał limit wydatków na obsługę zadłużenia, które nie mogły być wyższe niż 15% dochodów[80].

Od 2014 roku obowiązują nowe zasady ograniczania zadłużenia, określone w art. 243 ustawy o finansach publicznych[81]. W świetle tych zapisów jednostka samorządu terytorialnego nie będzie mogła uchwalić budżetu, którego realizacja spowoduje, że w roku budżetowym oraz w każdym roku następującym po danym roku budżetowym relacja łącznej kwoty przypadających na dany rok budżetowy wydatków na obsługę zadłużenia do planowanych dochodów budżetu ogółem przekroczy średnią arytmetyczną z obliczonych w odniesieniu do ostatnich trzech lat relacji jej dochodów bieżących, powiększonych o dochody ze sprzedaży majątku oraz pomniejszonych o wydatki bieżące, do dochodów budżetu ogółem.

W ustawie opisano tę relację następującym wzorem:

$$\left(\frac{R+O}{D}\right) \leq \left(\frac{Db_{n-1}+Sm_{n-1}-Wb_{n-1}}{D_{n-1}} + \frac{Db_{n-2}+Sm_{n-2}-Wb_{n-2}}{D_{n-2}} + \frac{Db_{n-3}+Sm_{n-3}-Wb_{n-3}}{D_{n-3}}\right), \tag{3.53}$$

gdzie użyte symbole oznaczają:

R – planowaną na rok budżetowy łączną kwotę z tytułu spłaty rat kredytów i pożyczek oraz wykupów papierów wartościowych wyemitowanych przez JST na cele określone w ustawie,

O – planowane na rok budżetowy odsetki od kredytów i pożyczek, odsetki i dyskonto od papierów wartościowych wyemitowanych przez JST na cele określone w ustawie oraz spłaty kwot wynikających z udzielonych poręczeń i gwarancji,

D – dochody budżetu ogółem w danym roku budżetowym,

Db – dochody bieżące,

Sm – dochody ze sprzedaży majątku,

[79] Wskaźnik zadłużenia (tzn. relacja zobowiązań do dochodów) w pierwszym, drugim i trzecim kwartale odnosi się do dochodów planowanych, w czwartym zaś – do uzyskanych w danym roku.

[80] Jeśli państwowy dług publiczny przekroczy 55% PKB, relacja ta obniżona zostanie do 12%.

[81] Na mocy ustawy z dnia 8 listopada 2013 roku o zmianie ustawy o finansach publicznych oraz niektórych innych ustaw złagodzona została tzw. reguła indywidualnego wskaźnika zadłużenia.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wb – wydatki bieżące,
n – rok budżetowy, wprzypadku którego ustalana jest relacja,
n – 1 – rok poprzedzający rok budżetowy, w przypadku którego ustalana jest relacja,
n – 2 – rok poprzedzający rok budżetowy o dwa lata,
n – 3 – rok poprzedzający rok budżetowy o trzy lata.

Reguła fiskalna opisana formułą (3.53) nakazuje, aby relacja kosztu obsługi zadłużenia do dochodów nie była większa niż (średnia) relacja nadwyżki dochodów nad wydatkami do samych dochodów.

3.10.7. Reguły czy instytucje? Rady polityki fiskalnej

Reguły fiskalne bywają często krytykowane za ich procykliczność, brak elastyczności i arbitralność ustalonych celów numerycznych (zob. np. Eichengreen, Hausmann i von Hagen 1999; Wyplosz 2001). Nawet gdy są dobrze skonstruowane, i przyczyniają się istotnie do wzmocnienia dyscypliny budżetowej, mogą powodować nadmierne koszty dostosowania w zmieniającej się sytuacji koniunkturalnej i zewnętrznej. Na przykład, w czasie recesji deficyt budżetowy powiększa się samorzutnie ze względu na spadek dochodów podatkowych i wzrost wydatków na ochronę socjalną (tzw. endogeniczny deficyt). Reguła fiskalna wymaga najczęściej w takiej sytuacji ograniczania wydatków publicznych, aby obniżyć deficyt do dozwolonego poziomu. Jednocześnie cięcia wydatków zmniejszają jednak zagregowany popyt i wpędzają gospodarkę w jeszcze głębszą recesję, a to z kolei o wiele bardziej obniża dochody budżetu i utrudnia równoważenie finansów publicznych.

Oprócz procykliczności, istotną wadą działania reguł fiskalnych pozostaje to, że jeśli ich celem jest stabilizacja finansów publicznych w długim okresie, to brakuje im niezbędnej elastyczności w krótkim okresie. Dlatego też w okresach słabej koniunktury, a zwłaszcza w czasie kryzysu, ograniczają one nadmiernie swobodę polityki fiskalnej w pobudzaniu produkcji i zatrudnienia. W efekcie sztywne reguły fiskalne mogą przyczyniać się do głębszego spowolnienia gospodarczego i przedłużania okresu recesji, co w rezultacie utrudnia proces przywracania równowagi budżetowej.

Potrzeba uwzględnienia krótkookresowej elastyczności przy jednoczesnym zapewnieniu długookresowej stabilności finansów publicznych skłania do poszukiwania alternatywnych rozwiązań. Zdaniem niektórych ekonomistów, lepszym rozwiązaniem byłoby powołanie specjalnych instytucji, czuwających nad dyscypliną

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

finansów publicznych. Takie niezależne instytucje fiskalne (*independent fiscal institutions* – IFI) bywają niekiedy nazywane radami polityki fiskalnej. Są to pozapartyjne, apolityczne ciała, inne niż ciała rządowe lub parlamentarne, choć finansowane z pieniędzy publicznych, przygotowujące analizy i prognozy makroekonomiczne dla celów budowy budżetu państwa, monitorujące sytuację finansów publicznych i formułujące rekomendacje odnośnie polityki fiskalnej.

Propozycja, aby powierzyć większy zakres odpowiedzialności za politykę gospodarczą ciałom niezależnym od instytucji politycznych i partyjnych, takich jak rząd czy parlament, pojawiła się po raz pierwszy w latach 70. w kontekście dyskusji nad zjawiskiem czasowej (dynamicznej) niespójności polityki pieniężnej. Zjawisko to, polegające na skłonności krótkowzrocznej władzy gospodarczej do dyskrecjonalnych działań i niedotrzymywania zobowiązań ze względu na doraźne cele polityczne, zostało po raz pierwszy omówione i wyjaśnione przez Kydlanda i Prescotta (1977)[82]. Wyniki przeprowadzonej przez nich analizy stały się silnym argumentem przemawiającym na rzecz powierzenia odpowiedzialności za politykę pieniężną niezależnym bankom centralnym.

Polityka fiskalna także cierpiała na niedostatki związane z upolitycznieniem i dyskrecjonalnością. Wśród głównych uchybień wymienia się najczęściej niespójność czasową związaną z niedostateczną świadomością długookresowych ograniczeń budżetowych, krótkowzroczność i przewagę doraźnych korzyści w podejmowaniu decyzji czy omówione wcześniej zjawisko politycznego cyklu wyborczego i problem wspólnego zasobu (zob. Calmfors 2010)[83]. W przypadku polityki fiskalnej dążenie do ograniczenia dyskrecjonalności nakierowało jednak najpierw uwagę ekonomistów na możliwość szerszego wykorzystania stabilnych reguł fiskalnych. Pozwalały one rzeczywiście uporać się z problemem niespójności czasowej i ograniczały dyskrecjonalność.

Lecz z upływem czasu reguły fiskalne ujawniły też swoje wady, takie jak procykliczność czy nadmierna sztywność. W związku z tym w latach 90. pojawiły się propozycje alternatywnych rozwiązań w postaci niezależnych instytucji fiskalnych. Instytucje takie funkcjonują obecnie w wielu krajach, tj. m.in. w USA (Kongresowe Biuro Budżetu – *Congressional Budget Office*), Holandii (Centralne Biuro Planowania – *Central Planning Bureau*), Szwecji (Szwedzka Rada Polityki Fiskalnej

[82] Problem niespójności czasowej polityki jest omówiony w rozdziale 5.6.2.
[83] Zobacz punkt 3.3.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

– *Swedish Fiscal Policy Council*), Belgii (Wysoka Rada Fiskalna – *High Council of Finance*) czy Danii (Rada Ekonomiczna – *Economic Council*).

Uważa się, że niezależne rady sprzyjają poprawie jakości polityki fiskalnej (Wyplosz 2005). W szczególności sporządzane przez nie fachowe opinie i prognozy dotyczące budżetu i polityki fiskalnej są z reguły bezstronne oraz pozbawione „urzędowego" optymizmu, charakterystycznego dla większości dokumentów rządowych. Istotną zaletą tych sprawozdań jest sygnalizowanie zagrożeń dla stabilności finansowej państwa, płynących ze zgłaszanych przez polityków nieodpowiedzialnych zapowiedzi zwiększenia wydatków czy obniżenia podatków. Niezależne instytucje fiskalne, działając na podstawie jasno określonego mandatu, wzbogacają rządowy punkt widzenia o bardziej długofalową analizę, uwzględniającą wymogi zachowania równowagi budżetowej w długim okresie.

IFI, czy też inaczej rady polityki fiskalnej, z reguły składają się z apolitycznych ekspertów, których zadaniem jest obiektywna, bezstronna analiza dokumentów rządowych dotyczących budżetu i polityki fiskalnej. Ich sformalizowane opinie powinny być, przynajmniej w pewnym zakresie, uwzględniane przez rząd w pracach nad budżetem. W zależności od przyjętych rozwiązań prawnych stanowisko rady polityki fiskalnej może mieć charakter ogłoszonej publicznie opinii, choć niewiążącej dla rządu, może też wymagać formalnej odpowiedzi ze strony rządu, udzielonej w określonym trybie, choć na ogół bez obowiązku uwzględniania tego stanowiska w dalszych pracach. Najdalej idące rozwiązanie zakłada wiążący dla rządu charakter opinii rady polityki fiskalnej. Przesądza ona wówczas o wielkości deficytu budżetowego bądź o tempie redukcji zadłużenia publicznego. W tym ostatnim przypadku rada ustala taki poziom deficytu (nadwyżki), który jest spójny z długookresową, bezpieczną relacją długu do PKB, zapewniającą stabilność finansową państwa.

W większości państw, w których funkcjonują IFI, przyjmują one charakter doradczy i opiniodawczy. Wyposażenie ich w dalej idące kompetencje pozostaje sprawą dyskusyjną z punktu widzenia demokratycznej odpowiedzialności (*democratic accountability*). Ponieważ IFI – w celu zapewnienia im politycznej niezależności – nie są wyłaniane w demokratycznych wyborach, a powoływane w trybie nominacji, ich mandat demokratyczny uprawniający do decydowania o gromadzeniu i podziale środków publicznych może być kwestionowany. Dlatego też uprawnienia IFI nie powinny z reguły obejmować zagadnień związanych z redystrybucyjną funkcją budżetu lub ze strukturą wydatków, a jedynie odnosić się do kwestii makroekonomicznych, takich jak wielkość deficytu czy poziom zadłużenia.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.11. Polityka konsolidacji finansów publicznych

Kryzys zadłużenia suwerennego w strefie euro w latach 2010–2013 postawił w centrum uwagi kwestię przywrócenia równowagi w finansach publicznych wśród krajów zadłużonych i odzyskania przez nie zdolności do zaciągania pożyczek na rynkach finansowych. W celu wypracowania właściwej polityki konsolidacji fiskalnej – czyli obniżenia deficytu – rządy tych krajów musiały pilnie odpowiedzieć przede wszystkim na trzy istotne pytania. Po pierwsze, jaka powinna być *struktura konsolidacji*, tj. czy ma ona polegać głównie na cięciu wydatków rządowych, a jeśli tak, to w jakich kategoriach wydatkowych, czy może raczej powinna koncentrować się na podnoszeniu podatków. Po drugie, jaka powinna być *sekwencja działań*, a więc czy najpierw należy dokonać zacieśnienia fiskalnego na rzecz jak najszybszego odzyskania równowagi budżetowej, czy raczej pobudzać gospodarkę poprzez stymulację fiskalną i dalsze przejściowe pogłębienie nierównowagi, aby odzyskać równowagę budżetową w późniejszym okresie, lecz mniejszym kosztem w postaci utraconego dochodu i zatrudnienia. Po trzecie wreszcie pytano, jakie powinno być *tempo konsolidacji*, a zatem czy niezbędne zacieśnienie fiskalne ma być dokonane szybko, czy może powinno być rozłożone w czasie. Te trzy kwestie zostaną kolejno omówione poniżej.

3.11.1. Struktura dostosowania fiskalnego: podnosić podatki czy obniżać wydatki?

Wspomnieliśmy już wcześniej, że, chcąc zmniejszyć deficyt budżetowy, rząd ma w zasadzie do wyboru dwie opcje – może ograniczyć wydatki lub podnieść podatki. Alesina i Perotti (1996) zbadali epizody konsolidacji fiskalnej na podstawie danych przekrojowych odnoszących się do 20 państw OECD w latach 1960–1994. Uzyskane wyniki wskazują, że lepsze i trwalsze efekty w przywracaniu zachwianej równowagi budżetowej daje ograniczanie wydatków rządowych niż podnoszenie podatków. Szczególnie skuteczna okazuje się w tym przypadku redukcja wydatków transferowych i kosztów zatrudnienia w sektorze publicznym. Autorzy wyjaśniają to tym, że alternatywny sposób konsolidacji, polegający na podnoszeniu podatków, łagodzi ograniczenia budżetowe i umożliwia wzrost wydatków w chwili, gdy koniunktura ulega poprawie. Poprawa fiskalna ma więc jedynie charakter tymczasowy i może być odwracalna. Co więcej, w krajach o wysokiej stopie fiskalizmu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

dalsze podnoszenie podatków nie tylko jest trudne z punktu widzenia polityki, lecz pociąga za sobą także duże koszty społeczne w postaci spowolnienia wzrostu na skutek nasilających się zniekształceń rynkowych (jest to tak zwany koszt jałowy podatku – zob. następny rozdział).

Wnioski te pozostają zgodne z wynikami wcześniejszych badań przeprowadzonych przez Giavazzi i Pagano (1990), którzy jako pierwsi analizowali skutki konsolidacji fiskalnej na konkretnym przykładzie dwóch zadłużonych krajów europejskich[84]. Udało im się wówczas wskazać, że kluczowe znaczenie w przypadku pobudzającego wpływu konsolidacji fiskalnej mają oczekiwania dotyczące poprawy sytuacji fiskalnej w przyszłości oraz ich wpływ na potencjalny poziom PKB. Również Perotti i in. (1998) potwierdzają te obserwacje, dodatkowo zwracając uwagę na to, że trwałe deficyty budżetowe najczęściej nie są wynikiem spadku dochodów podatkowych, a raczej wzrostu wydatków.

W kolejnym badaniu, przeprowadzonym 15 lat później, Alesina i Ardagna (2010) znajdują potwierdzenie dla wyników otrzymanych wcześniej przez Alesinę i Perottiego (1996). Autorzy analizują skutki dużych egzogenicznych zmian dokonujących się w polityce fiskalnej wśród grupy państw OECD w latach 1970–2007. Okazuje się, że w okresach recesji cięcia podatków przynoszą o wiele bardziej ekspansywne skutki niż zwiększanie wydatków. Z kolei w okresach konsolidacji fiskalnej ograniczanie wydatków ma zdecydowanie silniejszy oraz trwalszy wpływ na spadek zadłużenia i przywracanie równowagi budżetowej niż podnoszenie podatków. Z perspektywy zagregowanego popytu konsolidacja fiskalna staje się ekspansywna, gdy uczestnicy rynku wierzą w to, że zapewni ona wyraźną poprawę sytuacji fiskalnej oraz pozwoli uniknąć jeszcze silniejszego i bardziej kosztownego dostosowania fiskalnego w przyszłości. Ponadto, wiarygodna konsolidacja przekłada się na obniżenie rentowności papierów skarbowych, co powoduje spadek rynkowych stóp procentowych, obniża koszty obsługi długu i przyczynia się do wzrostu inwestycji. W przypadku zagregowanej podaży konsolidacja fiskalna oddziałuje poprzez wpływ obniżki wydatków na zwiększenie podaży pracy, zgodnie z założeniami neoklasycznymi[85].

Eichengreen i in. (2011) zbadali epizody skutecznej konsolidacji w krajach OECD i krajach strefy euro w latach 1999–2006. Uzyskane przez nich wyniki

[84] Chodzi o dostosowania fiskalne przeprowadzone w Danii i Irlandii w latach 80. XX wieku.

[85] Obniżenie wydatków rządowych wiąże się ze zmniejszeniem zatrudnienia w sektorze publicznym, a to zmniejsza presję na wzrost wynagrodzeń i, zgodnie z neoklasycznym modelem, skłania do wzrostu podaży pracy.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

potwierdzają hipotezę dotyczącą tego, że ograniczanie wydatków jest skuteczniejszą metodą konsolidacji fiskalnej. Wyniki tych ostatnich analiz zostały przedstawione w tabeli 3.11.

Tabela 3.11. Ciąć wydatki czy podnosić podatki? Epizody skutecznej konsolidacji budżetowej w krajach strefy euro (EA) i krajach OECD w latach 1999–2006 (udziały % w całości epizodów)

Dominujący kierunek działania	Kraje strefy euro (EA)	Kraje OECD (bez EA)
Ograniczenie wydatków	67	46
Podnoszenie podatków	33	31
Oba działania	0	23
Razem	100	100

Źródło: Eichengreen i in. (2011), s. 59.

Hipoteza o wpływie struktury konsolidacji na poziom produkcji została ponownie poddana ekonometrycznej weryfikacji przez Alesinę, Favero i Giavazziego (2012), którzy wykorzystali zgromadzone przez MFW dane dotyczące epizodów konsolidacji fiskalnej w 17 krajach OECD z lat 1978–2009 (zob. DeVries i in. 2011). Nowatorstwo tego badania polegało na odmiennym zdefiniowaniu wstrząsu fiskalnego – zamiast rzeczywistych danych opisujących zmiany poziomu wydatków (lub podatków) miernikiem wstrząsów stały się oficjalne rządowe zapowiedzi dokonania określonych zmian w wydatkach lub podatkach (tzw. wstrząsy narracyjne – *narrative shocks*). Autorzy argumentowali, że dostosowanie sektora prywatnego do zapowiedzi *ex-ante* jest lepszą miarą reakcji gospodarki niż adaptacja do nieoczekiwanych zmian, dokonywana *ex-post*. Otrzymane wyniki wskazują, że skutki konsolidacji zależą od jej struktury (podnoszenie podatków bądź obniżanie wydatków) oraz od spójności (trwałości) konsolidacji w czasie. Konsolidacji opartej na cięciu wydatków towarzyszyły łagodne recesje lub brak recesji w ogóle, z kolei w parze z konsolidacją opartą na podwyżce podatków szły recesje głębokie i długotrwałe. Z konsolidacją trwałą wiązały się zaś niższe koszty w postaci utraty produkcji niż miało to miejsce w przypadku konsolidacji krótkotrwałej lub przerywanej (zob. Alesina, Favero i Givazzi 2012).

Autorzy przeprowadzonego badania podkreślają, że rozbieżności dotyczące reakcji na oba typy konsolidacji są na tyle wyraźne, że nie da się ich wyjaśnić różnicami w towarzyszącej im polityce pieniężnej (która mogła być bardziej lub mniej akomodująca). Pozostają one przede wszystkim skutkiem odmiennej reakcji sektora

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

prywatnego na wstrząsy fiskalne – zwłaszcza reakcji na poziomie wydatków inwestycyjnych, w mniejszym zaś stopniu konsumpcyjnych. Również reakcja w postaci poprawy nastrojów przedsiębiorstw i konsumentów jest silniejsza w przypadku konsolidacji opartej na spadku wydatków niż na wzroście podatków.

Istotnym argumentem przemawiającym na rzecz konsolidacji opartej na ograniczaniu wydatków w przypadku strefy euro jest malejąca wraz z ich wzrostem krańcowa społeczna użyteczność wydatków publicznych. W krajach strefy euro wydatki publiczne w relacji do PKB należą do najwyższych na świecie, a ich malejąca użyteczność krańcowa może nie rekompensować rosnących kosztów społecznych, związanych z koniecznością podnoszenia podatków. Obserwacje te potwierdzają wyniki badań przeprowadzonych przez Afonso i in. (2003). Autorzy ci wskazują, między innymi, że spadek rozmiarów sektora publicznego w latach 1990–2002 w krajach o najwyższym udziale wydatków publicznych w PKB, takich jak Szwecja, Dania czy Belgia, był właśnie skutkiem podjęcia świadomych działań mających na celu ograniczenie zakresu nadmiernie rozbudowanego sektora publicznego. U podstaw tej zmiany leżało przekonanie, że w warunkach wysokiego udziału wydatków publicznych w PKB ich krańcowa społeczna użyteczność jest niewielka, a nawet może być ujemna.

Teza mówiąca o tym, że konsolidacja fiskalna powinna bardziej koncentrować się na stronie wydatkowej budżetu niż na jego stronie dochodowej, pośrednio znajduje potwierdzenie w wynikach badań nad wpływem wydatków i dochodów budżetowych na tempo wzrostu PKB. Na podstawie badań panelowych przeprowadzonych wśród szerokiej grupy 155 krajów rozwiniętych i rozwijających się, obejmujących lata 1970–2008, Afonso i Jalles (2013) dowodzą, że wzrost wydatków budżetowych ma statystycznie istotny ujemny wpływ na tempo wzrostu PKB na mieszkańca, z kolei związek tempa wzrostu PKB z dochodami budżetowymi jest generalnie znacznie słabszy. Otrzymane wyniki potwierdzają także znaną prawidłowość: wyższym poziomom PKB na mieszkańca odpowiadają większe udziały wydatków publicznych w PKB, przy czym związek przyczynowy przebiega od wzrostu PKB do wzrostu wydatków (a nie odwrotnie), co można interpretować jako pośrednie potwierdzenia działania omówionego wcześniej prawa Wagnera.

3.11.2. Sekwencja działań – co czynić najpierw: oszczędzać czy pobudzać gospodarkę?

Druga istotna kwestia dotyczy sekwencji działań na rzecz konsolidacji fiskalnej. Kryzys zadłużenia suwerennego oznacza, że – z punktu widzenia uczestników

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

rynków finansowych – poziom zadłużenia danego kraju przekroczył granicę bezpieczeństwa, gwarantującą wypłacalność państwa. Aby odzyskać zaufanie inwestorów, należy obniżyć poziom długu w relacji do PKB, a przynajmniej przekonać uczestników rynku, że zadłużenie nie będzie rosnąć, a raczej zacznie spadać. Oznacza to konieczność nie tylko ograniczenia bieżącego deficytu, lecz także wypracowania nadwyżki pierwotnej w wysokości umożliwiającej obsługę zadłużenia, zgodnie z równaniem (3.26). W przeciwnym razie kraj nie będzie w stanie zaciągać pożyczek na rynkach finansowych.

Lecz efekty makroekonomiczne takiej konsolidacji fiskalnej nie muszą być jednoznaczne. Zgodnie z tradycyjnym poglądem keynesowskim, konsolidacja polegająca na ograniczeniu wydatków i/lub podniesieniu podatków oznacza spadek zagregowanego popytu. Wynikający stąd spadek PKB nie pozwala na zmniejszenie relacji długu do PKB, a niekiedy nawet powoduje wzrost tej relacji. Zmusza to kraj do dalszego zacieśniania polityki fiskalnej, co z kolei prowadzi – poprzez mechanizm mnożnika wydatków – do dalszego spadku PKB (Blanchard i Leigh 2013a, 2013b; Eyraud i Weber 2013). Na skutek tych negatywnych sprzężeń zwrotnych konsolidacja fiskalna w czasie kryzysu może powodować głęboką i przewlekłą recesję, która utrudnia lub uniemożliwia wyjście z pułapki zadłużenia i wiąże się z ryzykiem całkowitego bankructwa. Zgodnie z tym poglądem, właściwą polityką okazuje się wówczas zastosowanie stymulacji fiskalnej, która – poprzez efekt mnożnika – wywoła wzrost zagregowanego popytu i uruchomi mechanizm wzrostu gospodarczego (zob. np. Corden 2010). Spowodowany stymulacją fiskalną wzrost bieżącego deficytu jest bowiem przejściowy, a w dalszej perspektywie samofinansujący, ponieważ wzrost PKB oznacza zwiększenie dochodów podatkowych, które pozwolą sfinansować pożyczki zaciągnięte w związku ze wzrostem wydatków. Mechanizm ten umożliwia więc wyjście z pułapki zadłużenia dzięki efektom mnożnika fiskalnego i wzrostowi PKB.

Wielu ekonomistów odnosi się jednak krytycznie do tego typowo keynesowkiego poglądu o „wychodzeniu" z pułapki zadłużenia dzięki dalszej ekspansji fiskalnej. W ramach poglądu alternatywnego, nazywanego niekiedy podejściem „niemieckim" (Fels i Froehlich 1986, Hellwig i Neumann 1987) lub koncepcją „ekspansywnego zacieśnienia fiskalnego" (*expansionary fiscal contraction*), kładzie się nacisk na mechanizm oczekiwań i wskazuje na negatywny wpływ wydatków publicznych na inwestycje prywatne. W krótkim okresie konsolidacja fiskalna wpływa wprawdzie bezpośrednio na spadek zagregowanego popytu, ale pośrednio przyczynia się też do wzrostu popytu poprzez korzystną zmianę oczekiwań, jeśli konsolidacja

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jest traktowana przez uczestników rynku jako trwałe obniżenie udziału sektora publicznego w PKB. Spadek wydatków i/lub wzrost podatków oznacza zmniejszenie potrzeb pożyczkowych rządu, a zatem także zmniejszenie przyszłych obciążeń podatkowych wśród konsumentów i przedsiębiorstw (efekt ten jest uwzględniany w decyzjach uczestników rynku w takim stopniu, w jakim obowiązuje równoważność rikardiańska), co przywraca zaufanie na rynkach. Wzrost tego czynnika przekłada się na spadek rentowności obligacji rządowych i rynkowych stóp procentowych, a to z kolei pobudza konsumpcję i inwestycje. Wynikający stąd wzrost gospodarczy ułatwia wyjście z pułapki zadłużenia.

Działanie mechanizmu „ekspansywnego zacieśnienia fiskalnego" potwierdziły wyniki badań przeprowadzonych m.in. przez Giavazziego i Pagano (1990), Blancharda (1990), Alesinę i Perottiego (1998), oraz Barry'ego i Devereux (2003). Z kolei Alesina i in. (1999), odwołując się do danych panelowych zebranych wśród krajów OECD, dowodzą, że wzrost wydatków publicznych ma silnie negatywny wpływ na inwestycje prywatne poprzez presję na wzrost wynagrodzeń w sektorze prywatnym, powstającą na skutek powiększania zatrudnienia w sektorze publicznym. Co więcej, to ujemne oddziaływanie jest silniejsze niż ujemny wpływ wzrostu podatków, niezbędny w celu sfinansowania wzrostu wydatków.

Spór między zwolennikami i przeciwnikami keynesowskiego pobudzania fiskalnego ciągnie się od lat. Pogląd keynesowski załamał się w latach 70. – z jednej strony na skutek zderzenia ze zjawiskiem stagflacji, z drugiej zaś pod wpływem krytyki ze strony przedstawicieli nowych nurtów ekonomii, w tym monetaryzmu i teorii racjonalnych oczekiwań. Od tamtej pory przeważało przekonanie, że stymulacja fiskalna finansowana długiem, zwłaszcza wśród krajów wysoko zadłużonych, wywołuje głównie efekty „niekeynesowskie", tj. nie przyspiesza wzrostu gospodarczego, a nawet może prowadzić do spadku popytu. Dzieje się tak nie tylko dlatego, że dana stymulacja pozwala oczekiwać podwyżki podatków w przyszłości w celu spłaty długu (efekt rikardiański), lecz także dlatego, że jej pobudzający wpływ na popyt jest z reguły neutralizowany równoległym wzrostem restrykcyjności polityki pieniężnej. Stymulacja fiskalna przy danym poziomie PKB oraz inflacji spowoduje bowiem zaostrzenie polityki pieniężnej, gdy niezależny bank centralny uzna, że może ona doprowadzić do wzrostu inflacji ponad poziom spójny z celem stabilności cen. W takich warunkach stymulacja fiskalna jest nieskuteczna, a mnożnik wydatków fiskalnych pozostaje zerowy (zob. wcześniejszy punkt 3.9.8).

W kontekście kryzysu zadłużenia suwerennego w niektórych krajach strefy euro spór o rolę i skuteczność stymulacji fiskalnej powrócił do głównego nurtu debaty

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ekonomicznej (zob. Krugman 2012). Pamiętając, że w chodzi ostatecznie o zredukowanie relacji długu do PKB, problem można też postawić następująco: czy, dążąc do obniżenia tej relacji, należy głównie zmniejszać licznik ułamka (zadłużenie), czy zwiększać jego mianownik (PKB). Uzyskanie jednoznacznego rozstrzygnięcia w tej kwestii jest prawdopodobnie niemożliwe. Można bowiem wykazać, że to, czy w warunkach kryzysu i dużego zadłużenia stymulacja fiskalna powinna, czy nie powinna mieć miejsca, w dużym stopniu zależeć będzie od konkretnej sytuacji danego kraju. W szczególności zwraca się uwagę na trzy czynniki, od których zależy skuteczność stymulacji fiskalnej.

Po pierwsze, stymulacja może mieć sens w sytuacji, gdy polityka pieniężna pozostaje nieskuteczna ze względu na to, że stopy procentowe zostały obniżone do poziomów bliskich zeru i bank centralny nie jest w stanie złagodzić polityki pieniężnej, mimo że uważa, iż byłoby to uzasadnione. W takiej sytuacji wzrost wydatków budżetowych nie wywołuje zaostrzenia polityki pieniężnej, co z kolei zwiększy wartość mnożnika wydatków (Christiano i in. 2009, Corsetti i in. 2010). Dodatkowo, podobny skutek ma istnienie dużych niewykorzystanych mocy produkcyjnych. Wyższa wartość mnożnika powoduje, że wzrost deficytu budżetowego zostanie sfinansowany *ex-post* z dodatkowych dochodów podatkowych, generowanych dzięki przyspieszeniu wzrostu gospodarczego (DeLong i Summers 2012).

Drugim istotnym argumentem przemawiającym na rzecz stymulacji fiskalnej w okresie kryzysu są efekty tak zwanej histerezy na rynku pracy, związane z przewlekłym bezrobociem, utratą kwalifikacji przez długotrwale bezrobotnych i wpływem tych zjawisk na obniżenie potencjalnego poziomu PKB. Można wykazać, że tolerowanie dużego bezrobocia przez dłuższy czas trwale obniża potencjalną wartość PKB (DeLong i Summers 2012). Z tego punktu widzenia stymulacja fiskalna, stabilizująca poziom produkcji i zatrudnienia, pozwala uniknąć szkodliwego efektu histerezy na rynku pracy.

Po trzecie, wiele zależy też od tego, czy wzrost wydatków rządowych ma charakter trwały czy przejściowy. Stałe zwiększanie wydatków rządowych okaże się bardziej skuteczne w przypadku pobudzania wzrostu, ponieważ uczestnicy rynku nie będą oczekiwać, że po przejściowym okresie zwiększonych wydatków nastąpi zacieśnienie fiskalne w celu przywrócenia równowagi w zakresie finansów publicznych. Można bowiem wykazać, że samo oczekiwanie konsolidacji fiskalnej po uprzedniej stymulacji przyniesie skutki recesyjne, ponieważ powoduje spadek oczekiwanej inflacji, a to oznacza zaś wzrost realnych stóp procentowych i w konsekwencji spadek popytu (Christiano i in. 2009, Corsetti i in. 2010).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Oddziaływanie tych czynników przebiega odmiennie w różnych krajach. O ile można przyjąć, że histereza daje generalnie podobne efekty w poszczególnych krajach, o tyle polityka pieniężna kształtuje się inaczej. W szczególności, jeśli w krajach, które cieszą się zaufaniem rynków finansowych, stopy procentowe – zarówno krótko-, jak i długookresowe – utrzymują się na bardzo niskim poziomie, oznacza to, że stymulacja fiskalna nie jest neutralizowana zaostrzeniem polityki pieniężnej i w rezultacie mnożniki fiskalne przyjmują relatywnie wysokie wartości. W tych krajach wyrażenie ($y - i$) pozostaje dodatnie, nawet przy wysokim poziomie zadłużenia, a polityka stymulacji fiskalnej okazuje się skuteczna w ograniczaniu recesji i pobudzaniu wzrostu gospodarczego. Przykładem takiego kraju były Stany Zjednoczone oraz – w nieco mniejszym stopniu – Japonia i Wielka Brytania. Wśród większości pozostałych państw zadłużonych stopy procentowe okazały się jednak wyższe, zwłaszcza rynkowe stopy oprocentowania długu. Oznacza to, że mnożniki fiskalne są niższe, a wyrażenie ($y - i$) pozostaje ujemne. W takich warunkach stymulacja fiskalna jest nieskuteczna i bardzo szybko napotyka na barierę finansowania.

Zdolność banku centralnego do utrzymywania niskich stóp procentowych, tak aby nie neutralizować skutków stymulacji fiskalnej, zależy od stopnia zaufania rynków finansowych. Kraj cieszący się dużym zaufaniem (kraj N na rysunku 3.3) może zadłużać się w większym stopniu i przy niższych stopach procentowych niż kraj o niższym zaufaniu (kraj M na rysunku 3.3). Wysoki poziom zaufania wywodzi się najczęściej z przekonania, że dany kraj w zasadzie nie może zbankrutować. Przeświadczenie to może wynikać z kilku powodów. Po pierwsze, dług danego kraju w całości bądź przynajmniej w dominującej części jest nominowany w walucie krajowej, której podaż jest w pełni kontrolowana przez dane państwo. Tego typu okoliczność występuje wśród krajów największych i najbardziej rozwiniętych, a także tych, których waluta pełni funkcję międzynarodowej waluty rezerwowej (USA, Wielka Brytania, Japonia, Szwajcaria)[86]. Wspólna waluta europejska stanowi również międzynarodową walutę rezerwową, lecz w przeciwieństwie do innych krajów, poszczególne kraje strefy euro nie kontrolują podaży euro – uprawnienia w tym zakresie posiada wyłącznie Europejski Bank Centralny. Z tego powodu zaufanie rynków finansowych do tych krajów będzie niższe. Po drugie, przekonanie o fundamentalnej wypłacalności kraju ma miejsce wówczas, gdy dług równoważony

[86] Kraje kontrolujące podaż własnej waluty mogą teoretycznie dostarczyć wszelką wymaganą ilość swojej waluty, aby zrealizować zapadające zobowiązania płatnicze. Kontrola nad podażą własnej waluty pozwala im też na obniżanie realnej wartości długu poprzez dopuszczenie do wyższej inflacji („inflacyjne" wyjście z długu).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jest wysokim poziomem wartościowych aktywów będących w posiadaniu danego kraju. Taki stan rzeczy występuje w krajach dysponujących dużymi rezerwami walut obcych (Chiny, Japonia), a także w krajach posiadających relatywnie duże zasoby bogactw naturalnych (Norwegia, kraje Zatoki Perskiej). Po trzecie wreszcie, ważne jest, aby rząd danego kraju posiadał wysoką zdolność opodatkowania większości swoich wierzycieli. Okoliczność ta wynika z udziału długu w posiadaniu krajowych rezydentów w całości długu – ponieważ rząd ma możliwość opodatkowania własnych obywateli i własnych przedsiębiorstw, wysoki udział rezydentów w strukturze podmiotowej długu zmniejsza prawdopodobieństwo niewypłacalności kraju. Wszystkie te czynniki występują w wielu krajach w różnym zakresie, co rzutuje na poziom zaufania rynków finansowych, a w konsekwencji przekłada się także na zróżnicowany poziom stóp procentowych.

Kolejny argument przemawiający na rzecz stymulacji fiskalnej w warunkach kryzysu spowodowanego nadmiernym zadłużeniem przedstawili Eggertson i Krugman (2012). Podkreślają oni, że interwencja rządu może być konieczna ze względu na asymetrię zachowań zadłużonych i niezadłużonych (lub bardziej zadłużonych i mniej zadłużonych) uczestników rynku. Przykładowo, gdy zadłużeni uczestnicy rynku ograniczają wydatki ze względu na konieczność „delewarowania" w celu naprawy swoich bilansów, niedobór popytu mogliby uzupełnić niezadłużeni uczestnicy rynku. W praktyce tak się jednak nie dzieje i mniej zadłużeni uczestnicy rynku generalnie nie zwiększają swoich wydatków w okresach kryzysowych. Właśnie ta asymetria zachowań powoduje spadek zagregowanego popytu. Dlatego też rządy poszczególnych państw powinny w takiej sytuacji więcej pożyczać i więcej wydawać, aby uzupełnić niedobór popytu ze strony sektora prywatnego, szczególnie wówczas, gdy sektor prywatny stara się obniżyć poziom swego zadłużenia kosztem bieżących wydatków. Rozumowanie Eggertsona i Krugmana nie uwzględnia jednak faktu, że możliwości zwiększania popytu ze strony sektora publicznego w krajach zadłużonych są niewielkie, ponieważ zdolność pożyczkowa rządów w tych krajach jest ograniczona. Przy dużym zadłużeniu i w czasie kryzysu, rząd może po prostu nie być w stanie przekonać potencjalnych inwestorów co do tego, że będzie mógł obsłużyć wyższy poziom zadłużenia.

Jak brzmi zatem poprawna odpowiedź na pytanie postawione w tytule niniejszego podrozdziału? Stymulacja fiskalna będzie wskazana wtedy, gdy polityka pieniężna nie będzie neutralizować skutków wzrostu wydatków i gdy recesja okaże się stosunkowo głęboka, ponieważ wówczas mnożniki wydatków fiskalnych pozostają relatywnie wysokie. Stymulacja będzie także wskazana w warunkach

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

występowania znaczących efektów histerezy na rynku pracy oraz w warunkach silnej asymetrii zachowań po stronie zadłużonych i niezadłużonych uczestników rynku. Gdy te warunki nie zostaną jednak spełnione lub gdy z innych powodów mnożniki wydatków nie są wysokie, wskazane jest wówczas raczej zacieśnienie fiskalne. Natomiast konsolidacja fiskalna w czasie kryzysu zadłużenia okazuje się szczególnie skuteczna w obniżaniu relacji długu do PKB, gdy redukcja deficytu powstaje na skutek ograniczenia wydatków i jest równocześnie wspierana przez reformy strukturalne, zwiększające potencjalne tempo wzrostu PKB. Na koniec należy dodać, że stymulacja fiskalna może być nieosiągalna w przypadku państwa, które ze względu na bardzo wysokie zadłużenie utraciło zaufanie rynków finansowych i nie ma dostępu do pożyczek. Wówczas konsolidacja fiskalna pozostaje jedyną możliwą opcją polityki.

3.11.3. Tempo konsolidacji fiskalnej – co jest korzystniejsze: więcej konsolidacji teraz i mniej później czy odwrotnie?

Gdy dany kraj uzna, że z jakiegoś powodu zastosowanie stymulacji fiskalnej nie jest możliwe (lub nie jest celowe) oraz że przywrócenie równowagi budżetowej wymaga zacieśnienia fiskalnego, powstaje pytanie o tempo konsolidacji. Czy należy ograniczyć deficyt szybko, czy raczej rozłożyć działania oszczędnościowe na dłuższy okres? Innymi słowy, czy korzystniej jest przeprowadzić konsolidację na większą skalę teraz, a w mniejszym stopniu później, czy odwrotnie?

Odpowiedź na to pytanie jest w zasadzie zbliżona do odpowiedzi na pytanie postawione w poprzednim punkcie rozdziału. Zacznijmy od czynników przemawiających na rzecz rozłożenia konsolidacji w czasie. Ogólnie rzecz biorąc, gdy mnożniki fiskalne są w danym momencie stosunkowo wysokie, skoncentrowana konsolidacja polegająca na cięciu wydatków i/lub podwyżce podatków będzie miała silny krótkookresowy negatywny wpływ na produkcję i zatrudnienie oraz zaledwie niewielki wpływ na relację długu do PKB. Gdy spadek produkcji okaże się większy niż spadek zadłużenia (do czego konieczna byłaby nadwyżka pierwotna), relacja ta ulegnie dalszemu pogorszeniu.

Wielu autorów zauważa, że w czasie ostatniego kryzysu (lata 2009–2013) mnożniki fiskalne były stosunkowo wysokie (Eyraud i Weber 2013, Blanchard i Leigh 2013a), i to z kilku powodów. Po pierwsze, w warunkach nominalnych stóp procentowych bliskich zeru nie należy oczekiwać szybkiego zaostrzenia polityki pieniężnej ze strony banków centralnych. Po drugie, jak wskazują Eggertson

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i Krugman (2012), trudniejsza dostępność kredytu bankowego (spowodowana koniecznością pokrycia strat i poprawienia bilansów przez banki) oznacza, że obecne wydatki konsumpcyjne zależą w większym stopniu od bieżących raczej niż przyszłych dochodów, podczas gdy inwestycje w większym stopniu pozostają zależne od bieżących raczej niż przyszłych zysków. W takich warunkach konsolidacja silnie ogranicza bieżące wydatki. Po trzecie, również wyniki wielu badań empirycznych potwierdzają, że mnożniki stają się wyższe w warunkach głębszej recesji (Auerbach i Gorodnichenko 2012).

Kolejnym argumentem przemawiającym na rzecz rozłożenia konsolidacji w czasie jest ryzyko wepchnięcia gospodarki w recesję. Zacieśnienie fiskalne ma tym silniejszy negatywny wpływ na produkcję, im niższy osiąga ona poziom, ponieważ przy niewielkim poziomie produkcji (lub w recesji) popyt prywatny staje się także niski. Redukcja deficytu budżetowego w warunkach słabego popytu prywatnego okazuje się więc generalnie bardziej kosztowna dla gospodarki (Sheets 2011, Ho i Yetman 2013).

Za rozłożeniem konsolidacji w czasie przemawia też omówiony wyżej efekt histerezy. Jak dowodzą DeLong i Summers (2012), natychmiastowa konsolidacja w warunkach recesji spowalnia wzrost gospodarczy w długim okresie, ponieważ rośnie wówczas udział bezrobocia strukturalnego w całej populacji osób bezrobotnych. To z kolei obniża potencjalne tempo wzrostu PKB.

Jakie argumenty przemawiają natomiast na rzecz ograniczenia deficytu od razu? Zasadniczym motywem przeprowadzenia szybkiej konsolidacji jest chęć przywrócenia zaufania na rynkach i odzyskania dostępu do finansowania poprzez obniżenie ryzyka związanego z wysokim zadłużeniem w relacji do PKB. Wysoki poziom zadłużenia oddziałuje negatywnie na wzrost gospodarczy kilkoma kanałami. Po pierwsze, oznacza konieczność wprowadzenia wyższych podatków w przyszłości, co zniekształca bodźce, obniża wydatki i powoduje utratę dobrobytu (zob. Ostry i in. 2015). Po drugie, wysoki poziom zadłużenia zmniejsza zaufanie do dłużnika, co przekłada się na wzrost premii za ryzyko niewypłacalności w rynkowych stopach procentowych, a to z kolei ogranicza kredyt i inwestycje sektora prywatnego (zob. Zoli 2013). Po trzecie, wysoki poziom zadłużenia zwiększa ryzyko wpadnięcia gospodarki w „złą" równowagę w warunkach istnienia „wielorakich równowag" (*multiple equilibria*) na rynku długu, co grozi kryzysem niewypłacalności[87].

[87] Mechanizm „wielorakich równowag" na rynku finansowym został omówiony w rozdziale 5 (patrz punkt 5.14.5).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W sytuacji ograniczonego zaufania nawet niewielkie pogorszenie nastrojów może wywołać załamanie na rynku długu poprzez mechanizm samospełniających się obaw. Gospodarka wpada w „złą" równowagę, co oznacza, że wyprzedaż obligacji powoduje spadek cen długu, a to z kolei przynosi dalszą wyprzedaż obligacji. W takiej sytuacji wskazana jest jak najszybsza konsolidacja, aby wyprowadzić gospodarkę ze strefy zagrożenia kryzysem (zob. więcej na ten temat w punkcie 3.8.4).

Kluczowe znaczenie ma więc w tym przypadku odbudowa zaufania i przywrócenie wiarygodności. Radykalne obniżenie długu pomaga odzyskać zaufanie, lecz jednocześnie może być trudne do przeprowadzenia w krótkim czasie. Działaniom konsolidacyjnym powinno więc towarzyszyć wiarygodne zobowiązanie do dokonania trwałej poprawy w finansach publicznych. Wówczas efekty „niekeynesowskie" będą stosunkowo silne, a spadkowi popytu rządowego będzie towarzyszyć wzrost popytu prywatnego. W takich warunkach szybka konsolidacja okazuje się korzystna.

Podsumowując, gdy mnożniki fiskalne są w danym momencie wyższe niż w normalnych warunkach, gdy zagrożenie niskim wzrostem (lub recesją) jest w danym momencie duże, gdy popyt sektora prywatnego okazuje się w danym momencie słaby i gdy efekty histerezy są silne, konsolidacja fiskalna powinna być raczej rozłożona w czasie. I odwrotnie, gdy mnożniki fiskalne pozostają w danym momencie niskie (a przynajmniej nie wyższe niż normalne), gdy popyt prywatny jest silny oraz gdy zagrożenie kryzysem pozostaje wysokie i występuje duże ryzyko „złej" równowagi, konsolidacja powinna być przeprowadzona szybko.

3.11.4. Tempo konsolidacji fiskalnej – próby weryfikacji statystycznej w przypadku strefy euro

DeLong i Summers (2012) argumentują, że sytuacja panująca w gospodarce Stanów Zjednoczonych w latach 2008–2012 uzasadniała stymulację fiskalną w celu przezwyciężenia recesji i przyspieszenia wzrostu, ponieważ zdolność polityki pieniężnej do pobudzania gospodarki była w tym czasie silnie ograniczona ze względu na to, że stopy procentowe pozostawały na poziomie praktycznie równym zeru, a zarazem istniały znaczne wolne moce produkcyjne. Ponadto USA kontrolują podaż swojej waluty, w której denominowane są zobowiązania rządowe, co zasadniczo zmniejsza ryzyko niewypłacalności. W takich warunkach wartości mnożnika wydatków budżetowych pozostają wysokie i stymulacja fiskalna okazuje się skuteczna. W efekcie dodatkowe wydatki budżetowe finansują się same dzięki wzrostowi dochodów podatkowych uzyskanych wskutek przprowadzenia stymulacji

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w warunkach stóp procentowych bliskich zeru. Choć nie dysponujemy podobnym opracowaniem w przypadku krajów strefy euro, można zakładać, że w większości z nich nie występują tak sprzyjające warunki, jak ma to miejsce w USA. Kraje te nie kontrolują podaży swojej waluty (ta znajduje się bowiem pod nadzorem EBC), a rynkowe stopy procentowe w większości z nich nie były w czasie kryzysu bliskie zeru.

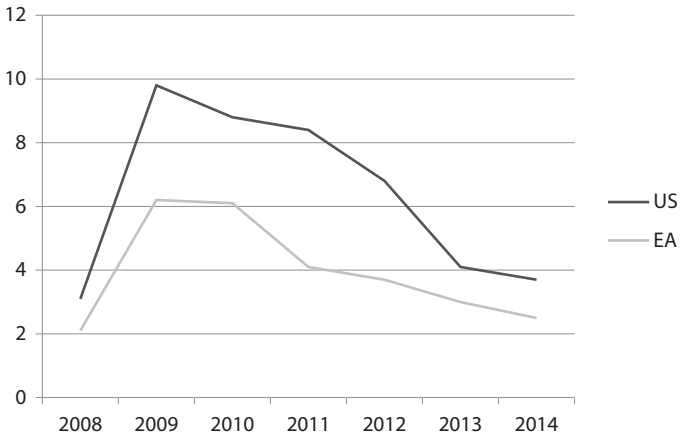
Odpowiedzią krajów strefy euro na kryzys zadłużenia suwerennego w latach 2009–2012 była przede wszystkim konsolidacja fiskalna. Zgodnie z postanowieniami unijnego Paktu Stabilności i Wzrostu, kraje, w których deficyt finansów publicznych przekroczył poziom 3% PKB, wdrożyły programy oszczędnościowe w ramach tzw. procedury nadmiernego deficytu. Ponadto, kraje o bardzo wysokim zadłużeniu publicznym w relacji do PKB nie miały praktycznie innego wyjścia. Ponieważ utraciły zdolność zaciągania pożyczek na rynkach finansowych (przynajmniej po umiarkowanych stopach procentowych) ze względu na wysokie ryzyko niewypłacalności (Grecja, Irlandia, Portugalia), musiały zwrócić się o pomoc finansową do pozostałych krajów strefy euro. Wsparcie to zostało im udzielone pod warunkiem przeprowadzenia konsolidacji fiskalnej (oraz innych reform), tak aby krajom tym udało się przywrócić równowagę w zakresie finansów publicznych i – po odzyskaniu zaufania wśród inwestorów – powrócić na rynki finansowe.

Wykres 3.16 odzwierciedla przebieg polityki fiskalnej w okresie kryzysu 2008–2014 w USA i w strefie euro, podczas gdy wykres 3.17 ilustruje zmiany, jakie w tym czasie dokonały się w poziomie PKB obu gospodarek. Przedstawione dane dowodzą, że w obu przypadkach miał miejsce wzrost deficytu budżetowego, skala tego zjawiska była jednak znacznie większa w przypadku USA. Wyniakło to stąd, że oprócz cyklicznego efektu spadku dochodów budżetowych rząd amerykański zastosował także dyskrecjonalne działania zwiększające wydatki i zmniejszające podatki w ramach programu ARRP (*American Recovery and Reinvestment Plan* – zob. ramka 3.11) na łączną sumę równą około 5% PKB, w strefie euro skala stymulacji fiskalnej była natomiast znacznie mniejsza. W rezultacie recesja w USA okazała się płytsza i gospodarka amerykańska szybciej powróciła na ścieżkę solidnego wzrostu gospodarczego. Z kolei znacznie ostrożniejszej polityce fiskalnej w strefie euro towarzyszyła głębsza recesja i wolniejsze tempo wychodzenia ze stagnacji[88].

[88] Należy zaznaczyć, że poza ekspansywną polityką fiskalną do szybszego przezwyciężenia recesji w USA przyczyniła się także bardzo ekspansywna polityka pieniężna. Ban centralny USA obniżył stopy procentowe prawie do zera już pod koniec 2008 roku i rozpoczął skup papierów wartościowych w ramach polityki „luzowania ilościowego" (*quantitative easing*) na początku 2009 roku, podczas gdy Europejski

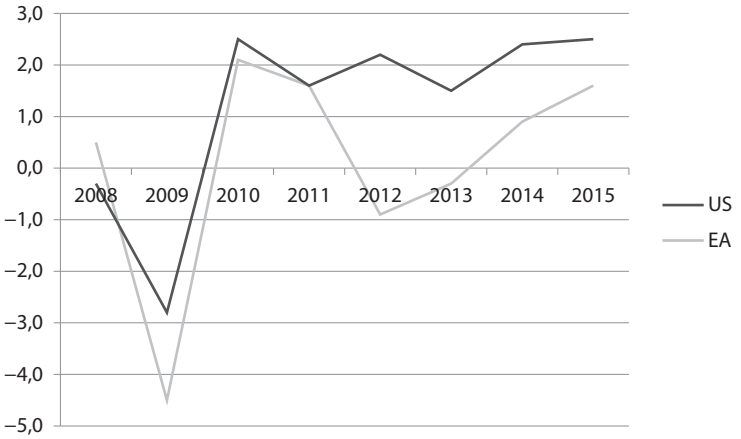
Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 3.16. Deficyt sektora publicznego w USA i w strefie euro w latach 2008–2014, (dane w % PKB)



Źródło: opracowanie własne na podstawie danych Eurostatu.

Wykres 3.17. Tempo wzrostu PKB w USA i w strefie euro w latach 2008–2015 (dane w %)



Źródło: opracowanie własne na podstawie danych Eurostatu.

Bank Centralny prowadził o wiele ostrożniejszą politykę i dopiero w 2014 roku zastosował zerowe stopy procentowe i zdecydował się na skup aktywów (zob. rozdział 5).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Zadłużone kraje ze strefy euro podjęły – w różnym stopniu – wysiłek konsolidacyjny, starając się ograniczyć wielkość deficytu budżetowego i zatrzymać przyrost zadłużenia. Recesja w tych krajach okazała się jednak głębszai dłuższa nic początkowo zakładano[89]. Tabela 3.12 pokazuje, że obniżaniu deficytu w krajach zadłużonych towarzyszyły głównie spadki PKB. Jedynie w Irlandii poziom PKB wzrósł przy jednocześnie bardzo wyraźnym obniżeniu deficytu o ponad 1/4 PKB. Najbardziej drastyczne załamanie produkcji miało miejsce w Grecji (o 17,5%, a w okresie 2007–2013 nawet o 26,4%), i to w warunkach braku jakiejkolwiek konsolidacji fiskalnej (deficyt budżetowy w istocie zwiększył się w latach 2010–2013 z 11,1% do 12,2% PKB). Jednocześnie we wszystkich krajach nastąpił w tym okresie wzrost zadłużenia w relacji do PKB – od około 15% PKB we Włoszech do 60% PKB w Irlandii.

Tabela 3.12. Konsolidacja fiskalna i tempo wzrostu PKB wśród zadłużonych krajów strefy euro w latach 2010–2013

Kraj	Skumulowana zmiana salda sektora publicznego (% PKB)[a]	Poziom PKB w 2013 roku (2010=100)	Dług publiczny na koniec 2009 roku (% PKB)	Dług publiczny na koniec 2013 roku (% PKB)
Irlandia	26,7	107,5	62,2	123,3
Grecja	–1,1	82,5	126,8	174,9
Hiszpania	2,6	97,4	52,7	92,1
Włochy	1,4	95,7	112,5	127,9
Portugalia	6,3	94,5	83,6	128,0

a) Saldo sektora instytucji rządowych i samorządowych (*net lending/borrowing of general government*), znak (–) oznacza wzrost deficytu budżetowego.

Źródło: opracowanie własne na podstawie danych Eurostatu.

Po kilku latach od momentu wystąpienia kryzysu powróciło więc pytanie o to, czy konsolidacja fiskalna była prawidłową reakcją na recesję w warunkach wysokiego zadłużenia w niektórych państwach strefy euro. Z jednej strony pozostaje rzeczą oczywistą, że oszczędności budżetowe i podwyżki podatków, mające na celu ograniczenie deficytu, zmniejszają zagregowany popyt i muszą wiązać się z recesją. Z drugiej zaś strony dalszy wzrost wydatków rządowych prowadzący do

[89] Badania nad skutkami działań oszczędnościowych (*austerity measures*) w pierwszym okresie kryzysu przeprowadzone przez Międzynarodowy Fundusz Walutowy wskazują, że skutki recesyjne tych działań mogły okazać się silniejsze niż skutki ekspansywne (Guajardo i in. 2011).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

postępującego wzrostu, i tak już wysokiego, zadłużenia może również pogłębiać recesję na skutek wystąpienia efektów antykeynesowskich (zagadnienie to zostało szczegółowo omówione w poprzednich podrozdziałach). Na koniec pojawia się też oczywiście praktyczne pytanie o możliwości sfinansowania tego wzrostu wydatków wobec braku dostępu do rynków finansowych lub przy bardzo wysokich stopach procentowych.

Badania prowadzone przez ekonomistów z Europejskiego Banku Centralnego potwierdzają zasadniczo ujemną zależność pomiędzy poziomem zadłużenia a skutkami stymulacji fiskalnej. Nickel i Tudyka (2013) testowali wpływ stymulacji fiskalnej na wzrost PKB przy różnych poziomach zadłużenia publicznego wśród 17 krajów europejskich w latach 1970–2010. Otrzymane wyniki wskazują, że o ile przy niskich i umiarkowanych poziomach zadłużenia publicznego wpływ stymulacji fiskalnej na wzrost gospodarczy jest dodatni oraz istotny, o tyle przy wyższych poziomach zadłużenia wpływ ten staje się ujemny, przy czym badana zależność ma wyraźnie nieliniowy charakter. Wnioski te są zgodne z wynikami otrzymanymi podczas wcześniejszych badań (np. Afonso i in. 2010, Checherita i Rother 2010), które wyraźnie wskazywały, że wpływ impulsu fiskalnego na produkcję w znacznym stopniu zależy (czyli występuje jako *state-dependent*) od stanu zadłużenia i koniunktury danej gospodarki. Oznacza to, że próby przeprowadzenia stymulacji fiskalnej przy wysokim poziomie zadłużenia prowadziłyby do pogorszenia, a nie do poprawy sytuacji.

Mniej jednoznaczne okazują się wyniki badań dotyczących wielkości i znaku mnożników fiskalnych w czasie recesji. Na przykład Taylor (1993) oraz Cogan i in. (2009) szacują, że nawet w okresie recesji mnożniki fiskalne są wyraźnie mniejsze od jedności, z kolei Auerbach i Gorodnichenko (2012) wykazują, że mnożniki podczas recesji stają się wyraźnie wyższe niż zazwyczaj. Starając się rozstrzygnąć tę kontrowersję, Gechert i Ranneberg (2014) oraz Gechert, Hallett i Ranneberg (2015) przeanalizowali 98 opracowań (z lat 1997–2013), których autorzy badali wpływ szoków fiskalnych na zmienne makroekonomiczne przy wykorzystaniu modeli ekonometrycznych. Następnie przeprowadzili meta-analizę regresji otrzymanych wówczas wielkości mnożników fiskalnych[90]. Uzyskane wyniki pozwoliły na sformułowanie następujących wniosków. Po pierwsze, mnożniki fiskalne są rzeczywiście wyraźnie

[90] Metaanaliza regresji polega na tym, że wyniki z poszczególnych badań są traktowane jako zbiór obserwacji, które następnie wykorzystuje się w ramach „normalnej" analizy regresji w celu oszacowania związków między zmiennymi fiskalnymi a zmiennymi makroekonomicznymi, w tym wielkości mnożników.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wyższe w okresie recesji niż w czasach „dobrych" i „normalnych". Na przykład, dla niewyspecyfikowanych wydatków rządowych mnożniki te stają się średnio wyższe o 0,6–0,8 jednostki. Po drugie, w czasach „dobrych" i „normalnych" mnożniki nie tylko okazują niższe, lecz także mniej różnią się między sobą w przypadku poszczególnych instrumentów fiskalnych. Po trzecie, wśród wybranych instrumentów wyjątkowo silnym dodatnim wpływem w czasie recesji charakteryzują się wydatki transferowe, a także wydatki na obrone narodową. Po czwarte, mnożniki dotyczące podatków są generalnie zasadniczo niewielkie, znacznie niższe niż mnożniki odnoszące się do wydatków, niezależnie od fazy cyklu koniunkturalnego. Większość otrzymanych oszacowań pozostaje statystycznie istotna, z wyjątkiem mnożnika wydatków konsumpcyjnych[91].

Uzyskane wyniki sugerują, że konsolidacja fiskalna w strefie euro mogła znacząco pogłębić recesję. Być może strefa euro byłaby w stanie szybciej powrócić na ścieżkę wzrostu, gdyby zastosowała stymulację fiskalną na większą skalę, zwłaszcza w latach 2011–2013. Trudno jednak ten wniosek zweryfikować, ponieważ kraje wysoko zadłużone nie miały praktycznie możliwości znacznego zwiększenia wydatków rządowych ze względu na ostre ograniczenia budżetowe i brak pełnego dostępu do rynków finansowych. Konsolidacja wydatków była więc prawdopodobnie w przypadku tych państw nieuniknioną koniecznością.

W ramce 3.15 zweryfikowana została hipoteza dotycząca wpływu stymulacji fiskalnej na wzrost konsumpcji indywidualnej wśród grupy krajów strefy euro w okresie kryzysu finansowego i gospodarczego (lata 2009–2014). Oszacowana zależność wskazuje, że wraz ze spadkiem salda budżetu – czyli razem ze wzrostem deficytu – tempo wzrostu konsumpcji maleje. Rezultat ten generalnie nie potwierdza dodatniego wpływu stymulacji fiskalnej na poziom konsumpcji.

[91] Komentując metodę meta-regresji zastosowaną przez Gecherta, Halletta i Ranneberga, należy pamiętać, że nie bierze się w tym przypadku pod uwagę różnic zachodzących w statystycznej jakości i istotności modeli pochodzących z różnych opracowań, a także nie uwzględnia się współzależności poszczególnych modeli. Otrzymane wyniki trzeba więc traktować z ostrożnością.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 3.15. Weryfikacja hipotezy o wpływie stymulacji fiskalnej na wzrost konsumpcji indywidualnej w warunkach kryzysu i przy dużym zadłużeniu: strefa euro w latach 2009–2014

Hipoteza o dodatnim (lub ujemnym) wpływie stymulacji fiskalnej na wzrost gospodarczy w warunkach kryzysu i przy dużym zadłużeniu została zweryfikowana statystycznie z wykorzystaniem danych empirycznych dotyczących państw strefy euro. Na wykresie 3.18 przedstawiono 96 obserwacji par zmiennych – salda strukturalnego sektora finansów publicznych (jako % PKB) i rocznych zmian konsumpcji indywidualnej (w %) – dla 16 państw członkowskich strefy euro[92] w latach 2009–2014. Wybrano deficyt strukturalny, a nie faktyczny, aby wyeliminować wpływ cyklu koniunkturalnego na saldo budżetu i lepiej ocenić skutki działań dyskrecjonalnych rządu. Na podstawie danych panelowych oszacowano równanie regresji (z pominięciem efektów stałych) o następującej postaci (w nawiasach wartości statystyki *t*):

$$y = 0{,}8616 - 0{,}2345x \tag{3.54}$$

$$(2{,}321) \quad (4.343)$$

$$R(kwadrat) = 0{,}1987; \quad F(stat) = 93{,}06; \quad SEE = 2{,}222.$$

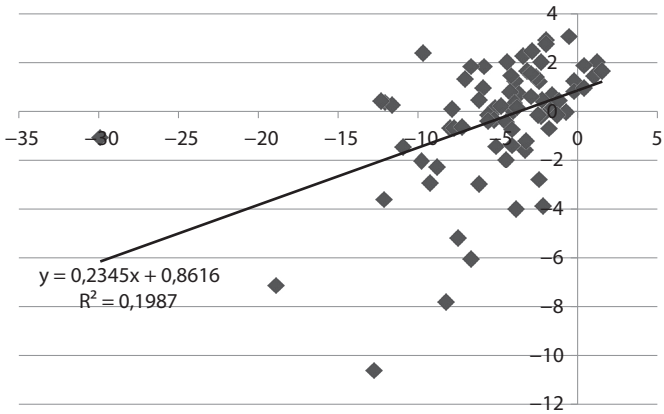
Wartości statystyki *t* i *F* wskazują, że zarówno oszacowane parametry, jak i całe równanie regresji są statystycznie istotne na poziomie 95%. Równanie pokazuje, że między zmiennymi istnieje niezbyt silna, lecz istotna statystycznie zależność dodatnia – im wyższe saldo budżetu (a zatem im mniejszy deficyt), tym szybsze tempo wzrostu konsumpcji indywidualnej, przy czym poprawie salda budżetu o 1% PKB towarzyszyło przyspieszenie tempa wzrostu konsumpcji o około 0,23 punktu procentowego. Dane te w znacznym stopniu wspierają hipotezę o „niekeynesowskim" (czy nawet „antykeynesowskim") efekcie stymulacji fiskalnej, przynajmniej w przypadku większości krajów, choć oszacowana zależność jest zbyt słaba

[92] Bez Estonii, Łotwy i Litwy.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i oparta o zbyt małą próbę, aby można było na jej podstawie formułować bardziej uniwersalne sądy[93].

Wykres 3.18. Saldo strukturalne budżetu (oś pozioma *x*) i tempo wzrostu konsumpcji prywatnej (opóźnione o 1 rok – oś pionowa *y*) wśród 16 krajów strefy euro w latach 2009–2014 (96 obserwacji rocznych) (dane w %)



Uwaga: saldo sektora finansów publicznych (*cyclically-adjusted net lending of general government*) jako % PKB. Tempo wzrostu konsumpcji prywatnej w % (r/r).

Źródło: obliczenia własne na podstawie danych Eurostatu.

Literatura cytowana

Acocella N., *Economic Policy in the Age of Globalization*, Cambridge University Press, Cambridge 2005.

Afonso A., Schuknecht L., Tanzi V., *Public Sector Efficiency: An International Comparison*, European Central Bank, "Working Paper Series" 2002, no. 242.

[93] Wnioski te są generalnie zgodne z wynikami innych badań. Nickel i Vansteenkiste (2008), analizując związek między wielkością długu publicznego a saldem rachunku bieżącego w krajach UE, stwierdziły, że korelacja ta jest dodatnia w przypadku długu nieprzekraczającego 90% PKB. Powyżej tego poziomu wzrost zadłużenia nie wywoływał wzrostu deficytu bieżącego. Świadczy to pośrednio o tym, że przy bardzo wysokim poziomie zadłużenia stymulacja fiskalna, finansowana dalszym wzrostem długu, może skłaniać gospodarstwa domowe do zwiększania oszczędności i nie powodować przyrostu deficytu rachunku bieżącego. Oznacza to, że w warunkach wysokiego zadłużenia gospodarstwa domowe zachowują się w sposób zgodny z równoważnością rikardiańską.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Afonso A., Jalles J.T., *Fiscal composition and long-term growth*, European Central Bank, "Working Papers Series" 2013, no. 1518.

Aiyagari R.S., Christiano L., Eichenbaum M., *The Output, Employment and Interest Rate Effects of Government Consumption*, "Journal of Monetary Economics" 1992, no. 30.

Alesina A., Perotti R., *Fiscal Discipline and the Budget Process*, "American Economic Review" 1996, no. 86.

Alesina A., Roubini N., Cohen G., *Political Cycles and the Macroeconomy*, MIT Press, Cambridge 1997.

Alesina A., Perrotti R., Tavares J., *The Political Economy of Fiscal Adjustments*, "Brookings Papers on Economic Activity" 1998, no. 1.

Alesina A., Ardagna S., Perotti R., Schantarelli F., *Fiscal policy, profits and investment*, "NBER Working Paper" 1999, no. 7207.

Alesina A., Ardagna S., *Large Changes in Fiscal Policy: Taxes versus Spending*, "Tax Policy and the Economy" 2010, vol. 24.

Alesina A., Favero C., Giavazzi F., *The Output Effect of Fiscal Consolidation*, "NBER Working Paper" 2012, no. 18336.

Armey D., *The Freedom Revolution*, Ragnery Publishing, Washington 1995.

Auerbach A.J., Gorodnichenko Y., *Measuring the Output Responses to Fiscal Policy*, "American Economic Journal: Economic Policy" 2012, vol. 4, no. 2.

Bajo-Rubio O., *A further generalization of the Solow growth model: the role of the public sector*, "Economic Letters" 2000, no. 68.

Balassone F., Franco D., *Fiscal Federalism and the Stability and Growth Pact: A Difficult Union*, 1 February, 2001, http://ssrn.com/abstract=2095225 (5.06.2014).

Barro R.J., *On the Determination of the Public Debt*, "Journal of Political Economy" 1979, vol. 87, no. 5.

Barro R.J., *A Cross-Country Study of Growth, Saving and Government*, "NBER Working Paper" 1989, no. 2855.

Barro R.J., *Government Spending in a Simple Model of Endogenous Growth*, "Journal of Political Economy" 1990, vol. 98, no. 5.

Barry F., Devereux M.B., *Expansionary fiscal contraction: A theoretical exploration*, "Journal of Macroeconomics" 2003, vol. 25.

Baxter M., King R.G., *Fiscal Policy in General Equilibrium*, "American Economic Review" 1993, no. 83.

Blanchard O.J., *Current and anticipated deficits, interest rates and economic activity*, "European Economic Review" 1984, vol. 25(1).

Blanchard O.J., *Comment*, "NBER Macroeconomics Annual" 1990.

Blanchard O.J., Perotti R., *An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output*, "The Quarterly Journal of Economics" 2002, vol. CXVII.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Blanchard O., Leigh D., *Growth Forecast Errors and Fiscal Multipliers*, "IMF Working Paper" 2013a, no. 1.

Blanchard O., Leigh D., *Fiscal consolidation. At what speed?*, VOX–CEPR's Policy Portal, 3 May, 2013b, www.voxeu.org/article/fiscal-consolidation-what-speed (13.10.2014).

Buchanan J.M., *The Limits of Liberty: Between Anarchy and Leviathan*, University of Chicago Press, Chicago 1975.

Buchanan J.M., Tullock G.W., *The Calculus of Consent*, Michigan University Press, Ann Arbor 1962.

Buiter W., Corsetti G., Roubini N., *Excessive Deficits: Sense and Nonsense in the Treaty of Maastricht*, "Economic Policy" 1993, no. 8.

Buiter W., Rahbari E., *Greece and the fiscal crisis in the Eurozone*, "Policy Insight" 2010, no. 51, CEPR.

Bulow J., Rogoff K., *Sovereign Debt: Is to Forgive to Forget?*, "American Economic Review" 1989, vol. 79.

Burda M., Wyplosz C., *Makroekonomia. Podręcznik europejski*, wydanie 2 zmienione, PWE, Warszawa 2000.

Burman L.E., Slemrod J., *Taxes in America. What Everyone Needs to Know*, Oxford University Press, Oxford 2012.

Burnside C., Eichenbaum M., Fisher J., *Fiscal Shocks and Their Consequences*, "NBER Working Paper" 2003, no. 9772.

Buti M., *Will the New Stability and Growth Pact Succeed? An Economic and Political Perspective*, European Commission, 9 January, 2006, mimeo.

Buti M., Eijffinger S., Franco D., *Revisiting the Stability and Growth Pact: Grand Design or Internal Adjustment?*, "European Economy – Economic Papers" 2003, no. 180.

Calmfors L., *What Remains of the Stability Pact and What Next?*, Swedish Institute for European Policy Studies, Report no. 8, November 2005.

Calmfors L., *The role of independent fiscal policy institutions*, Report to the Prime Minister's Office, Institute for International Economic Studies, Stockholm University, Finland, 1 September, 2010.

Canes-Wrone B., Ponce de Leon C., *Electoral cycles and democratic development*, Princeton University, Princeton 2015, mimeo.

Cecchetti S.G., Mohanty M.S., Zampolli F., *The future of public debt: prospects and implications*, "BIS Working Papers" 2010, no. 300.

Cecchetti S.G., Mohanty M.S., Zampolli F., *The real effects of debt*, "BIS Working Papers" 2011, no. 352.

Checherita C., Rother Ph., *The impact of high and growing government debt on economic growth: An empirical investigation for the euro area*, "Working Paper Series" 2010, no. 1237, European Central Bank.

Chobanov D., Mladenova A., *What Is the Optimum Size of Government*, Institute for Market Economics, Sofia, August 2009.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Christiano L.J., Eichenbaum M., *Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations*, "American Economic Review" 1992, vol. 82.

Christiano L.J., Eichenbaum M., Evans C.L., *Nominal Rigidity and the Dynamic Effects of a Shock to Montary Policy*, "Journal of Political Economy" 2005, vol. 113, no. 1.

Christiano L., Eichenbaum M., Rebelo S., *When is the government spending multiplier large?*, "NBER Working Paper" 2009, no. 15394.

Coeré B., Pisani-Ferry J., *Fiscal Policy in EMU: Towards a Sustainability and Growth Pact*, "Oxford Review of Economic Policy" 2005.

Cogan J.F., Cwik T., Taylor J.B., Wieland V., *New Keynesian versus Old Keynesian Government Spending Multipliers*, "NBER Workin Paper" 2009, no. 14782.

Corden M.W., *The theory of the fiscal stimulus: How will a debt-financed stimulus affect the future?*, "Policy Insight" 2009, no. 34, CEPR.

Corsetti G., Mueller G., *Sovereign risk, macroeconomic instability*, 12 August, 2011, www.voxeu.org/indexphp?q=node/6857 (12.06.2014).

Corsetti G., Kuestner K., Meier A., Mueller G., *Debt consolidation and fiscal stabilization of deep recession*, 2010, mimeo.

Cottarelli C., Forni L., Gottschalk J., Mauro P., *Default in Today's Advanced Economies: Unnecessary, Undesirable and Unlikely*, IMF Staff Position Note, SPN/10/12, 1 September, 2010.

Darvas Z., *Debt Restructuring in the Euro Area: A Necessary but Manageable Evil?*, "Bruegel Policy Contribution" 2011, no. 7.

De Grauwe P., *Economics of Monetary Union*, Oxford University Press, Oxford 2000.

De Grauwe P., Ji Y., *Mispricing of Sovereign Risk and Multiple Equilibria in the Eurozone,* CEPS Working Documents, Brussels 2012.

DeLong B., Summers L., *Fiscal Policy in a Depressed Economy*, NBER, March 2012, mimeo.

Deroose S., Lagendijk S., *Economic policy coordination in EMU: accomplishments and challenges*, w: *EMU and Economic Policy in Europe. The Challenge of the Early Years*, red. M. Buti, A. Sapir, European Commission, Edward Elgar, Cheltenham 2002.

DeVries P., Guajardo J., Leigh D., Pescatori A., *A New Action-based Dataset of Fiscal Consolidation*, "IMF Working Paper" 2011, no. 128.

Eggertsson G.B., Krugman P., *Debt, Deleveraging, and the Liquidity Trap. A Fisher-Minsky-Koo Approach*, "Quarterly Journal of Economics" 2012, vol. 127, no. 3.

Eichengreen B., Wyplosz C., *The Stability Pact. More than a Minor Nuisance?*, "Economic Policy" 1998, no. 26.

Eichengreen B., Hausmann R., von Hagen J., *Reforming Budgetary Institutions in Latin America*, "Open Economies Review" 1999, vol. 10, no. 4.

Eichengreen B. i in., *Public Debts: Nuts, Bolts and Worries*, Geneva Reports on the World Economy, International Center for Monetary and Banking Studies and Centre for Economic Policy Research, Geneva 2011.

EMU and Economic Policy in Europe. The Challenge of the Early Years, red. M. Buti, A. Sapir, European Commission, Edward Elgar, Cheltenham 2002.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

European Commission, *European Economy, Supplement A. Budgetary Surveillance in the EMU*, March 1999.

European Commission, *Public Finance in EMU, European Commission*, "European Economy" 2009, no. 5.

European Commission, *Reinforcing economic policy coordination*, Communication from the Commission to the European Parliament, the European Council, the Council, the European Central Bank, the Economic and Social Committee and the Committee of the Regions, Brussels, 12 May, 2010.

European Commission, *Focus: Debt dynamics and the sustainability of the euro area*, "Quarterly Report on the Euro Area" 2011, vol. 10, issue 3.

European Commission, *Report on Public Finances in EMU – 2014*, "European Economy" 2014, no. 9.

Eyraud L., Weber A., *The Challenge of Debt Reduction during Fiscal Consolidation*, "IMF Working Paper" 2013, no. 67.

Fatas A., Mihov I., *Fiscal policy and EMU: challenges of the early years*, w: *EMU and Economic Policy in Europe. The Challenge of the Early Years*, red. M. Buti, A. Sapir, European Commission, Edward Elgar, Cheltenham 2002.

Fatas A., Mihov I., *On Constraining Fiscal Policy Discretion in EMU*, "Oxford Review of Economic Policy" 2003, no. 19(1).

Fels G., Froehlich H.P., *Germany and the world economy: A German view*, "Economic Policy" 1986, no. 4.

Galí J., Perotti R., *Fiscal Policy and Monetary Integration in Europe*, "Economic Policy" 2003, no. 37.

Galí J., López-Salida J.D., Vallés J., *Understanding the Effects of Government Spending on Consumption*, "Journal of European Economic Association" 2007, vol. 5, no. 1.

Gechert S., Ranneberg A., *Are Fiscal Multipliers Regime-Dependent? A Meta Regression Analysis*, "IMK Working Paper" 2014, no. 139.

Gechert S., Hughes-Hallett A., Ranneberg A., *Fiscal multipliers in downturns and the effects of Eurozone consolidation*, "CEPR Policy Insight" 2015, no. 79.

Ghali K.H., *Government size and economic growth: evidence from a multivariate cointegration analysis*, "Applied Economics" 1998, no. 31.

Giavazzi F., Pagano M., *Can severe fiscal contractions be expansionary? Tales from two small European economies*, "NBER Working Paper" 1990, no. 3372.

Gruber J., *Public Finance and Public Policy*, Worth Publishers, New York 2007.

Guajardo J., Leigh D., Pescatori A., *Expansionary Austerity: new International Evidence*, "IMF Working Paper" 2011, no. 158.

Gwartney J., Holcombe R., Lawson R., *The Scope of Government and the Wealth of Nations*, "Cato Journal" 1998, vol. 18, no. 2.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Hallerberg M., von Hagen J., *Electoral Institutions, Cabinet Negotiations and Budget Deficits in the European Union*, w: *Fiscal Institutions and Fiscal Performance*, red. J. Poterba, J. von Hagen, NBER, University of Chicago Press, Chicago 1999.

Heller P., *Back to Basics – Fiscal Space: What It Is and How to Get It*, "Finance and Development" 2005, vol. 42, no. 2.

Hellwig M., Neumann M.J.M., Vaubel R., Miller M., *Economic policy in Germany: Was there a turnaround?*, "Economic Policy" 1987, no. 5.

Ho W.H.A., Yetman J., *Do Economies Stall? The International Evidence*, "BIS Working Paper" 2013, no. 407.

IMF, *Toward a Fiscal Union for the Euro Area*, IMF Staff Discussion Note, SDN/13/09, September 2013.

Irons J., Bivens J., *Government Debt and Economic Growth*, "Economic Policy Institute Briefing Paper" 2010, no. 271.

Kalecki M., *Parę uwag o teorii Keynesa*, „Ekonomista" 1936, nr 3.

Kalecki M., *Essays in the Theory of Economic Fluctuations*, New York 1939.

Kontopoulos Y., Perotti R., *Government Fragmentation and Fiscal Policy Outcomes: Evidence from OECD Countries*, w: *Fiscal Institutions and Fiscal Performance*, red. J. Poterba, J. von Hagen, University of Chicago Press, Chicago 1999.

Kopits G., *Fiscal Policy Rules: A Useful Framework or an Unnecessary Ornament?*, 2001, www.bancaditalia.it/studiricerche/convegni/.../059-084_kopits.pdf (30.09.2012).

Kopits G., Symansky S., *Fiscal policy rules*, "IMF Occasional Paper" 1998, no. 162.

Krugman P., *End this depression now!*, W.W. Norton and Company, New York 2012.

Levitt M.R., Brass C.T., Nicola T.J., Nuschler D., *Reaching the debt Limit: Background and Potential Effects on Government Operations*, Congressional Research Service, US Congress, Washington, 4 January, 2013.

Magazzino C., Forte F., *Optimal size of government and economic growth in EU-27*, "Munich Paper" 2010, no. 26669.

McKinnon R., *Default Risk in Monetary Union*, Background Report for the Swedish Government Commission on EMU, Stockholm 1996.

McLure C.E. Jr., *The Tax Assignment Problem: Ends, Means, and Constraints*, "Public Budgeting and Financial Management" 1998, vol. 9, no. 4.

Melitz J., *Some cross-country evidence about fiscal policy behavior and consequences for EMU*, "European Economy Reports and Studies" 2000, no. 2.

Melitz J., *Eurozone reform. A proposal, Centre for Economic Policy Research*, "Policy Insight" 2010, no. 48.

Milesi-Feretti G.M., *Good, Bad or Ugly? On the Effects of Fiscal Rules with Creative Accounting*, "Journal of Public Economics" 2003, no. 88.

Musgrave R.A., *The Theory of Public Finance: A Study in Public Economy*, McGraw Hill, New York 1959.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Nickel C., Vansteenkiste I., *Fiscal Policies, the Current Account and Ricardian Equivalence*, European Central Bank, "Working Paper Series" 2008, no. 935.

Nickel C., Tudyka A., *Fiscal Stimulus in Times of High Debt. Reconsidering Multipliers and Twin deficits*, European Central Bank, "Working Papers Series" 2013, no. 1513.

Nordhaus W.D., *The political business cycle*, "The Review of Economic Studies" 1975, no. 42(2).

North D.C., *Institutions, Institutional Change, and Economic Performance*, Cambridge University Press, Cambridge 1990.

Oates W.E., *Fiscal Federalism*, Harcourt Brace Jovanovich, New York 1972.

Oates W.E., *The Economics of Fiscal Federalism and Local Finance*, Edward Elgar, Northampton 1998.

Ostry J.D., Ghosh A.R., Kim J.I., Qureshi M.S., *Fiscal Space*, IMF Staff Discussion Note, SDN/10/11, September 2010.

Ostry J.D., Gosh A.R., Espinoza R., *When Should Public Debt Be Reduced?*, IMF Staff Discussion Note, SDN/15/10, June 2015.

Parenteau R., *Fiscal consolidation is not expansionary after all*, 30 October, 2011, http://www.eurointelligence.com (12.05.2014).

Pecorino P., *Tax rates and tax revenues in a model of growth through human capital accumulation*, "Journal of Monetary Economics" 1995, no. 36(3).

Perotti R., Strauch R., von Hagen J., *Sustainability of Public Finances*, CEPR, London 1998.

Pisani-Ferry J., *The EMU's economic policy and principles: words and facts*, w: *EMU and Economic Policy in Europe. The Challenge of the Early Years*, red. M. Buti, A. Sapir, European Commission, Edward Elgar, Cheltenham 2002.

Pisani-Ferry J., *Fiscal Discipline and Policy Coordination in the Euro area: Assessment and Proposals*, w: *Budgetary Policy in E(M)U, Design and Challenges*, Conference proceedings, Dutch Ministry of Finance, The Hague 2003.

Poterba J.M., *State Responses to Fiscal Crises: The Effects of Budgetary Institutions and Politics*, "Journal of Political Economy" 1994, no. 102(4).

Rahn R.W., Fox H.W., Fox L.H., *What is the Optimum Size of Government*, Vernon Krieble Foundation, Denver 1996.

Ramey V.A., *Identifying Government Spending Shocks: It's All in the Timing*, "NBER Working Paper" 2009, no. 15464.

Ramey V.A., Shapiro M.D., *Costly Capital Reallocation and the Effects of Government Spending*, "Carnegie Rochester Conference Series on Public Policy" 1998, no. 48.

Reinhart C.M., Rogoff K.S., *Growth in a Time of Debt*, "American Economic Review" 2010, no. 100(2).

Rogoff K., Sibert A., *Elections and macroeconomic policy cycles*, "Review of Economic Studies" 1988, no. 55(1).

Romer D., *Makroekonomia dla zaawansowanych*, Wydawnictwo Naukowe PWN, Warszawa 2000.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Romer C., Bernstein J., *The Job Impact of the American Recovery and Reinvestment Plan*, 8 January, 2009, mimeo.

Rosati D.K., *Kierunki zmian w systemie zarządzania fiskalnego w strefie euro*, „Ekonomista" 2010, no. 5.

Rosati D.K., *Konsolidacja fiskalna a kryzys zadłużenia w strefie euro*, w: *Kryzys finansów publicznych. Przyczyny, mechanizm, drogi wyjścia*, red. L. Oręziak, D.K. Rosati, Uczelnia Łazarskiego, Warszawa 2013.

Sapir A. i in., *An Agenda for a Growing Europe. The Sapir Report*, Oxford University Press, European Commissions, Oxford 2004.

Samuelson P.A., *The Pure Theory of Public Expenditure*, "Review of Economics and Statistics" 1954, vol. 36, no. 4.

Schuknecht L., *Political business cycles and fiscal policies in developing countries*, "Kyklos" 1996, no. 49.

Schultz K.A., *The politics of the political business cycle*, "British Journal of Political Science" 1995, vol. 25.

Scully G.W., *What is the Optimal Size of Government in the United States*, National Centre for Policy Analysis, "Policy Report" 1994, no. 188.

Scully G.W., *Measuring the Burden of High Taxes*, National Centre for Policy Analysis, "Policy Report" 1998, no. 215.

Sheets N., *Is the US Economy Approaching "Stall Speed"?*, "Citi Empirical and Thematic Perspectives", 11 November, 2011.

Shi M., Svensson J., *Political budget cycles in developed and developing countries*, IIES, Stockholm University, 2002, mimeo.

Shi M., Svensson J., *Political Budget Cycles: A Review of Recent Developments*, "Nordic Journal of Political Economy" 2003, vol. 29.

Stark J., *Genesis of a Pact, in: Brunilla*, w: *The Stability and Growth Pact – The Architecture of Fiscal Policy in EMU*, red. M. Buti, D. Franco, Palgrave, Basingstoke 2001.

Stiglitz J.E., *Ekonomia sektora publicznego*, Wydawnictwo Naukowe PWN, Warszawa 2004.

Sulmicki P., *La théorie du multiplicateur d'investissement et l'experiance allemande 1933–1937*, Librairie du Recueil Sirey, Paris 1945 (wydanie polskie: *Teoria mnożnika inwestycyjnego a doświadczenia niemieckie 1933–1937*, Oficyna Wydawnicza SGH, Warszawa 2014).

Tabellini G., *The Politics of Intergenerational Redistribution*, "Journal of Political Economy" 1991, vol. 99.

Tanzi V., Schuknecht L., *Public Spending in the 20th Century: A Global Perspective*, Cambridge University Press, Cambridge 2000.

Von Hagen J., *Fiscal Rules, Fiscal Institutions, and Fiscal Performance*, "The Economic and Social Review" 2002, no. 33(3).

Von Hagen J., Harden I., *Budget Processes and Commitment to Fiscal Discipline*, "IMF Working Paper" 1996, no. WP/96/78.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Von Hagen J., Mundschenk S., *The functioning of economic policy coordination*, w: *EMU and Economic Policy in Europe. The Challenge of the Early Years*, red. M. Buti, A. Sapir, European Commission, Edward Elgar, Cheltenham 2002.

Von Hagen J., Guntram B.W., *What do deficits tell us about debt? Empirical evidence on creative accounting with fiscal rules in the EU*, Deutsche Bundesbank, "Discussion Paper Series 1: Studies of the Economic Research Centre" 2004, no. 38.

Wyplosz C., *Fiscal Policy: Institutions vs. Rules,* w*: Stebiliseringspolitik i valutaunionen, Swedish Government Commission on Stabilization Policy in the EMU*, Frietzes, Stockholm 2002a.

Wyplosz C., *Fiscal Discipline in EMU: Rules or Institutions?*, Paper prepared for the meeting of Group of Economic Advisers of the European Commission, April 2002b.

Wyplosz C., *Fiscal Policy: Institutions versus Rules*, "National Institute Economic Review" 2005, no. 191.

Zoli E., *Italian Sovereign Spreads: Their Determinants and Pass-through to Bank Funding Costs and Lending Conditions*, "IMF Working Paper" 2013, no. 13(84).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozdział 4

POLITYKA PODATKOWA

4.1. Podstawowe pojęcia i definicje

Podatki – obok wydatków budżetowych – są najważniejszym instrumentem polityki fiskalnej. To głównie dzięki nim władza gospodarcza każdego kraju jest w stanie gromadzić środki na finansowanie wydatków budżetowych, które pozwalają na realizację ważnych celów społecznych. Podatki – czy też szerzej: daniny publiczne – funkcjonują od najdawniejszych czasów, stanowiąc podstawowe źródło dochodów władzy publicznej, przeznaczanych głównie na zaspokajanie jej potrzeb oraz zapewnianie bezpieczeństwa wewnętrznego i zewnętrznego danej społeczności.

Podatek definiuje się jako świadczenie pieniężne ustalone jednostronnie (czyli arbitralnie, a nie w drodze negocjacji), o charakterze nieodpłatnym i bezzwrotnym, pobierane powszechnie na rzecz władzy publicznej. Dawniej władzę tę reprezentował monarcha, obecnie jest to zaś Skarb Państwa, a często także samorząd lokalny, w zależności od typu rozwiązań ustrojowych przyjętych w danym państwie. Podatek stanowi przymusową, legalną, bezzwrotną daninę, płaconą przez wszystkich członków danej społeczności lub przez określone grupy społeczne. Przymusowość oznacza, że podatek może być egzekwowany przy zastosowaniu środków przymusu, jeśli podatnik unika płacenia należności. Z kolei legalność wskazuje na to, że obowiązek rozliczania podatku reguluje określone prawo. Współcześnie musi to być akt prawny wysokiej rangi, przyjmujący postać co najmniej ustawy. Ma to istotne znaczenie z perspektywy uzyskania społecznej akceptacji dla podatków, ponieważ każdy podatek jest w istocie przejęciem przez władzę publiczną własności prywatnej bez odszkodowania. Legalność odróżnia więc podatki od innych, bezprawnych form konfiskaty mienia. Wreszcie bezzwrotność oznacza, że świadczenie pieniężne jest jednostronne, a zatem nie podlega zwrotowi, i nie może być źródłem konkretnego roszczenia.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Formy podatku ulegały zmianom w miarę rozwijania się form gospodarowania. Początkowo przybierały postać świadczenia w naturze (zboża, bydło), stanowiącego część produkcji wytwarzanej przez poszczególnych członków danej społeczności. W ten sposób wykształcił się podatek w postaci dziesięciny – czyli jednej dziesiątej zbiorów lub produktów hodowli, uiszczany na rzecz panującego władcy. Odmianą podatku w naturze była pańszczyzna, tj. nieodpłatne świadczenie pracy na rzecz panującego, oraz tak zwany szarwark, czyli nieodpłatne świadczenie określonych usług na rzecz panującego. Dopiero w XVII–XVIII wieku wśród najbardziej rozwiniętych krajów upowszechniła się nowoczesna, pieniężna forma opodatkowania.

Współcześnie nałożenie każdego podatku wymaga ustalenia jego zakresu, poziomu, obszaru zastosowania i charakteru. Te podstawowe parametry musi regulować ustawa (lub akt prawny równorzędnej rangi). Opis zasadniczych elementów przepisu podatkowego zawiera tabela 4.1.

Tabela 4.1. Zasadnicze elementy przepisu podatkowego

Zasadniczy element podatku	Opis
Obowiązek podatkowy	Wynikająca z ustawy ogólna (nieskonkretyzowana – bez ustalenia wysokości itp.) powinność przymusowego świadczenia pieniężnego w związku z zaistnieniem zdarzenia określonego w ustawie.
Przedmiot podatku	Zdarzenie, którego zaistnienie powoduje powstanie obowiązku podatkowego (na przykład transakcja kupna-sprzedaży, uzyskanie dochodu, posiadanie nieruchomości).
Podmiot podatku	Podatnik, czyli osoba fizyczna, osoba prawna lub jednostka nieposiadająca osobowości prawnej, lecz podlegająca obowiązkowi podatkowemu.
Płatnik podatku	Osoba fizyczna lub prawna, obowiązana do obliczenia i pobrania podatku od podatnika, a następnie wpłacenia go w określonym terminie organowi podatkowemu.
Podstawa opodatkowania (baza podatkowa)	Wartościowo lub ilościowo ujęty (skwantyfikowany) przedmiot podatku (na przykład wartość transakcji, wartość otrzymanego dochodu, wartość nieruchomości).
Stawka podatkowa	Parametr kwotowy lub procentowy pozwalający ustalić należny podatek w zależności od podstawy opodatkowania.

Źródło: opracowanie własne.

Ponieważ podatki są głównym źródłem dochodów państwa, troska o utrzymanie tych dochodów na odpowiednim poziomie zawsze absorbowała uwagę każdej władzy publicznej. W większości państw obowiązek podatkowy jest zapisany w konstytucji. Art. 84 polskiej ustawy zasadniczej stanowi, że „każdy jest obowiązany

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

do ponoszenia ciężarów i świadczeń publicznych, w tym podatków określonych w ustawie". Zapis ten nadaje podatkom legalny charakter, a zarazem, podkreślając ustawową rangę obowiązku podatkowego, chroni podatników przed samowolą administracji. Jeden z Ojców Założycieli Stanów Zjednoczonych, Benjamin Franklin, zwykł był mawiać, że „w życiu człowieka istnieją jedynie dwie pewne rzeczy: płacenie podatków i własny pogrzeb". Polityk podkreślał w ten sposób bezwzględny i uniwersalny charakter obowiązku podatkowego, co miało szczególne znaczenie w procesie budowy podstaw ustrojowych nowego państwa[1].

4.2. Zasady opodatkowania

4.2.1. Potrzeba równoważenia skutków podatku

Choć w pewnym zakresie większość ludzi jest skłonna do altruistycznych zachowań, to, gdy mowa o podatkach, na ogół niechętnie oddaje się część zarobionych przez nas pieniędzy fiskusowi. A skoro już musimy płacić podatki, staramy się aby były one możliwie jak najniższe i nie wyższe niż podatki płacone przez innych podatników. Dlatego też decyzje dotyczące regulacji podatkowych są zawsze bardzo wrażliwym politycznie elementem polityki państwa i budzą mieszane odczucia. Pod względem ekonomicznym podatki wywierają wpływ na zachowania uczestników rynku – na ich skłonność do pracy, gromadzenia oszczędności oraz podejmowania działalności gospodarczej – i w ten sposób oddziałują na poziom produkcji, poziom zatrudnienia i wzrost gospodarczy. Wprowadzenie każdego podatku pociąga za sobą dwa bezpośrednie skutki – przynosi dochód władzy publicznej (efekt fiskalny) i ogranicza działalność objętą podatkiem (efekt ekonomiczny). W związku z tym dobry system podatkowy powinien być skonstruowany w taki sposób, aby, zapewniając odpowiedni poziom dochodów budżetowych, stanowić podatki możliwie jak najmniej szkodliwe dla procesów gospodarczych.

Negatywny wpływ podatków – zwłaszcza wysokich – na działalność gospodarczą dostrzegano od dawna. Tomasz z Akwinu (1225–1274) przestrzegał przed wprowadzaniem zbyt wysokich podatków, wskazując na to, że „podatek nie powinien

[1] Franklin doskonale wiedział, o czym mówił. Trzeba bowiem pamiętać, że w końcu XVIII wieku obywatele Stanów Zjednoczonych okazywali się często uciekinierami z państw europejskich, ściganymi przez władze m.in. za unikanie płacenia podatków.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

naruszać substancji majątkowej opodatkowanego" i zauważając, że nadmierne podatki mogą prowadzić do ruiny kupców i wytwórców. Wypowiadał się on zatem na rzecz podatku rozumianego raczej jako udział w nowo wytworzonym dochodzie niż jako narzędzie konfiskaty zgromadzonego majątku. Podobny pogląd głosił XVIII-wieczny ekonomista niemiecki Johan Henrich Gottlob von Justi (1705–1771), który ostrzegał panujących przed zbyt wysokimi podatkami, które hamują rozwój aktywności gospodarczej. Myśli te rozwijał dalej David Ricardo (1772–1823). Sformułowana przez niego tzw. reguła edynburska zaleca, aby opodatkowanie ustalać na takim poziomie, który pozwoli każdemu podatnikowi pozostać we względnie takiej samej sytuacji jak przed opodatkowaniem.

Z kolei Adam Smith (1723–1790) sformułował cztery zasady dobrych podatków. Zasada równości opodatkowania wymaga, aby ludzie będący w podobnej sytuacji finansowej płacili takie same podatki. Zasada określoności (pewności) opodatkowania mówi o tym, że reguły i stawki podatku powinny być stabilne i podawane do wiadomości z odpowiednim wyprzedzeniem, tak aby ludzie mogli bezpiecznie planować swoją przyszłość. Zasada dogodności opodatkowania wymaga, aby system podatkowy nie był nadmiernie skomplikowany i aby sprzyjał zarówno płaceniu podatków, jak i ich poborowi. Wreszcie zasada taniości podatku obliguje do tego, aby system podatkowy i jego funkcjonowanie nie generowały zbyt dużych kosztów ponoszonych przez podatników. Bardzo zbliżone zasady podatkowe sformułował także wspomniany w poprzednim rozdziale niemiecki ekonomista Adolf Wagner (Wagner 1877, cyt. za: Sosnowski 2012).

Również John Stuart Mill (1806–1873) podkreślał, że ciężary podatkowe muszą być względnie równo rozłożone wśród członków społeczeństwa. Zdefiniował on tzw. zasadę równości ofiar, zgodnie z którą uczucie przykrości z tytułu płacenia podatku powinno być odczuwane przez każdego obywatela w tym samym stopniu. W myśl tej zasady osoby zamożniejsze powinny być zobligowane do tego, aby wpłacać do wspólnej kasy więcej niż osoby uboższe.

4.2.2. Zasada równości poziomej i zasada równości pionowej

Klasyczne zasady podatkowe Smitha i Milla są obecnie traktowane również jako ważne wytyczne przy tworzeniu współczesnych systemów podatkowych. Podstawę prawidłowego, tj. sprawiedliwego, określenia ciężaru podatkowego stanowi we współczesnej nauce o podatkach pojęcie „zdolności do płacenia podatku" przez daną osobę fizyczną lub prawną. Zasada równości obywateli wobec obowiązku

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

podatkowego jest obecnie definiowana w dwóch wymiarach – jako równość pozioma (*horizontal equity*) i równość pionowa (*vertical equity*). Pierwsza z nich wymaga jednakowego opodatkowania dla osób mających te same dochody, dokonujących tych samych transakcji lub dysponujących majątkiem tej samej wartości. Innymi słowy jest to zasada przeciwdziałająca dyskryminacji. Z kolei równość pionowa oznacza ważne odstępstwo od zasady równości poziomej ze względu na zróżnicowaną zdolność do płacenia podatków w grupie osób posiadających takie same dochody lub taki sam majątek. Taki stan rzeczy może wynikać z różnego poziomu społecznie uzasadnionych wydatków lub kosztów bądź odmiennego udziału w dostarczanych dobrach publicznych. W pierwszym przypadku można przykładowo wskazać na porównanie sytuacji materialnej dwóch kobiet, które pracują i otrzymują takie samo wynagrodzenie, przy czym jedna z nich jest osobą samotną, a druga to samotna matka utrzymująca troje dzieci. Jest rzeczą oczywistą, że ze względu na konieczność ponoszenia kosztów związanych z wychowaniem dzieci (wychowanie dzieci jest uznawane za ważne zadanie społeczne), to druga z kobiet wykazuje niższą zdolność do płacenia podatków, w związku z czym powinna uiszczać niższy podatek. Z kolei w drugim przypadku można wskazać na porównanie sytuacji dwóch osób o jednakowych zarobkach, z których jedna porusza się po drogach publicznych wyłącznie samochodem, a druga korzysta tylko z transportu zbiorowego. W takiej sytuacji należy stwierdzić, że pierwsza z osób powinna uiszczać wyższy podatek drogowy, zawarty w cenie paliwa, ponieważ w większym stopniu korzysta indywidualnie z dróg publicznych. Dlatego też projektowanie systemu podatkowego z perspektywy sprawiedliwego rozłożenia ciężaru podatków musi odwoływać się w pierwszej kolejności do zasady równości poziomej, a następnie korygować ją pod względem ważnych celów społecznych, tak aby zapewnić także równość pionową, czyli rzeczywistą zdolność do płacenia podatków wśród poszczególnych osób (Stiglitz 2004, Gruber 2007).

4.3. Rodzaje podatków

W klasyfikacji podatków bierze się pod uwagę przede wszystkim przedmiot opodatkowania, lecz niekiedy dokonuje się jej także ze względu na cel opodatkowania, metodę obliczania oraz sposób rozłożenia ciężaru podatkowego między członków społeczeństwa.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

4.3.1. Podatki bezpośrednie

Zasadniczo rozróżnia się podatki bezpośrednie i podatki pośrednie. Pierwsze z nich, jak sama nazwa wskazuje, obciążają bezpośrednio dochody lub majątek. Najważniejsze z nich to podatek od dochodów indywidualnych, zwany w Polsce podatkiem dochodowym od osób fizycznych (*Personal Income Tax* – PIT), podatek dochodowy od osób prawnych, czyli od przedsiębiorstw (*Corporate Income Tax* – CIT), oraz podatek od dochodów kapitałowych (*capital gain tax*). W niektórych krajach oprócz podatku PIT (albo zamiast niego) stosuje się podatek od funduszu płac (*payroll tax*), uiszczany przez pracodawcę. Podatek PIT obciąża z reguły wszystkie dochody z tytułu jakiejkolwiek działalności danego podatnika, po uwzględnieniu pewnej kwoty wolnej od opodatkowania i po potrąceniu ewentualnych zwolnień, ulg oraz odliczeń. Bardzo podobnie liczony jest podatek CIT, który obciąża zysk przedsiębiorstwa, po odliczeniu ściśle określonych przepisami kosztów uzyskania przychodu. Podatek od dochodów kapitałowych dotyczy zaś takich środków jak dywidendy, odsetki od lokat bankowych, a także dochód ze sprzedaży aktywów, liczony jako różnica między ceną nabycia a ceną sprzedaży w określonym czasie.

Podatkiem bezpośrednim jest także podatek majątkowy, który może przybierać różne formy, występując na przykład w postaci podatku od nieruchomości (domy, działki budowlane, ziemia rolna, lasy) lub podatku od posiadanych dóbr luksusowych (jachty, baseny, domy letniskowe). Podstawą wymiaru podatku bezpośredniego jest wartość dochodu lub wartość majątku. Aby uniknąć nieporozumień co do wysokości obciążenia podatkowego, odpowiednie przepisy muszą określać w sposób precyzyjny metodę wyliczania podstawy opodatkowania.

4.3.2. Podatki pośrednie

Podatki pośrednie odprowadza się od dokonywanych transakcji kupna dóbr i usług, a niekiedy także aktywów, w tym finansowych. Najważniejsze z nich to podatek od wartości dodanej (*Value-Added Tax* – VAT), nazywany w Polsce podatkiem od towarów i usług (PTU), podatek obrotowy (*sales tax, turnover tax*) i podatek akcyzowy (*excise tax*). Podatek VAT jest podatkiem netto, ponieważ podstawą jego wymiaru jest wartość dodana na każdym etapie obrotu gospodarczego. Ze względu na sposób konstrukcji, podatek VAT pozwala uniknąć kumulowania

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

obciążeń podatkowych w poszczególnych fazach przetwórstwa (tzw. efekt kaskadowy) i przeciwdziała skłonności do ukrywania sprzedaży, ponieważ warunkiem odzyskania podatku zapłaconego w cenie materiałów stosowanych do produkcji jest uregulowanie podatku zawartego w cenie produktu końcowego (zob. ramka 4.1). Z kolei podatek obrotowy jest podatkiem brutto, uiszczanym najczęściej na każdym etapie obrotu gospodarczego. Pozostaje on prostszy i tańszy w administrowaniu niż podatek VAT, choć w większym stopniu zarazem zniekształca relacje cen i sprzyja ukrywaniu dochodów. W Europie powszechnie stosuje się podatek VAT, ale w niektórych krajach pozaeuropejskich nadal funkcjonuje podatek obrotowy (m.in. w USA, Kanadzie, Republice Południowej Afryki).

Ramka 4.1. Podatek VAT

Podatek VAT jest nowoczesną odmianą podatku obrotowego, opracowaną i zaproponowaną w 1954 roku przez francuskiego ekonomistę Maurice'a Laurégo. Jego najbardziej charakterystyczną cechą jest to, że obciąża on tylko wartość dodaną w danej fazie obrotu gospodarczego, a pomija wartość zużytych przy produkcji materiałów i usług. Rozliczanie podatku VAT polega więc na tym, że od podatku należnego od całej wartości sprzedaży podatnik może odliczyć podatek naliczony i zapłacony w poprzednich fazach obrotu, dzięki czemu unika się efektu kaskadowego (jest to tzw. podatek netto). O ile zwykły podatek obrotowy (podatek brutto) skłania do ukrywania przychodów i tworzenia pionowych struktur organizacyjnych w ramach jednego przedsiębiorstwa (ponieważ obrót wewnętrzny w przedsiębiorstwie jest zwolniony z podatku), o tyle podatek VAT sprzyja ujawnianiu kosztów i przychodów, tak aby zyskać możliwość odliczenia podatku naliczonego i zapłaconego w cenie zakupionych do produkcji dóbr i usług. Technicznie rzecz biorąc, rozliczanie podatku VAT polega na tym, że sprzedawca wpłaca do organu podatkowego podatek *należny* za dany okres (liczony jako iloczyn wartości sprzedaży w tym okresie i stawki procentowej podatku), pomniejszony o podatek *naliczony*, tj. o kwoty podatku wykazane na fakturach dostawców za dany okres. Przykładowy sposób rozliczania podatku VAT przedstawia tabela 4.2.

W większości krajów stosuje się zróżnicowane stawki podatku VAT – poza stawką podstawową, obciążającą większość dóbr i usług, występują też stawki obniżone, dotyczące dóbr i usług szczególnie ważnych, których spożycie jest pożądane lub niezbędne, oraz stawka zerowa (na przykład w odniesieniu do eksportu). W Polsce

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

podstawowa stawka podatku VAT od 1 stycznia 2011 roku jest ustalona na poziomie 23%, podczas gdy stawka obniżona wynosi 8% i dotyczy żywności przetworzonej oraz usług budowlanych w ramach budownictwa objętego społecznym programem mieszkaniowym. Ponadto stosowana jest tzw. stawka superobniżona w wysokości 5%, którą obciążone są podstawowe produkty żywnościowe (np. chleb, nabiał, przetwory mięsne, produkty zbożowe, mąka, kasze, makarony, soki), a także książki drukowane i wydawane na dyskach, taśmach i innych nośnikach oraz czasopisma specjalistyczne. W przypadku niektórych usług (transport lotniczy, turystyka) oraz w eksporcie stawka VAT wynosi 0%. Stawka zerowa w eksporcie (oraz w wewnątrzwspólnotowym obrocie towarami w ramach UE) pozwala unikać podwójnego opodatkowania z uwagi na to, że polski eksport jest obciążony stawką VAT w wysokości obowiązującej w kraju przeznaczenia. Warto podkreślić, że stawka zerowa w eksporcie nadal umożliwia eksporterom odliczanie podatku naliczonego.

Tabela 4.2. Przykład rozliczania podatku VAT przy stawce 22% (dane w PLN, jeśli nie podano inaczej)

Pozycja	Cena netto	Stawka VAT (22%)	Cena brutto	Podatek VAT
Sprzedaż	100	22	122	22,00
Zakup: materiał 1	40	22	48,8	8,80
Zakup: materiał 2	10	7	10,7	0,70
Podatek należny	x	x	x	22,00
Podatek naliczony	x	x	x	8,80 + 0,70 = 9,50
Podatek do zapłacenia	x	x	x	22,00 – 9,50 = 12,50

Źródło: opracowanie własne.

Akcyza jest szczególnym rodzajem podatku obrotowego, który obciąża tylko niektóre towary. Dotyczy to na ogół trzech grup produktów, takich jak:

a) wyroby, których konsumpcja jest niepożądana społecznie (niesie ujemne efekty zewnętrzne), jak na przykład wyroby alkoholowe i tytoniowe;
b) wyroby, których wytwarzanie i konsumpcja wymagają zużycia ograniczonych zasobów naturalnych, a jednocześnie popyt na te wyroby jest względnie sztywny, dzięki czemu ich sprzedaż zapewnia stabilne źródło dochodów w ramach budżetu państwa (jak na przykład paliwa samochodowe, węgiel, energia elektryczna);
c) wyroby luksusowe (na przykład samochody, jachty).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W ten sposób podatek akcyzowy pełni zarówno funkcję ekonomiczno-społeczną, zniechęcając do konsumpcji niektórych dóbr, jak i funkcję fiskalną, zapewniając stabilne dochody w ramach budżetu. Akcyza stanowi podatek jednofazowy (co oznacza, że pobierana jest tylko na jednym szczeblu obrotu), a jej stawka przyjmuje najczęściej postać kwotową na jednostkę towaru (jest to tzw. stawka specyficzna), choć dopuszcza się także stawki procentowe od wartości towaru (tzw. stawka *ad valorem*).

Wśród pozostałych form opodatkowania stosowanych w Polsce wymienia się podatek od gier hazardowych, podatki od eksploatacji zasobów naturalnych, podatki od środków transportu, a czasami także inne tytuły podatkowe (na przykład podatek od aktywów bankowych). Funkcje podatków pełnią również nakładane przez władze lokalne rozmaite opłaty nie będące formalnie podatkami, takie jak opłaty turystyczne, klimatyczne, manipulacyjne czy opłaty od posiadania zwierząt. Niekiedy do podatków (pośrednich) zalicza się także cła i opłaty wyrównawcze w handlu zagranicznym, choć mają one inny zasięg stosowania i niosą za sobą inne skutki ekonomiczne niż powszechnie obowiązujące podatki krajowe. Również składki na ubezpieczenia społeczne mogą być traktowane jako swoiste podatki (lub parapodatki), choć trzeba pamiętać o tym, że składka na ubezpieczenia społeczne upoważnia – w przeciwieństwie do zwykłego podatku – do otrzymywania konkretnego świadczenia ze strony państwa w postaci emerytury lub renty. W ramce 4.2 wymienione są wszystkie rodzaje podatków stosowanych w Polsce.

Ramka 4.2. Rodzaje podatków stosowanych w Polsce: tytuły podatkowe – stan na koniec 2016 roku

Polski system podatkowy zawiera łącznie 13 tytułów podatkowych objętych materialnym prawem podatkowym. Należą do nich:

1. Podatki pośrednie:
   1.1. Podatek od towarów i usług (VAT)
   1.2. Podatek akcyzowy
2. Podatki bezpośrednie dochodowe:
   2.1. Podatek dochodowy od osób prawnych (CIT)
   2.2. Podatek dochodowy od osób fizycznych (PIT)
3. Podatki od dochodów kapitałowych:
   3.1. Podatek od dochodów kapitałowych (odsetki bankowe, dywidendy)

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

3.2. Podatek od zbycia nieruchomości, papierów wartościowych, itp.
4. Podatek majątkowy
   4.1. Podatek od nieruchomości
   4.2. Podatek rolny
   4.3. Podatek leśny
5. Podatek od czynności cywilno-prawnych
6. Podatek od spadków i darowizn
7. Podatek od gier
8. Podatek tonażowy
9. Podatek od wydobycia niektórych kopalin
10. Podatek od aktywów bankowych.

Największe znaczenie w przypadku dochodów budżetu państwa mają podatki pośrednie (VAT i akcyza) oraz podatki dochodowe (PIT i CIT). W 2014 roku dochody z tych czterech tytułów podatkowych stanowiły 88,1% wszystkich dochodów budżetu państwa i 98,7% dochodów z tytułu podatków.

4.3.3. Efekt zniekształcenia i podatki korekcyjne

Jak zauważono już wcześniej, każdy podatek, dostarczając dochody do budżetu, ogranicza jednocześnie skalę działalności gospodarczej objętej podatkiem, a więc pociąga za sobą koszty w postaci utraconej produkcji i dochodu. Mówi się w związku z tym, że każdy podatek zniekształca funkcjonowanie mechanizmu rynkowego (*distortionary taxes*), ponieważ sprawia, że produkcja i dochód (a także zatrudnienie, ceny oraz płace) kształtują się w warunkach istnienia opodatkowania na poziomie odbiegającym od optimum. Przyczyny tego stanu rzeczy zostaną wyjaśnione w następnym punkcie rozdziału. Obecnie należy zapamiętać jedynie, że, ponieważ każdy podatek powoduje odchylenie wielkości zmiennych ekonomicznych od wielkości optymalnych obciążenie takie zawsze będzie wiązało się z pewnym kosztem ekonomicznym. Na koszt ten składają się zarówno straty społeczne w postaci utraconej produkcji oraz dochodu, jak i koszt poboru podatków. Wynika więc stąd, że nakładając (lub podnosząc) podatki, zawsze należy oszacować związane z tym koszty, a następnie porównać je z korzyściami, jakie ma przynieść opodatkowanie. Może się bowiem okazać, że – z perspektywy poziomu dobrobytu ogólnospołecznego – nałożenie takiego czy innego podatku przynosi w ostatecznym rozrachunku więcej szkody niż pożytku.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Zdarzają się jednak sytuacje, w których ta immanentna, wydawałoby się, wada podatku staje się jego istotną zaletą. Ma to miejsce wówczas, gdy głównym celem podatku nie jest dostarczenie dochodów do budżetu, lecz raczej ograniczenie określonej działalności. Dotyczy to tych wszystkich sytuacji, w których mamy do czynienia z niedoskonałościami rynku (*market failures*). Gdy dana działalność gospodarcza przynosi określone ujemne skutki zewnętrzne (*negative externalities*), jej opodatkowanie przekłada się na zwiększenie indywidualnego (prywatnego) kosztu produkcji do poziomu rzeczywistego kosztu społecznego. Wzrost kosztów skłania do zredukowania skali szkodliwej działalności i pozwala ewentualnie pokryć wywołane przez nią straty społeczne. Przykładem tego rodzaju działalności jest produkcja przemysłowa powodująca zanieczyszczanie środowiska. Podobnie dzieje się w przypadku pewnych dóbr, których konsumpcja wywołuje negatywne efekty zewnętrzne (alkohol, papierosy). Opodatkowanie spożycia tych dóbr przekłada się na zwiększenie indywidualnego (prywatnego) kosztu konsumpcji do poziomu rzeczywistego kosztu społecznego (uwzględniającego m.in. koszty leczenia osób spożywających produkty szkodliwe dla zdrowia). Tego typu podatki dotyczące zarówno produkcji, jak i konsumpcji, nazywa się podatkami korekcyjnymi (*corrective taxes*) lub podatkami piguwiańskimi (*Pigouvian taxes*) – od nazwiska brytyjskiego ekonomisty Arthura Pigou, który jako pierwszy wyjaśnił ich działanie (Pigou 1920).

Podatki korekcyjne nie zapewniają jednak zupełnej eliminacji skutków niedoskonałości rynkowych, choćby dlatego, że ich pobór kosztuje. Najlepszym rozwiązaniem (*first-best*) byłoby po prostu zaprzestanie szkodliwej działalności. Często okazuje się jednak, że jest to w praktyce niemożliwe, w związku z czym władza gospodarcza zmuszona jest sięgać po instrumenty podatkowe. Ponieważ stosowanie podatków zawsze pociąga za sobą również pewne koszty administracyjne i ekonomiczne, podatki korekcyjne traktowane są jako rozwiązanie suboptymalne, czasem zwane też rozwiązaniem następnym po najlepszym (lub rozwiązaniem drugim w kolejnoci – *second-best*).

4.4. Stawki podatkowe

4.4.1. Podatki ryczałtowe, proporcjonalne i progresywne

Stawka podatkowa jest relacją wielkości podatku do podstawy opodatkowania. Są one najczęściej wyrażone w procentach podstawy opodatkowania, choć niekiedy

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

stosuje się także stawki kwotowe na określoną jednostkę towaru (m.in. w podatku akcyzowym). Przykładowo w 2015 roku stawka podatku CIT w Polsce wynosiła 19% zysku, podstawowa stawka VAT stanowiła 23% ceny netto sprzedaży, a stawka podatku akcyzowego na benzynę silnikową była ustalona na poziomie 1540 PLN na 1000 litrów paliwa.

W szczególnych przypadkach stosuje się podatki ryczałtowe (*lump-sum taxes*), czyli określone jako należna kwota przypadająca na daną osobę bądź wskazaną transakcję. W średniowieczu powszechnie występował podatek pogłówny, stanowiący określoną kwotę (wyrażoną w pieniądzu lub w naturze), którą musiał uiszczać każdy poddany niezależnie od swojej sytuacji materialnej. Ponieważ podatek pogłówny nie wymagał szacowania ani wysokości dochodu, ani wartości majątku, jego pobór był stosunkowo prosty i tani. Odmianą podatku ryczałtowego było stałe myto pobierane za przejazd drogą lub mostem. Obecnie stosuje się podatki ryczałtowe w przypadku najmniejszych zakładów rzemieślniczych i usługowych, zwalniając je w ten sposób z obowiązku rejestrowania dochodów i prowadzenia dokumentacji księgowej.

W odniesieniu do podatków bezpośrednich (dochodowych) stawka podatkowa może być stała lub zmienna. Pierwsza z nich oznacza, że mamy do czynienia z podatkiem proporcjonalnym (nazywanym także podatkiem liniowym – *flat tax*), druga może mieć zaś charakter progresywny lub regresywny. Stawki progresywne zwiększają się w miarę wzrostu podstawy opodatkowania, podczas gdy stawki regresywne maleją wraz ze wzrostem podstawy opodatkowania (w praktyce stosuje się tylko stawki proporcjonalne lub progresywne). Stosowanie stawek progresywnych wymaga ustalenia progów progresji (wielkości granicznych), których przekroczenie powoduje zastosowanie wyższej stawki, a kolejno wyznaczane progi określają zakres progresji.

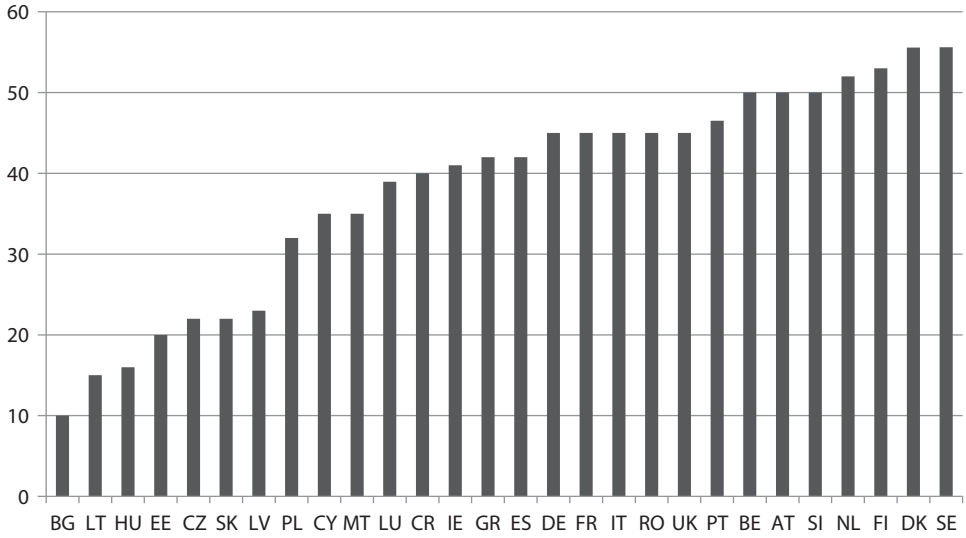
Progresja podatkowa stosowana jest głównie odnośnie do podatku PIT. W większości krajów Europy Zachodniej istnieje kilka stawek podatku dochodowego od osób fizycznych, przy czym stawki krańcowe w niektórych państwach są bardzo wysokie i przekraczają nawet 50% (na przykład w Szwecji wartość ta sięga 55,6%, w Danii 55,56%, w Finlandii 53%, a w Holandii 52%)[2]. W Europie Środkowo-Wschodniej popularny jest zaś podatek proporcjonalny (liniowy), który stosują m.in. Estonia, Litwa, Łotwa, Czechy, Węgry, Rumunia, Bułgaria, a także kraje nienależące

[2] Ogólnie rzecz biorąc, krańcowe stawki podatku PIT w Europie uległy znacznemu obniżeniu w okresie ostatnich czterdziestu lat. Jeszcze w latach 60. i 70. w wielu państwach stawki krańcowe przekraczały 80% (dotyczyło to m.in. Wielkiej Brytanii oraz w krajów skandynawskich).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

do UE, takie jak Ukraina i Rosja. Na wykresie 4.1 pokazane są wysokości stawek krańcowych podatku PIT w krajach UE – uderza tu zwłaszcza duże zróżnicowanie stawek, które wahają się od 10% (w Bułgarii) do ponad 55% (w Szwecji i Danii)[3].

Wykres 4.1. Krańcowe stawki podatku PIT w krajach Unii Europejskiej – stan na rok 2013 (dane w %)



Źródło: opracowanie własne na podstawie danych OECD i Eurostatu.

Stawki podatku CIT są mniej zróżnicowane, choć, podobnie jak w przypadku stawek PIT, są z reguły wyższe w krajach Europy Zachodniej niż wśród państw Europy Środkowej i wahają się między 10% (w Bułgarii) a 35% na Malcie (zob. wykres 4.2). Skąd właściwie wynika takie zróżnicowanie poziomów opodatkowania w ramach ugrupowania integracyjnego, jakim jest Unia Europejska? Głównym powodem są różnice w poziomie rozwoju. Podobnie jak w przypadku podatku PIT, kraje uboższe stosują z reguły również niższe stawki podatku CIT niż państwa zamożniejsze. Wysokie podatki silnie zachęcają do ich unikania, co staje się łatwiejsze w krajach o słabiej rozwiniętych systemach administracji podatkowej. Ma to też związek z kulturą płacenia podatków, znacznie silniej zakorzenioną w krajach Europy Zachodniej niż w wielu krajach Europy Środkowej, gdzie samodzielna państwowość i reguły gospodarki rynkowej wykształciły się później.

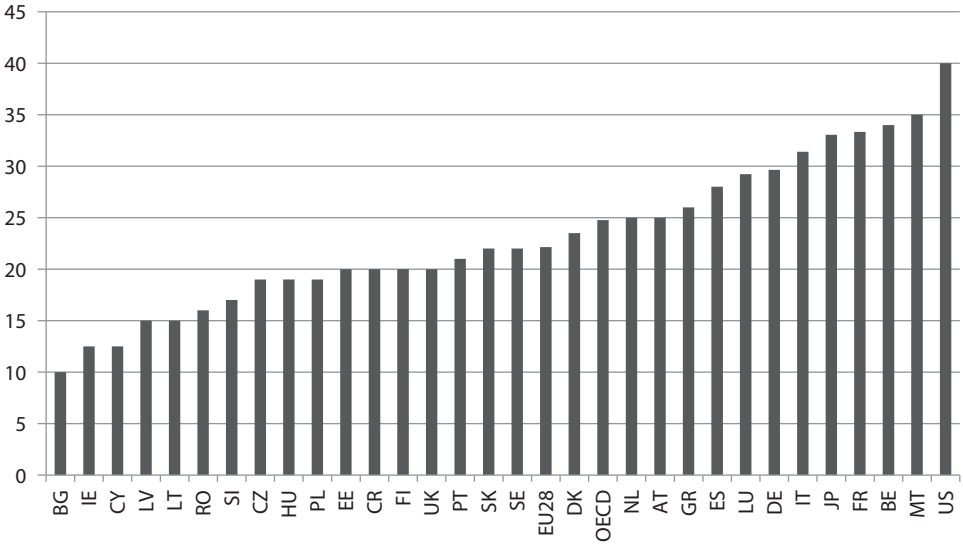
[3] W Danii podatek PIT zawiera także składki na ubezpieczenia społeczne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ponadto relatywnie niższe stawki podatków dochodowych odzwierciedlają także stosunkowo niższą podaż dóbr publicznych (infrastruktura transportowa, administracja publiczna, prawo i sądownictwo, opieka zdrowotna) w krajach uboższych niż wśród zamożniejszych państw. W warunkach swobody przepływu kapitału między państwami trudno jest krajom uboższym obciążać firmy wysokimi podatkami, gdy jakość i dostępność usług publicznych znajduje się tam na niższym poziomie niż w państwach zamożnych.

Zróżnicowanie poziomu opodatkowana w ramach UE wynika także z odmiennego stosunku do podatków wśród określonych partii politycznych. Partie lewicowe oraz socjaldemokratyczne są tradycyjnie zwolennikami wyższych (i bardziej progresywnych) podatków dochodowych w związku z dążeniem do finansowania większego zakresu usług publicznych. W przeciwieństwie do nich partie konserwatywne, a zwłaszcza liberalne, pozostają zwolennikami większej wolności gospodarczej, „małego" państwa i niskich podatków. Pewną rolę odgrywa również tradycja: w państwach skandynawskich podatki zawsze były stosunkowo wysokie, podczas gdy w krajach anglosaskich (Irlandia) oraz bałtyckich (Estonia, Litwa, Łotwa) kształtowały się zazwyczaj na niskim poziomie, co jest w pewnym stopniu wynikiem ogólnie bardziej liberalnego podejścia do polityki gospodarczej.

Wykres 4.2. Statutowe stawki podatku CIT w krajach UE, USA i Japonii – stan na rok 2014 (dane w %)



Źródło: opracowanie własne na podstawie danych OECD i Eurostatu.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

4.4.2. Nominalne i efektywne stawki podatkowe

Same nominalne stawki podatkowe nie informują jeszcze w pełni o poziomie rzeczywistych obciążeń podatkowych. Obciążenia te zależą także od takich czynników, jak zakres i skala ulg podatkowych, odliczeń oraz zwolnień, które łącznie wpływają na wysokość podstawy opodatkowania oraz wysokość płaconego podatku. Są kraje, w których stawki nominalne pozostają stosunkowo wysokie, lecz ze względu na dużą liczbę możliwych odliczeń od dochodu rzeczywiste obciążenie podatkiem okazuje się mniejsze niż w krajach stosujących niższą statutową stawkę podatku. W takiej sytuacji mówi się o różnicy między nominalną a efektywną stopą opodatkowania. Gdy wziąć pod uwagę wpływ na gospodarkę, wówczas ta druga z nich ma oczywiście większe znaczenie.

Tabela 4.3. Nominalne i efektywne obciążenie podatkiem PIT (wartości średnie) – dane za 2015 rok

Przedziały dochodów w PLN	Nominalna stawka podatkowa	Obciążenie dochodu w %			Różnica w obciążeniu (punkty procentowe) (3–4)	Różnica w obciążeniu (punkty procentowe) (3–5)
		nominalne	efektywne			
			łącznie ze składką na ubezpieczenie zdrowotne	po odliczeniu składki na ubezpieczenie zdrowotne		
1	2	3	4	5	6	7
Do 85 528	18% minus kwota zmniejszająca podatek 556, 02 PLN	15,49	14,25	6,82	1,24	8,67
Powyżej 85 528	14 839, 02 PLN plus 32% nadwyżki ponad 85 528 PLN	22,99	22,02	15,41	0,97	7,58

Źródło: opracowanie własne na podstawie danych Ministerstwa Finansów.

W tabeli 4.3 przedstawiono różnicę między obciążeniem nominalnym a obciążeniem efektywnym w odniesieniu do stosowanego w Polsce podatku PIT przy nominalnych stawkach 18% i 32%, biorąc pod uwagę wpływ odliczeń z tytułu kwoty wolnej oraz składki na ubezpieczenia społeczne i składki na ubezpieczenia zdrowotne, lecz bez uwzględniania innych ulg i odliczeń funkcjonujących w polskim prawie podatkowym. Jak widać, osoby osiągające dochody nieprzekraczające kwoty 85 528 PLN, a zatem płacące nominalnie podatek według pierwszej stawki 18%, po potrąceniu odliczeń płacą efektywnie stawkę podatku średnio w wysokości

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

6,82% wartości dochodu. Z kolei osoby zarabiające ponad 85 528 PLN płacą efektywną stawkę podatku w wysokości 15,41%.

W odniesieniu do podatków pośrednich stałe stawki podatku są z reguły traktowane jako stawki regresyjne. Wynika to z tego, że stała stawka podatku pośredniego nie zmienia się wraz z wysokością dochodu danego podatnika. Na przykład jeśli osoba, która miesięcznie zarabia 2000 PLN (netto), przeznacza na niezbędne zakupy 100% swoich dochodów, a średnia stawka VAT wynosi 20%, rzeczywiste obciążenie podatkiem VAT tej osoby wynosi 333,3 PLN / 2000 PLN = 16,7%. Z kolei osoba, która zarabia 6000 PLN i przeznacza na zakupy tylko 50% swoich dochodów, jest obciążona podatkiem VAT w wysokości zaledwie 561,0 PLN, co oznacza rzeczywiste obciążenie podatkiem VAT w wysokości 561,0 PLN / 6000 PLN = 9,3%, mimo że jej realna konsumpcja jest o połowę większa niż w przypadku osoby uboższej. Tak implikowana regresywność podatku jest głównym uzasadnieniem do stosowania obniżonych stawek VAT na artykuły pierwszej potrzeby, takie jak m.in. żywność, których udział w wydatkach uboższych grup społecznych jest największy. Obniżone stawki VAT neutralizują więc efekt regresywności i oznaczają niższe efektywne opodatkowanie dochodu uboższej części społeczeństwa.

Zróżnicowane stawki VAT nie są jednak najlepszym rozwiązaniem, ponieważ powodują zniekształcenie cen rynkowych i mogą prowadzić do nieoptymalnej struktury konsumpcji na skutek efektu substytucji[4]. Ponadto z obniżonych stawek VAT korzystają także zamożne grupy społeczne, co trudno uzasadnić społecznie. Alternatywnym sposobem ograniczenia stopnia regresywności podatku VAT w warunkach stosowania jednolitej stawki podatkowej jest wprowadzenie systemu bezpośredniego subsydiowania dochodów grup najuboższych.

4.5. Pożądane cechy systemu podatkowego

Podstawowym celem stosowania podatków jest zapewnienie środków na pokrycie niezbędnych wydatków budżetowych. Oprócz celu fiskalnego podatki mogą jednak służyć też zaspokajaniu innych ważnych celów, takich jak bardziej sprawiedliwy podział dochodów czy zniechęcanie do podejmowania pewnych rodzajów działań uznanych za społecznie niepożądane. Stosowanie podatków niesie za

[4] Efekt substytucji jest wynikiem tego, że spożycie towarów obciążonych niższą stawka VAT rośnie kosztem spadku spożycia towarów obciążonych wyższą stawka VAT.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

sobą zarazem rozmaite konsekwencje ekonomiczne, przy czym część z nich może być szkodliwa dla rozwoju gospodarczego. Budowa sprawnego, efektywnego systemu podatkowego pozostaje więc trudnym zadaniem, ponieważ wymaga pogodzenia wielu często rozbieżnych celów i priorytetów. Ogólnie rzecz biorąc, władza publiczna stara się tak skonstruować system podatkowy, aby wykazywał on cztery podstawowe cechy: efektywność fiskalna, efektywność ekonomiczna, niski koszt funkcjonowania i sprawiedliwość społeczna[5].

Jak już zauważono wcześniej, każdy podatek, dostarczając dochody do budżetu, ogranicza jednocześnie skalę działalności gospodarczej objętej podatkiem, a więc pociąga za sobą koszty w postaci utraconej produkcji i dochodu. Władza publiczna musi więc tak zaprojektować system podatkowy, aby był on dostatecznie wydolny fiskalnie, czyli aby zapewniał niezbędne dochody do budżetu i zarazem nie powodował nadmiernych strat ekonomicznych. Dlatego też wymaga się od systemu podatkowego, aby był efektywny zarówno fiskalnie, jak i ekonomicznie.

Trzeba następnie pamiętać o tym, że system podatkowy pociąga za sobą określone koszty. Gromadzenie podatków wymaga stworzenia administracji podatkowej, odpowiedzialnej za wymiar, pobór i kontrolę ściągalności podatków. Powołanie i funkcjonowanie administracji podatkowej absorbuje zasoby ludzkie i materialne, które w ten sposób wyłączane są z działalności gospodarczej i nie mogą być wykorzystane przy produkcji dóbr i usług. Ponadto sama realizacja obowiązków podatkowych wiąże się z ponoszeniem określonych kosztów przez uczestników życia gospodarczego. W związku z tym kolejne pożądane cechy systemu podatkowego to prostota administracyjna oraz odpowiednio niski koszt zarządzania, określany także jako zasada taniości opodatkowania.

Wreszcie, po czwarte, podatki są także ważnym instrumentem redystrybucji dochodu na obszarze danego kraju. Spełniając tę funkcję, system podatkowy ma realizować postulat uwzględnienia sprawiedliwości społecznej, czyli zapewniać w pożądanym zakresie równość poziomą i równość pionową.

Poniżej omówione zostaną szczegółowo cztery wymienione cechy systemu podatkowego.

[5] Joseph Stiglitz wyróżnia pięć podstawowych cech „dobrego" systemu podatkowego, nazywanych też przez niego „zasadami opodatkowania": efektywność ekonomiczna, prostota administracyjna, elastyczność, odpowiedzialność polityczna i sprawiedliwość (Stiglitz 2004, s. 550–551). Brakuje wśród nich efektywności fiskalnej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

4.5.1. Efektywność fiskalna

Przez pojęcie efektywności fiskalnej rozumiemy zdolność systemu podatkowego do gromadzenia dochodów w wysokości wymaganej przez zaplanowane w budżecie wydatki w czasie zapewniającym terminową realizację zobowiązań budżetowych. Na efektywność fiskalną wpływ ma wiele czynników, takich jak łatwość poboru podatku, zakres i stabilność bazy podatkowej, rodzaj podstawy opodatkowania, zróżnicowanie stawek oraz zakres preferencji i ulg podatkowych. Ogólnie rzecz biorąc, im mniej skomplikowany jest system podatkowy, tym łatwiejszy staje się pobór podatku i stabilniejsze dochody. Podobnie też im szersza i bardziej stabilna jest baza podatkowa (czyli zakres dochodów lub zakres transakcji gospodarczych objętych opodatkowaniem), tym łatwiej przychodzi pozyskiwać podatki i tym stabilniejsze są dochody podatkowe.

Pod tym względem podatki pośrednie wykazują wyższość nad podatkami dochodowymi. Podstawę opodatkowania w przypadku podatków pośrednich stanowi zasadniczo sprzedaż detaliczna, tj. poziom konsumpcji w gospodarce, w przeciwieństwie do podatków dochodowych, gdzie podstawą opodatkowania jest zysk firmy lub dochód indywidualny. Konsumpcja pozostaje zmienną znacznie bardziej stabilną w przebiegu cyklu koniunkturalnego niż zyski firm bądź dochody indywidualne. W okresie złej koniunktury zyski silnie maleją, a część firm ponosi straty, co silnie obniża wartość podatku od zysku. Podobnie też dzieje się w przypadku dochodów indywidualnych, które w czasie recesji maleją silniej niż konsumpcja, ponieważ cześć osób traci pracę i przestaje zarabiać, co odbija się znacząco na dochodach podatkowych. Z kolei poziom konsumpcji zmienia się w niewielkim stopniu dzięki wykorzystywaniu oszczędności i tendencji do „wygładzania" konsumpcji w czasie. Dlatego też dochody budżetu państwa są bardziej stabilne i mniej wrażliwe na przebieg koniunktury, gdy pochodzą głównie z podatków pośrednich.

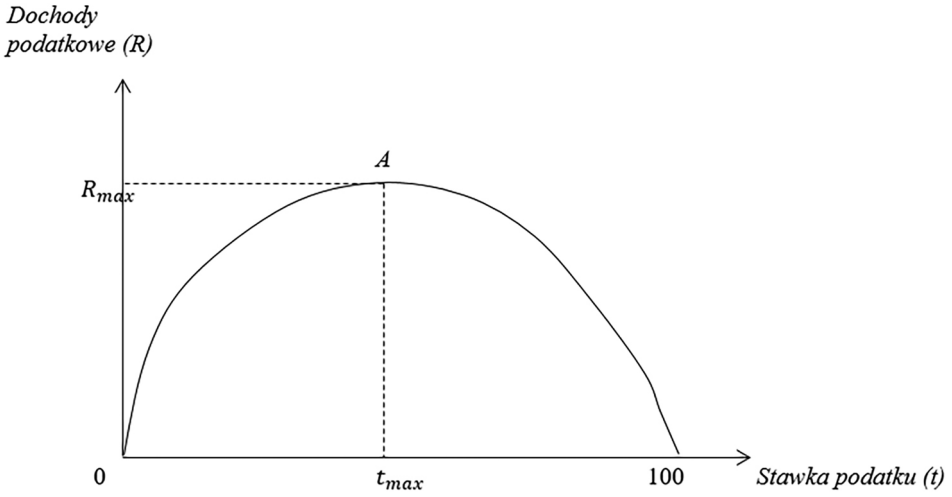
Na przykład dane dotyczące Polski z lat 2000–2014 potwierdzają tezę o wyższej zmienności w czasie dochodów pochodzących z podatków bezpośrednich. Porównanie zmian udziału podatków pośrednich i bezpośrednich w PKB pokazuje, że w okresie tym stosunek odchylenia standardowego do średniej wyniósł w przypadku podatków bezpośrednich 0,112 i był ponad dwa razy większy niż w przypadku podatków pośrednich (0,049)[6].

[6] Obliczenia przeprowadzono na podstawie danych o wykonaniu budżetu państwa – zob. http://www.mf.gov.pl/ministerstwo-finansow/dzialalnosc/finanse-publiczne/budzet-panstwa/wykonanie-budzetu-panstwa/sprawozdanie-z-wykonania-budzetu-panstwa-roczne (3.11.2015).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Efektywność fiskalna zależy też od poziomu opodatkowania. Ekonomiści już dawno zauważyli, że o dochodach budżetu decyduje wysokość stawek podatkowych – w miarę jak stawka podatku rośnie, dochody najpierw wzrastają, lecz po przekroczeniu pewnego poziomu opodatkowania zaczynają maleć. Zjawisko to określa się mianem efektu Laffera, czego graficzną ilustrację stanowi słynna krzywa Laffera (podobna do omówionej wcześniej krzywej Barro), nazwana tak na cześć amerykańskiego ekonomisty Arthura Laffera, który spopularyzował to zjawisko (rysunek 4.1).

Rysunek 4.1. Krzywa Laffera (dane w %)



Źródło: opracowanie własne na podstawie Laffer (2004).

Krzywa Laffera pokazuje znaną zależność dochodów budżetowych od wysokości stawki podatkowej. Oznaczmy dochody podatkowe jako R, gdzie $R = t \cdot B$, t jest stawką podatku (w %), a B stanowi bazę podatkową, czyli dochód do opodatkowania lub inaczej podstawę opodatkowania. Gdy stawka podatkowa t wynosi 0%, dochody podatkowe R będą równe zeru bez względu na wielkość B. Z kolei, gdy stawka podatkowa t wynosi 100%, dochody podatkowe też będą równe zeru, ponieważ nikt nie będzie skłonny pracować i produkować w sytuacji, gdy całość wytworzonego dochodu jest odbierana w formie podatku – wówczas baza podatkowa B okaże się równa zeru. Między tymi skrajnościami zachodzi *continuum* stawek podatkowych, zapewniających różny poziom dochodów podatkowych. Skoro

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w przypadku granicznych wartości danego przedziału $R = 0$, to wewnątrz niego musi istnieć jakaś wartość t, dla której R osiąga maksimum.

Zauważmy, że wielkość R jest funkcją dwóch zmiennych, t i B, które są od siebie zależne w sposób odwrotnie proporcjonalny – gdy t rośnie, B spada, i odwrotnie. Laffer uważał, że charakter tej zależności polega na tym, że w miarę wzrostu stawki podatkowej do pewnego jej poziomu efekt wzrostu stawki przeważa nad efektem spadku bazy podatkowej, w rezultacie czego R rośnie, podczas gdy powyżej tego poziomu efekt spadku bazy podatkowej przeważa nad efektem wzrostu stawki podatkowej, w związku z czym R spada. Dlatego też krzywa Laffera ma kształt paraboli otwartej do dołu (dla wartości $R > 0$).

Efekt Laffera był wielokrotnie przedmiotem badań empirycznych, które generalnie potwierdzają istnienie zależności opisanej krzywą Laffera, choć ich wyniki różnią się co do wysokości stawki zapewniającej największe dochody. Na przykład Stuart (1981) i Pecorino (1995) szacują t_{max} na około 70%. Z kolei Hsing (1996) wylicza, że stawka ta jest dwa razy niższa (oscyluje między 32,67% a 35,21%). Nowsze badania potwierdzają raczej wcześniejsze szacunki, pozwalające stwierdzić, że krzywa Laffera jest niesymetryczna, a jej punkt maksimum mieści się w przedziale 60–70% (zob. Trabandt i Uhlig 2009).

Na efektywność fiskalną oddziałuje również elastyczność i odporność systemu podatkowego. Przez to pojęcie rozumiemy zdolność do szybkiego reagowania systemu na szoki zewnętrzne, powodujące spadek dochodów budżetowych na skutek zmniejszenia się bazy podatkowej. Jak już wiadomo, najprostsze rozwiązanie, polegające na podniesieniu stawek podatkowych, jest w praktyce trudne do zastosowania, ponieważ decyzję o podwyżce podatków (lub o nałożeniu nowego podatku) podejmuje parlament, co wymaga politycznych uzgodnień i z reguły zajmuje dużo czasu. Ponadto, w odniesieniu do podatków bezpośrednich, w państwach demokratycznych obowiązuje zasada niezmieniania podatków w trakcie roku fiskalnego oraz zasada *vacatio legis*, dotycząca zachowania pewnego odstępu czasu między publikacją aktu prawnego wprowadzającego podwyżkę podatków a jego wejściem w życie, co ma zapewnić podatnikom czas niezbędny na przystosowanie się do nowych przepisów. Nie bez znaczenia wreszcie pozostaje opóźnienie w spływie dochodów podatkowych, wynikające z potrzeby dokonania pełnych rozliczeń i konieczności zamknięcia okresów sprawozdawczych. Wszystko to silnie ogranicza możliwość podnoszenia podatków w krótkim czasie w sytuacji nagłego niedoboru dochodów.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Jak zatem zapewnić niezbędną elastyczność i odporność systemu podatkowego? Najlepszym sposobem jest w tym przypadku poszerzanie bazy podatkowej, tak aby obejmowała ona możliwie wiele sektorów, rodzajów działalności i grup społecznych, ponieważ wówczas rośnie szansa na to, że skutki podatkowe szoków egzogenicznych mogą się wzajemnie znosić (pojęcie optymalnej bazy podatkowej zostało omówione w ramce 4.3). Aby uniknąć jednak przy tym nadmiernych obciążeń podatkowych w skali kraju, stawki podatku powinny być ustalane na umiarkowanym poziomie. Jeśli podatki są bowiem wysokie, to nie tylko ograniczają one produkcję i dochody, lecz nasilają także bodźce skłaniające do ich unikania, co dodatkowo obniża dochody podatkowe. Zbyt wysokie podatki przekładają się też na wzrost zaangażowania zasobów potencjalnie produkcyjnych w nieprodukcyjne czynności związane z poszukiwaniem możliwości unikania opłat podatkowych. Oznacza to spadek potencjalnego PKB w kraju. Drugim sposobem jest uzależnienie systemu podatkowego w jak największym stopniu od podatków pośrednich oraz podatków majątkowych, w przypadku których baza podatkowa pozostaje znacznie bardziej stabilna niż baza podatków dochodowych. Odporność systemu podatkowego jest również wyższa w sytuacji, gdy stosuje się podatki jednolite, z możliwie najmniejszą liczbą wyjątków i odstępstw od reguł, co osłabia bodźce skłaniające do unikania podatków poprzez wykorzystywanie istniejących ulg oraz luk i nieszczelności w przepisach.

Ramka 4.3. Optymalna baza podatkowa

Baza podatkowa to określona wielkość ekonomiczna (dochód, majątek, wartość sprzedaży), która jest obciążona podatkiem. W odniesieniu do dochodu za najbardziej poprawną teoretycznie uważana jest całościowa definicja dochodu Schanza-Haiga-Simonsa (*Schanz-Haig-Simons comprehensive income definition* – SHS), określająca podstawę podatku dla danej osoby jako jej „zdolność do płacenia podatku" (*individual's ability to pay taxes*). Szacowana zdolność jest równa rocznej „potencjalnej konsumpcji" w przypadku danej osoby, czyli jej rocznej całkowitej odpłatnej konsumpcji, powiększonej o wzrost zasobu majątku w tym okresie oraz o bezpłatne lub częściowo płatne świadczenia, które wybrana osoba otrzymuje od pracodawcy, takie jak na przykład samochód służbowy, telefon służbowy, ubezpieczenie zdrowotne, deputat węglowy czy paczki świąteczne. Definicja dochodu obejmuje więc nie tylko dochód pieniężny danej osoby, lecz także

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wszystkie składniki konsumpcji bezpłatnej oraz przyrost netto majątku tej osoby (zob. Gruber 2007, s. 525–526).

Definicja SHS spełnia postulat równości poziomej, ponieważ zapewnia te same obciążenia ludzi osiągających tę samą konsumpcję, niezależnie od tego, czy jest ona finansowana z otrzymanego dochodu pieniężnego, czy dostarczana bezpłatnie. Spełniony zostaje również postulat równości pionowej, ponieważ uwzględnia się rzeczywistą zdolność od płacenia podatku.

Jednakże stosowanie w praktyce definicji SHS wymaga przyjmowania odstępstw i wyjątków. Przykładowo uwzględnienie w podstawie opodatkowania ubezpieczenia zdrowotnego opłacanego przez pracodawcę może budzić wątpliwości w sytuacji, gdy ubezpieczenie pokrywa koszty wypadku czy niezbędnego leczenia. Ponadto definicja SHS nie uwzględnia kosztów uzyskania dochodów. Wreszcie też przy zastosowaniu wskazanej reguły nie bierze się pod uwagę tego, że w wielu krajach część wydatków odejmuje się od podstawy opodatkowania wówczas, gdy przeznaczone są na pożądane cele społeczne, takie jak na przykład cele charytatywne.

4.5.2. Koszt funkcjonowania systemu podatkowego

Jak wspomniano już wyżej, zarządzanie systemem podatkowym pociąga za sobą określone koszty (Slemrod 1990). Obejmują one wydatki na utrzymanie instytucji administracji podatkowej, zajmującej się wymiarem, poborem i kontrolą ściągalności podatków. Nazywa się je kosztami *bezpośrednimi*, ponieważ są stosunkowo łatwe do policzenia. Wydatki na administrację podatkową zostają każdorazowo zapisane w budżecie państwa[7].

Lecz poza kosztami bezpośrednimi występują także koszty *pośrednie*, związane z wykonywaniem zobowiązań podatkowych przez samych podatników (zob. Stiglitz 2004, s. 560–562). Ten rodzaj wydatków jest trudniejszy do oszacowania. O ile bowiem w dużych firmach można wyliczyć koszt utrzymania działów księgowości lub wielkość wydatków na doradców podatkowych, o tyle w małych firmach

[7] W Polsce organami administracji publicznej odpowiedzialnymi za te działania są wojewódzkie, powiatowe i dzielnicowe urzędy skarbowe, izby skarbowe pełniące wobec urzędów funkcje nadzorcze, kontrolne i odwoławcze, oraz urzędy kontroli skarbowej. Łączne zatrudnienie w tych organach władzy podatkowej wynosiło w 2013 roku około 43 tysięcy pracowników, a bezpośredni koszt utrzymania tego aparatu (uwzględniając także urzędników Ministerstwa Finansów zatrudnionych w departamentach związanych z podatkami) przekroczył 4 mld PLN – równowartość około 1,25% wartości podatków gromadzonych przez budżet.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wyodrębnienie tych kosztów jest trudne lub wręcz niemożliwe do przeprowadzenia. Ponadto w grę wchodzą także koszty kontroli skarbowych ponoszone przez podatników, koszty sporządzania i przechowywania niezbędnej dokumentacji podatkowej, koszty postępowań prawnych i sądowych, wreszcie koszt „uczenia się" oraz błędów popełnianych przez firmy i osoby fizyczne w związku z koniecznością wykonywania zobowiązań podatkowych. Ocenia się, że w rozwiniętych gospodarkach, łącznie rzecz biorąc, pośrednie koszty funkcjonowania systemu podatkowego są kilka razy wyższe niż koszty bezpośrednie (Burmann i Slemrod 2012). Na koniec trzeba też wspomnieć o korupcji, która powoduje spadek dochodów państwa wskutek utraconych podatków, a na dodatek działa jak swoisty „podatek" nałożony na działalność gospodarczą który obniża ogólną efektywność gospodarowania, ponieważ zniekształca bodźce rynkowe, stwarzając nierówne warunki konkurencji między firmami.

Wszystkie te koszty należy odjąć od uzyskanych dochodów podatkowych, aby wyliczyć dochody netto do budżetu. W krajach o niesprawnych, skorumpowanych i opresyjnych administracjach podatkowych koszty funkcjonowania systemu są duże w relacji do uzyskiwanych dochodów. W rezultacie państwo nie posiada środków na sfinansowanie najpilniejszych potrzeb społecznych, a spora część ograniczonych zasobów jest zarazem zaangażowana w mało produkcyjne lub nieprodukcyjne działania administracyjne kosztem działalności bezpośrednio produkcyjnej.

Jak sprawić zatem, aby system podatkowy nie był zbyt kosztowny, a jednocześnie spełniał postulat efektywności fiskalnej? Za główne czynniki kształtujące koszty uznaje się strukturę systemu podatkowego (podatki pośrednie *vs* bezpośrednie) oraz jego złożoność. Pod względem struktury najtańsze w poborze są podatki pośrednie, zwłaszcza te najprostsze, takie jak podatek obrotowy i akcyza. Z kolei złożoność systemu to funkcja stopnia zróżnicowania stawek podatkowych oraz liczby ulg, zwolnień i odliczeń. Im więcej pojawia się wyjątków i odstępstw od ogólnych reguł, tym bardziej skomplikowany staje się system i tym większe pociąga to za sobą koszty wymiaru, poboru, kontroli oraz możliwości nadużyć, w tym korupcji.

Nadmiernie złożone systemy podatkowe otwierają też dodatkowe możliwości unikania płacenia podatków. Podatnicy wykorzystują luki w przepisach oraz ich wewnętrzną niespójność w celu obniżenia należnych podatków. Tego rodzaju legalne unikanie płacenia należnych podatków (*tax avoidance*) nazywa się też niekiedy optymalizacją podatkową lub agresywnym planowaniem podatkowym (*aggressive tax planning*). Z kolei nadmiernie opresyjne systemy podatkowe, cechujące się wysokimi stawkami podatku, zachęcają do ukrywania dochodów i przenoszenia

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

działalności gospodarczej do tzw. szarej strefy. Mamy wówczas do czynienia z nielegalnym unikaniem podatku (*tax evasion*), co stanowi działalność naruszającą prawo.

Dlatego też niezbyt kosztowny i zarazem wydajny system podatkowy to system o jednolitych, umiarkowanych stawkach, możliwie najmniejszej liczbie ulg, zwolnień i odliczeń oraz w jak największym stopniu zależny od podatków pośrednich.

4.5.3. Efektywność ekonomiczna: wpływ podatków na produkcję i konsumpcję

Jak już wielokrotnie wspominano, każdy podatek wywołuje podwójny efekt – z jednej strony przynosi dochód władzy publicznej, z drugiej zaś ogranicza skalę działalności objętej podatkiem. Dlatego też system podatkowy powinien być skonstruowany w taki sposób, aby przynosił państwu niezbędne dochody możliwie jak najmniejszym kosztem ekonomicznym. Poniżej wyjaśnimy, dlaczego podatki przekładają się na straty produkcji i dochodu, a następnie pokażemy, jak można je ograniczyć.

Jak pamiętamy z rozdziału pierwszego, gospodarka znajduje się w punkcie optymalności Pareto, gdy dochód krańcowy (cena) zrównuje się z kosztem krańcowym. Wprowadzenie podatku zaburza tę równowagę i pociąga za sobą zniekształcenia w strukturze bodźców zachęcających do produkcji, pracy i oszczędności. Ludzie nie pozostają bierni wobec obowiązku uiszczania podatków – gdy państwo zabiera im część wytworzonego dochodu lub zwiększa koszt nabywania dóbr i usług, dostosowują się do tej sytuacji poprzez zmianę dotychczasowego zachowania, z reguły ograniczając skalę produkcji, podaży pracy bądź oszczędności. Podatki wywołują też zmiany w *strukturze* produkcji, inwestycji oraz konsumpcji. Wszystkie te zniekształcenia powodują spadek efektywności ekonomicznej – gospodarka wytwarza i konsumuje mniej, niż byłoby to potencjalnie możliwe, gdyby podatków nie nakładano.

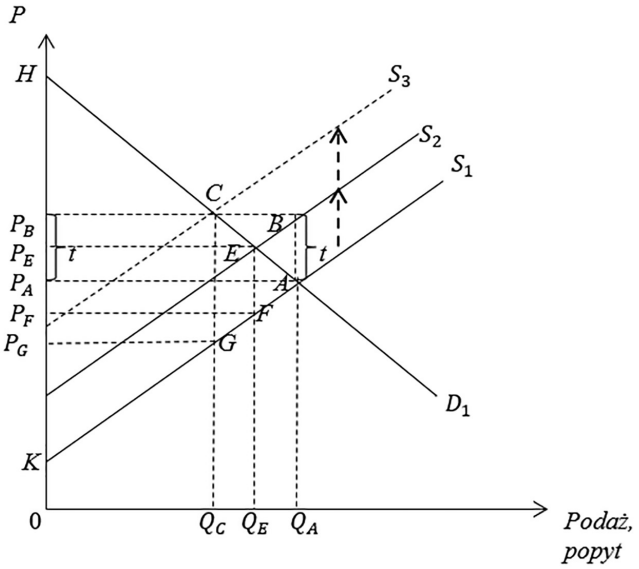
Rysunek 4.2 ilustruje wpływ nałożenia podatku na podaż, popyt i cenę na rynku określonego produktu *x*. Początkowo równowaga zachodzi w punkcie przecięcia krzywej podaży S_1 i krzywej popytu D_1 (punkt *A*). Cena równowagi wynosi P_A. Co stanie się, gdy na cenę sprzedaży nałożymy podatek w wysokości *t* PLN?

Zastosowanie podatku *t* przesuwa w pierwszym momencie krzywą podaży z punktu S_1 do S_2, ponieważ producenci początkowo usiłują podnieść cenę o wielkość podatku, do poziomu P_B. Przy tej cenie popyt spada jednak do poziomu Q_C, co zmusza producentów do obniżenia ceny. Przekłada się to na stopniowy wzrost

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

popytu i rynek stabilizuje się wówczas w punkcie E, przy cenie P_E. Skutkiem nałożenia podatku jest spadek produkcji i popytu (z Q_A do Q_E), wzrost ceny płaconej przez odbiorców (z P_A do P_B) i wreszcie spadek cen otrzymywanych przez dostawców (z P_A do P_F). Między ceną dla odbiorców a ceną dla dostawców pojawia się klin podatkowy (*tax wedge*), w wysokości równej $t = AB = EF$.

Rysunek 4.2. Wpływ nałożenia podatku na podaż, popyt i cenę



Źródło: opracowanie własne.

Można zauważyć więc, że dostarczane i nabywane ilości dobra x zmalały, a ponadto konsumenci płacą wyższe ceny, podczas gdy producenci uzyskują niższe ceny niż poprzednio. Wynika z tego, że – zgodnie z wcześniej postawioną tezą – nałożenie podatku ogranicza działalność nim obciążoną, tj. produkcję i konsumpcję dobra x[8].

[8] Podatek zmniejsza dochód „na krańcu", gdzie – zgodnie z zasadą malejącej krańcowej użyteczności dochodu – relacja użyteczności krańcowego dochodu do czasu wolnego (*leisure*) jest najmniejsza. Dlatego też nałożenie podatku skłania do ograniczenia pracy i działalności gospodarczej na rzecz czasu wolnego. Każdy podatek obniża więc efektywność gospodarowania, ponieważ ogranicza podaż pracy (spada zatrudnienie i maleje dochód z pracy) oraz skalę działalności gospodarczej (spada poziom zarówno produkcji, jak i uzyskiwanego dochodu).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Lecz dzięki zastosowaniu podatku rząd otrzymuje dodatkowe dochody, których suma równa jest iloczynowi tQ_E. Mamy więc do czynienia z przesunięciem zasobów od konsumentów i producentów do rządu. Czy można wówczas określić, jaki jest ogólny bilans tych zmian – czy społeczeństwo jako całość ma się lepiej po nałożeniu podatku, czy gorzej? Odpowiemy na to pytanie, wykorzystując pojęcie nadwyżki konsumenta (*consumer surplus*) i nadwyżki producenta (*producer surplus*).

Można przyjąć, że początkowo wszyscy konsumenci płacą za jednostkę dobra *x* jednakową cenę P_A. Przebieg krzywej popytu pokazuje jednak, że wielu z nich byłoby gotowych zapłacić cenę wyższą od P_A – na przykład przy wyższej cenie P_B popyt wynosi nadal Q_C. Oznacza to, że wszyscy konsumenci, którzy gotowi byliby zapłacić więcej niż P_A, odnoszą w tej sytuacji dodatkową korzyść, ponieważ kupują dobra po cenie P_A. Suma tych dodatkowych korzyści nazywana jest nadwyżką konsumenta i pozostaje równa powierzchni trójkąta AHP_A, wyznaczonego przez krzywą popytu D_1, linię ceny P_A oraz pionową oś *P*. Podobnie jest z producentami – przebieg krzywej podaży wskazuje, że wielu z nich byłoby gotowych sprzedawać swoje produkty poniżej ceny P_A – na przykład przy niższej cenie P_G podaż nadal wynosiłaby Q_C. Oznacza to, że wszyscy producenci, którzy gotowi byliby zaakceptować cenę niższą niż P_A, odnoszą w tej sytuacji dodatkową korzyść, ponieważ sprzedają swoje dobra po cenie P_A. Suma tych dodatkowych korzyści nazywana jest nadwyżką producenta i pozostaje równa powierzchni trójkąta AKP_A, wyznaczonego przez krzywą popytu S_1, linię ceny P_A oraz pionową oś *P*. Suma nadwyżki konsumenta i nadwyżki producenta odpowiada wielkości społecznego dobrobytu (*social welfare*), którym cieszy się dane społeczeństwo dzięki produkcji i konsumpcji dobra *x* na poziomie wyznaczonym przez punkt *A* – jest to powierzchnia trójkąta *AHK*.

Nałożenie podatku zmienia sytuację. Nadwyżka konsumenta zmniejsza się, ponieważ jest teraz równa powierzchni trójkąta EHP_E, a cena rośnie do $P_E > P_A$. Nadwyżka producenta również zmniejsza się, gdyż jest teraz równa powierzchni trójkąta FKP_F, a cena spada do $P_F < P_A$. Owszem, rząd odnosi wprawdzie korzyść z uzyskanych podatków – wielkość dochodów z tego tytułu jest równa iloczynowi tQ_E, czyli powierzchni prostokąta EFP_FP_E. Lecz łatwo zauważyć, że łączna suma nadwyżki konsumenta, nadwyżki producenta oraz dochodów budżetu jest teraz mniejsza niż początkowa wielkość dobrobytu reprezentowana przez powierzchnię trójkąta *AHK*. Różnica ta jest równa powierzchni trójkąta *AEF*. Dowodzi to, że nałożenie podatków zmniejsza poziom dobrobytu społecznego. Powierzchnia trójkąta *AEF* stanowi koszt jałowy (zbędny) podatku, zwany także kosztem martwym (*deadweight loss*), czyli stratę społeczną spowodowaną nałożeniem podatku.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Jak zmienia się koszt jałowy podatku, gdy rząd podnosi stawkę podatkową? Załóżmy, że stawka podatkowa ulega podwojeniu, z początkowej stawki t do stawki $2t$. Krzywa podaży przesuwa się – o wielkość t – do położenia S_3, a producenci znowu chcą podnieść cenę o wielkość dodatkowego podatku. Przy tym poziomie cen popyt spada jednak jeszcze bardziej, co zmusza producentów do obniżenia ceny. Ostatecznie nowym punktem równowagi staje się C, nowa cena, którą płacą konsumenci wynosi P_B, a cena jaką otrzymują producenci wynosi P_G. Popyt i podaż spadają do poziomu Q_C. Koszt jałowy pozostaje teraz równy powierzchni trójkąta *ACG* i – jak łatwo policzyć – jest cztery razy większy od pola trójkąta *AEF*. Można udowodnić, że koszt jałowy rośnie proporcjonalnie do kwadratu wzrostu stawki podatkowej[9]. Wynika z tego praktyczny wniosek odnośnie do prowadzenia polityki podatkowej. Ponieważ koszt jałowy rośnie bardziej niż proporcjonalnie na skutek wzrostu podatków, zdecydowanie opłaca się zastosować więcej małych podatków niż mniej dużych.

Zmiany wielkości kosztu jałowego w odpowiedzi na zmiany stawki podatku pokazują też, że *krańcowy* koszt jałowy (*marginal dead-weight loss*) rośnie w miarę wzrostu podatku. Krańcowy koszt jałowy oznacza *wzrost* kosztu jałowego na *jednostkę wzrostu* podatku (odległość między krzywą S_1 a krzywą S_2). Rośnie on wraz ze stawką podatku, ponieważ straty społeczne są coraz większe (odcinek *CG* jest dłuższy niż odcinek *EF*). Ogólnie rzecz biorąc, im bliżej początku układu znajdują się nowe punkty Q w stosunku do początkowej wartości równowagi Q_1, tym większą mają wartość społeczną i tym większy jest koszt jałowy podatku.

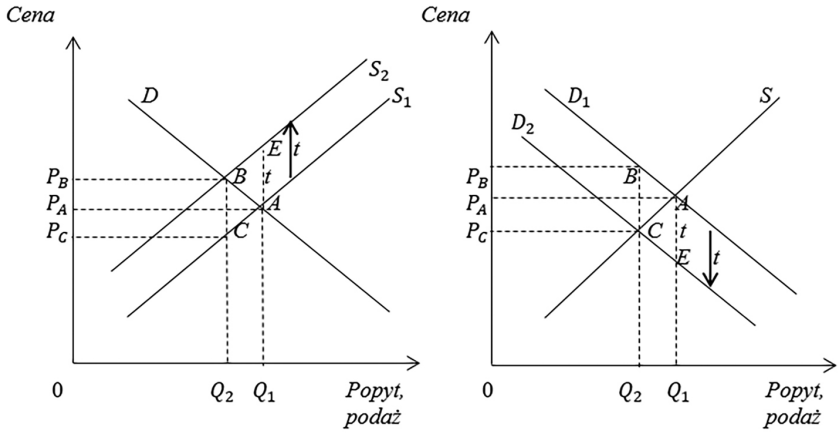
Rysunek 4.2 pokazuje, że podatek wpływa na sytuację ekonomiczną zarówno konsumentów, jak i producentów. Czy wobec tego jest istotne, na kogo władza gospodarcza nałoży podatek – na konsumentów czy na producentów? Rysunek 4.3 ponownie ilustruje wpływ podatku na podaż, popyt i cenę, przy czym lewa strona rysunku (stanowiąca replikę rysunku 4.2) przedstawia sytuację, w której podatek obciąża producentów, a prawa strona przedstawia sytuację, gdy podatek obciąża konsumentów. Można łatwo ustalić, co decyduje o tym, kto poniesie większy ciężar z tytułu podatku – producenci czy konsumenci.

---

[9] Zauważmy, że trójkąty *AEH* i *ACG* są podobne, ponieważ linie *S* są do siebie równoległe, podobnie jak odcinki *EF* i *CG* względem siebie. Pole trójkąta *AEH* wynosi $y = (th)/2$, gdzie t jest stawką podatku, a h oznacza wysokość trójkąta. Podwojenie stawki podatkowej sprawia, że koszt jałowy staje się równy polu trójkąta *ACG*. Pole trójkąta *ACG* wynosi $z = (2t \cdot 2h)/2 = (4th)/2 = 2^2y = 4y$. Ogólnie można wykazać, że wzrost stawki podatku o x razy oznacza wzrost kosztu jałowego o $(xt \cdot xh)/2 = x^2\left(\frac{th}{2}\right)$, czyli o x^2 razy.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 4.3. Wpływ nałożenia podatku na producenta i na konsumenta



Źródło: opracowanie własne.

Lewa strona rysunku 4.3 pokazuje, że w punkcie wyjścia cena P_A zapewnia równowagę między podażą a popytem (punkt A). Nałożenie podatku *t* na producentów przesuwa krzywą podaży w górę o odcinek *t*, do położenia S_2. Producenci początkowo chcą nadal otrzymywać cenę P_A, powiększoną jednak o podatek *t* (punkt *E*). lecz popyt przy tej cenie maleje i producenci są zmuszeni obniżyć cenę do poziomu P_B. Nowym punktem równowagi rynkowej jest *B*, w którym konsumenci płacą cenę P_B, a producenci otrzymują, po zapłaceniu podatku, cenę P_C. Jak widać, ciężar podatku jest ponoszony zarówno przez konsumentów (ponieważ $P_B > P_A$), jak i przez producentów (ponieważ $P_C < P_A$).

Spójrzmy teraz na prawą stronę rysunku 4.3. Nałożenie podatku na konsumentów przesuwa krzywa popytu w dół do położenia D_2. Konsumenci chcą nadal płacić cenę P_A, w którą będzie już wliczony podatek. To oznacza jednak, że producenci otrzymają tylko cenę P_C, co przełoży się na spadek podaży i konsumenci będą zmuszeni zgodzić się na wyższą cenę. Nowym punktem równowagi rynkowej jest *B*, gdzie konsumenci płacą cenę P_B, z której odprowadzają podatek, podczas gdy producenci otrzymują cenę P_C. Jak widać, ciężar podatku jest ponoszony zarówno przez konsumentów (ponieważ $P_B > P_A$), jak i przez producentów (ponieważ $P_C < P_A$).

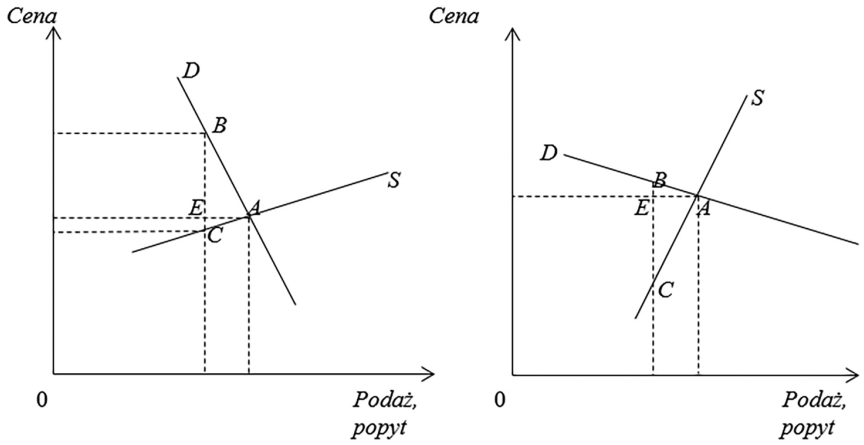
Tak więc w obu przypadkach koszty wprowadzenia podatku rozkładają się między producentów i konsumentów bez względu na to, na kogo formalnie nakładany jest podatek. Wynika to stąd, że klin podatkowy wpływa na ceny, a zmiana cen

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wywołuje zmianę zarówno popytu, jak i podaży. Czy można jednak powiedzieć coś na temat tego, jak rozłożą się proporcje wpływu podatku na konsumentów i producentów? Owszem, jesteśmy w stanie wykazać, że to, kto i w jakim stopniu ponosi ciężar nałożenia podatku, zależy wyłącznie od stosunku elastyczności podaży oraz elastyczności popytu.

Warto w tym celu przyjrzeć się rysunkowi 4.4. Lewa strona odzwierciedla sytuację, w której krzywa podaży jest płaska, a krzywa popytu pozostaje stroma, co oznacza, że elastyczność podaży jest większa od elastyczności popytu, $\eta_S > \eta_D$. Nałożenie podatku w wysokości *BC* przekłada się na wzrost ceny dla konsumentów o wielkość *BE* oraz spadek ceny dla producentów o wielkość *EC*. Wyraźnie widać, że ze względu na to, iż elastyczność popytu jest mniejsza niż elastyczność podaży (krzywa popytu pozostaje bardziej stroma niż krzywa podaży), konsumenci ponoszą znacznie większy ciężar z tytułu nałożenia podatku niż producenci, ponieważ wzrost ceny, jaką przychodzi zapłacić konsumentom, jest większy niż spadek ceny otrzymywanej przez producentów, czyli *BE* > *EC*.

Rysunek 4.4. Kto ponosi koszty nałożenia podatku?



Źródło: opracowanie własne.

Prawa strona rysunku 4.4 ilustruje sytuację odwrotną, w której to krzywa podaży jest stroma, natomiast krzywa popytu jest płaska, co oznacza, że elastyczność podaży jest mniejsza od elastyczności popytu, $\eta_S < \eta_D$. Nałożenie podatku w wysokości *BC* wywołuje wzrost ceny dla konsumentów o wielkość *BE* oraz spadek ceny dla producentów o wielkość *EC*. Wyraźnie widać, że ze względu na to, iż

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

elastyczność popytu jest większa niż elastyczność podaży (bo krzywa podaży pozostaje bardziej stroma niż krzywa popytu), producenci ponoszą znacznie większy ciężar z tytułu nałożenia podatku niż konsumenci, ponieważ spadek ceny otrzymywanej przez producentów jest większy niż wzrost ceny, jaką przychodzi zapłacić konsumentom, a więc *BE* < *EC*.

Można obecnie podsumować analizę wpływu podatku na podaż, popyt i cenę. Niezależnie od tego, czy podatek jest nałożony na producenta, czy na konsumenta, powoduje on zawsze spadek produkcji oraz popytu i wzrost ceny rynkowej, a także powstanie kosztu jałowego, który oznacza stratę dobrobytu społecznego. W rzeczywistości na nałożenie podatku reagują zarówno konsumenci, jak i producenci, a więc i krzywa *S*, i krzywa *D* ulegają przesunięciu w lewo. Z kolei proporcje, w jakich producenci i konsumenci ponoszą ciężar podatku, zależą od stosunku elastyczności podaży i popytu: gdy $\eta_S > \eta_D$, więcej tracą konsumenci, gdy zaś $\eta_S < \eta_D$, wówczas więcej tracą producenci. Tak więc nie ma znaczenia, na „kogo" nakłada się podatek, liczy się jedynie stosunek elastyczności podaży i popytu.

Obecnie zbadamy wpływ podatku pośredniego na wielkość i strukturę konsumpcji. Załóżmy, że mamy do dyspozycji dwa dobra konsumpcyjne, *x* (na przykład wino) oraz *y* (na przykład piwo), a konsumenci maksymalizują funkcję użyteczności o postaci ogólnej:

$$U = f(x,y). \tag{4.1}$$

Przy czym *U'* > 0, a *U"* < 0, co oznacza, że krańcowa użyteczność konsumpcji poszczególnych dóbr maleje w miarę wzrostu ich konsumpcji. Przyjmijmy też, że wydatki konsumentów spełniają ograniczenie budżetowe o postaci:

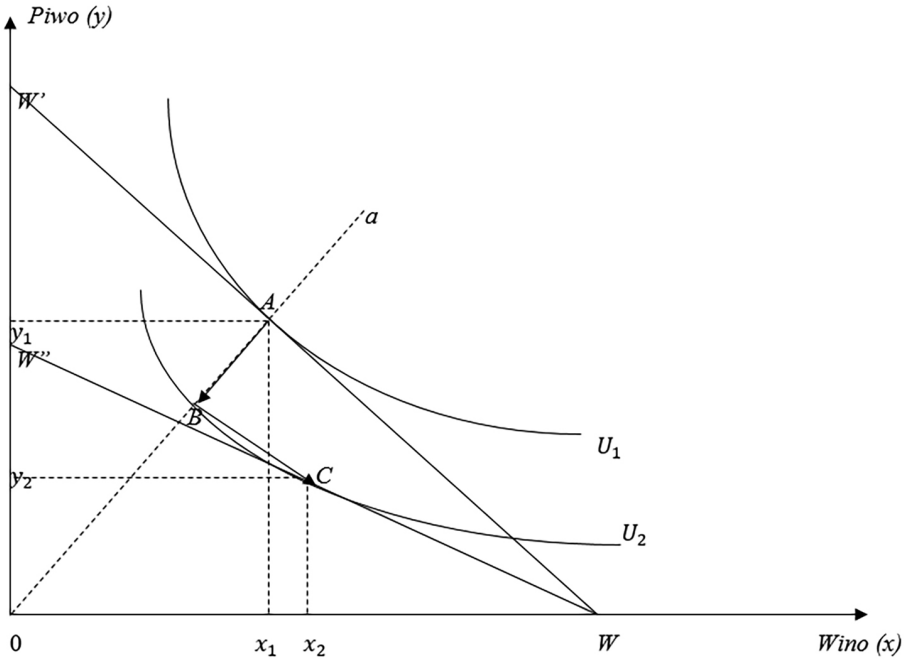
$$p_x x + p_y y = M\ , \tag{4.2}$$

gdzie p_x, p_y są cenami jednostkowymi dóbr *x* i *y*, a *M* jest pieniężnym dochodem konsumenta. Rysunek 4.5 pokazuje, że konsumenci maksymalizują użyteczność w punkcie *A* o współrzędnych x_1, y_1, gdzie linia budżetu *WW'* jest styczna do najdalej położonej – czyli najkorzystniejszej – krzywej użyteczności U_1. Poziomy użyteczności reprezentowane przez krzywe *U* znajdujące się poniżej U_1 są z definicji niższe, a poziomy użyteczności reprezentowane przez krzywe *U* położone powyżej U_1 są niedopuszczalne, ponieważ leżą powyżej linii budżetu. Linia budżetu *WW'* ma nachylenie równe stosunkowi cen $\frac{p_x}{p_y}$. W punkcie *A* nachylenie krzywej U_1

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jest dokładnie równe nachyleniu linii budżetu. W ramce 4.4 przeprowadzono formalny dowód na to, że funkcja U osiąga maksimum w punkcie styczności A.

Rysunek 4.5. Wpływ podatku pośredniego na wielkość i strukturę konsumpcji: przypadek dwóch dóbr konsumpcyjnych



Źródło: opracowanie własne.

Ramka 4.4. Maksimum funkcji użyteczności

Załóżmy, że mamy do dyspozycji dwa dobra, x i y, oraz że funkcja użyteczności konsumpcji ma postać opisaną równaniem (4.3), jej graficzną ilustracją jest krzywa U_1 zlokalizowana na rysunku 4.5, a ograniczenie budżetowe ma postać równania (4.4). Otrzymujemy więc:

$$U = f(x,y), \tag{4.3}$$

$$p_x x + p_y y = M. \tag{4.4}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Nachylenie linii budżetu jest równe stosunkowi cen $\frac{p_x}{p_y}$, podczas gdy nachylenie krzywej U zmienia się i w każdym jej punkcie pozostaje równe krańcowej stopie substytucji (*marginal rate of substitution – MRS*), która wskazuje, z ilu jednostek dobra y konsumenci są gotowi zrezygnować w zamian za jednostkę dobra x. W miarę tego, jak przesuwamy się w lewo wzdłuż krzywej U_1, rośnie udział dobra y w danej konsumpcji, a spada udział dobra x. W efekcie dobro x staje się relatywnie coraz bardziej pożądane i wartość *MRS* rośnie.

Z definicji mamy więc $MRS = -\frac{dy}{dx}$. Aby obliczyć wartość *MRS* dla danego punktu na krzywej U_1, różniczkujemy funkcję (4.3):

$$dU = \frac{\partial U}{\partial x}dx + \frac{\partial U}{\partial y}dy. \tag{4.5}$$

Przyjmując, że $dU = 0$ (ponieważ wartość U wzdłuż krzywej U_1 pozostaje taka sama), otrzymujemy na podstawie równania (4.5):

$$\frac{\frac{\partial U}{\partial x}}{\frac{\partial U}{\partial y}} = -\frac{dy}{dx} = MRS. \tag{4.6}$$

Aby wykazać, że funkcja U osiąga maksimum w punkcie, w którym nachylenie linii budżetu jest równe nachyleniu funkcji U, budujemy funkcje Lagrange'a:

$$L = U(x, y) + \lambda\left(M - p_x x - p_y y\right). \tag{4.7}$$

Warunki pierwszego rzędu dla $U = max$ są następujące:

$$\left.\begin{aligned} \frac{\partial L}{\partial x} &= \frac{\partial U}{\partial x} - \lambda p_x = 0 \\ \frac{\partial L}{\partial y} &= \frac{\partial U}{\partial y} - \lambda p_y = 0 \\ \frac{\partial L}{\partial \lambda} &= M - p_x x - p_y y = 0. \end{aligned}\right\} \tag{4.8}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozwiązując dwa pierwsze warunki dla λ, otrzymujemy:

$$\lambda = \frac{\frac{\partial U}{\partial x}}{p_x} = \frac{\frac{\partial U}{\partial y}}{p_y}. \tag{4.9}$$

Przekształcając i wykorzystując równanie (4.6), mamy zaś:

$$\frac{\frac{\partial U}{\partial x}}{\frac{\partial U}{\partial y}} = MRS = \frac{p_x}{p_y}. \tag{4.10}$$

tak więc warunkiem maksimum funkcji *U* jest równość krańcowej stopy substytucji i stosunku cen. Jedynym punktem na krzywej U_1, który spełnia ten warunek, jest punkt *A*, będący punktem styczności linii budżetu *WW'* i najdalej położonej krzywej użyteczności U_1.

Załóżmy teraz, że władza gospodarcza nakłada podatek na dobro *y* (piwo) w wysokości 50%. Oznacza to, że gdyby konsumenci chcieli wszystkie posiadane środki przeznaczyć na zakup piwa, w ramach niezmienionych możliwości budżetowych mogą obecnie kupić o połowę mniej piwa niż poprzednio. Ich możliwości zakupu wina pozostały bez zmian. Wskutek nałożenia podatku linia budżetu przesuwa się w dół do położenia *WW"*. Nowym punktem równowagi staje się punkt *C*, w którym wartość użyteczności w ramach nowego ograniczenia budżetowego osiąga maksimum.

Zauważmy, że zastosowanie podatku wywołuje podwójny efekt. Po pierwsze, konsumenci przenoszą się na nową, gorszą krzywą użyteczności U_2. Wynika to stąd, że podatek odbiera im część siły nabywczej, a więc efektywnie zmniejsza ich dochody i zmusza do obniżenia poziomu konsumpcji. Dlatego też ten efekt związany z nałożeniem podatku nazywamy efektem dochodowym (*income effect*). Po drugie, konsumenci przesuwają się *wzdłuż* nowej krzywej U_2 w prawo. Podatek nałożony na piwo powoduje wzrost jego ceny w relacji do ceny wina (za równowartość tego samego dochodu mogę kupić o połowę mniej piwa), co zmniejsza atrakcyjność konsumpcji piwa w relacji do wina. W rezultacie zmienia się struktura konsumpcji – konsumenci spożywają mniej piwa, a za to relatywnie więcej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wina. Ponieważ piwo zastępowane jest winem, ten efekt związany z nałożeniem podatku nazywamy efektem substytucji (*substitution effect*).

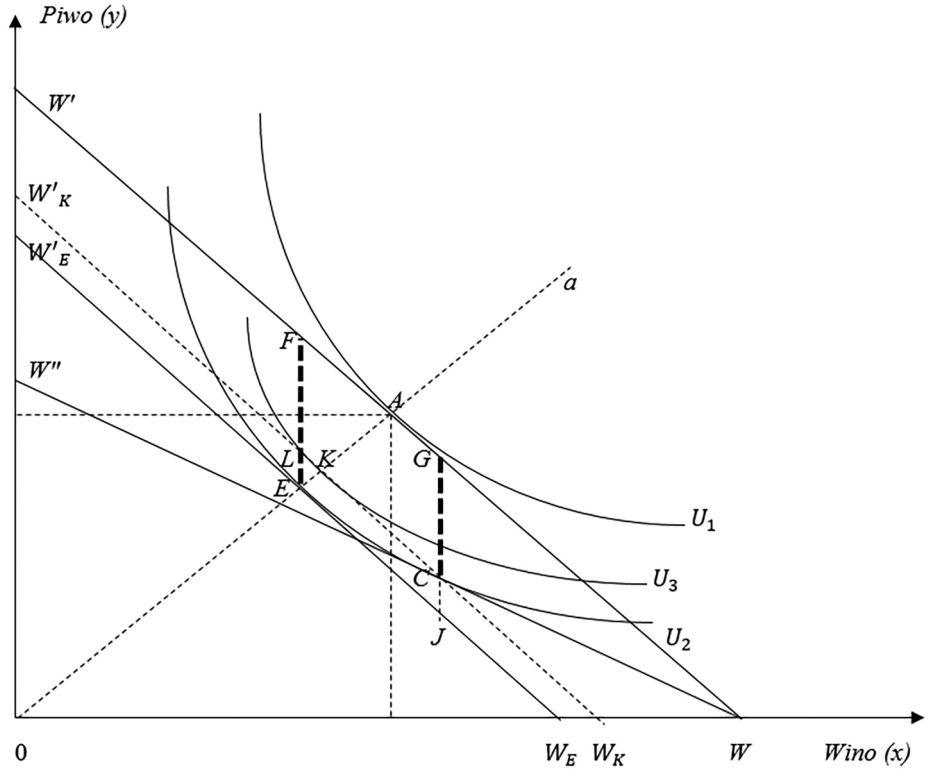
Rysunek 4.6 pokazuje, że nałożenie podatku na jedno z dóbr powoduje ogólny spadek konsumpcji w wyniku efektu dochodowego (odcinek *AB*) oraz pogorszenie jej struktury pod wpływem efektu substytucji (odcinek *BC*). Efekt substytucji zawsze oznacza odejście od optymalnej struktury konsumpcji, którą ilustruje linia *a* (łącząca punkty na kolejnych krzywych użyteczności o takim samym nachyleniu, czyli stałej relacji *y/x*). To odchylenie od optymalnej struktury nazywane jest kosztem „jałowym" podatku (*deadweight loss*). Jego wielkość nigdy nie jest znana *ex ante*, ponieważ zależy od stawki podatku oraz od przebiegu krzywych użyteczności. Czy można uniknąć efektu substytucji przy nakładaniu podatku? Owszem, jest to możliwe, gdy zastosuje się podatek ryczałtowy (kwotowy) lub nałoży się podatek na oba dobra w jednakowej wysokości. Rysunek 4.6 pozwala porównać działanie podatku pośredniego (wyrażonego w proporcji do ceny) i podatku ryczałtowego.

Gdy nie ma podatków, punktem równowagi – tak jak na rysunku 4.5 – jest *A*, a zatem punkt styczności linii budżetu *WW'* i najdalej położonej krzywej użyteczności U_1. Nałożenie podatku pośredniego na piwo w wysokości 50% przesuwa linię budżetu w dół do punktu *WW"*, a nowym punktem równowagi staje się *C*. Skutki są następujące. Po pierwsze, następuje spadek użyteczności z U_1 do U_2 (z punktu *A* do punktu *E*) – jest to efekt dochodowy. Po drugie, pojawia się zmiana struktury konsumpcji w wyniku przesunięcia wzdłuż krzywej U_2 z punktu *E* do *C* – jest to efekt substytucji. Odcinek *GC* mierzy wielkość dochodów budżetu z tytułu podatku pośredniego na piwo.

Teraz z kolei wyobraźmy sobie, że władza gospodarcza nakłada na konsumentów podatek ryczałtowy w wysokości równej odcinkowi *EF*, czyli równej odległości pomiędzy krzywą U_1 a krzywą U_2. Linia budżetu przesuwa się w dół do położenia $W_E W'_E$, a nowym punktem równowagi staje się *E*. Podatek ryczałtowy pozostawia konsumentów na tej samej krzywej U_2 co podatek pośredni na piwo, lecz pozwala uzyskać większe dochody budżetowe, ponieważ odcinek *EF* jest większy niż odcinek *GC* (różnica długości pozostaje równa odcinkowi *CJ*). Innymi słowy, podatek ryczałtowy jest bezwzględnie lepszym rozwiązaniem niż wcześniej omówiony podatek pośredni, ponieważ poprawia sytuację budżetu, nie pogarszając jednocześnie sytuacji konsumentów. Ten korzystny rezultat uzyskuje się dzięki eliminacji efektu substytucji – punkt równowagi *C* znajduje się na linii *a*, reprezentującej optymalną strukturę konsumpcji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 4.6. Porównanie podatku pośredniego i podatku ryczałtowego



Źródło: opracowanie własne.

Alternatywnie władza gospodarcza może nałożyć podatek ryczałtowy w wysokości równej *GC*, czyli niższy niż podatek ryczałtowy *FE*. Nową linią budżetu jest linia $W_K W'_K$, przechodząca przez punkt *C*, co oznacza, że dochody budżetu pozostają równe odcinkowi *GC* i są takie same jak w przypadku podatku pośredniego na piwo. Lecz tym razem konsumenci znajdą się na lepszej, wyżej położonej krzywej użyteczności U_3, a nowym punktem równowagi stanie się *K*. Również w tym przypadku podatek ryczałtowy zapewnia bezwzględnie lepsze rozwiązanie niż podatek pośredni na piwo, ponieważ poprawia sytuację konsumentów (użyteczność w punkcie *K* jest większa niż w punkcie *C*), pozostawiając jednocześnie dochody budżetu bez zmian. Podobnie jak poprzednio, wynik ten osiąga się skutkiem eliminacji efektu substytucji, ponieważ punkt konsumpcji *K* leży na linii optymalnej struktury konsumpcji *a*.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wreszcie też identyczny efekt ma zastosowanie podatku równomiernego, czyli obciążającego w jednakowym stopniu oba produkty. Wówczas linia budżetu przesuwa się w dół równolegle tak, jak w przypadku podatku ryczałtowego. Opodatkowanie równomierne jest neutralne w odniesieniu do struktury konsumpcji i dzięki temu nie pojawia się efekt substytucji. Z obserwacji tej wynika kolejny praktyczny wniosek odnośnie do polityki podatkowej. Ponieważ zróżnicowane stawki podatku prowadzą do powstania efektu substytucji, odchylającego strukturę konsumpcji od struktury optymalnej, należy stosować w miarę równomierne stawki podatkowe. Wówczas system podatkowy jest neutralny w stosunku do działalności gospodarczej, co oznacza, że nie wywołuje szkodliwych efektów substytucji.

4.5.4. Efektywność ekonomiczna: wpływ podatków na zatrudnienie i dochód

Dotychczas omawialiśmy działanie podatków pośrednich, obciążających towary w obrocie rynkowym. Obecnie przychodzi natomiast zająć się podatkami bezpośrednimi, obciążającymi dochód, oraz zbadać wpływ stosowania tych podatków na podaż pracy i poziom zatrudnienia. Rysunek 4.7 przedstawia zależność między podażą pracy, wysokością dochodów z pracy a poziomem konsumpcji. Podaż pracy L jest mierzona w jednostkach czasu (godziny, dni) i równa iloczynowi liczby zatrudnionych E oraz średniej liczby godzin pracy na jednego zatrudnionego n:

$$L = nE. \tag{4.11}$$

Jak widać, wielkość L może zmieniać się zarówno na skutek zmian poziomu zatrudnienia E, co określa się mianem ekstensywności pracy (*extensive margin*), jak i pod wpływem zmian liczby godzin pracy przypadających na zatrudnionego, co określa się mianem intensywności pracy (*intensive margin*).

Wypukłe krzywe użyteczności U ilustrują malejącą krańcową użyteczność konsumpcji uzyskanej dzięki dochodom z pracy – im wyższy poziom konsumpcji, tym większy musi być przyrost konsumpcji na jednostkę dodatkowej pracy. Innymi słowy, im wyższy jest poziom podaży pracy (im wyższe zatrudnienie), tym większa staje się krańcowa użyteczność czasu wolnego (*leisure*) i tym wyższy musi być przyrost konsumpcji, aby zrównoważyć utraconą jednostkę czasu wolnego.

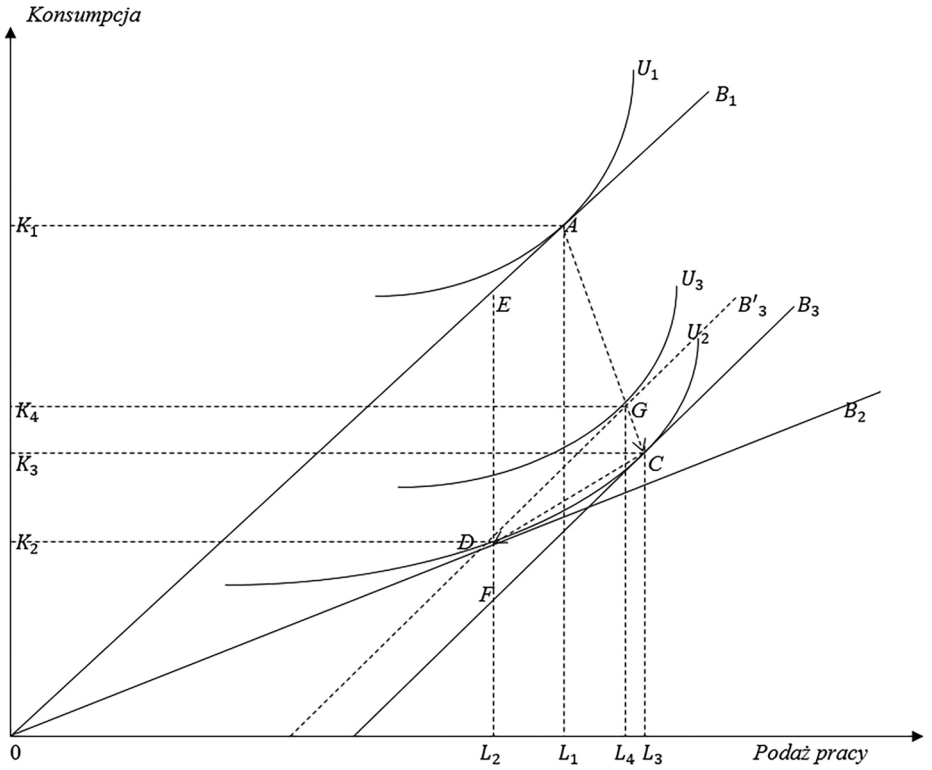
Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Związek między podażą pracy a konsumpcją K jest opisany równaniem budżetu:

$$K = wL, \tag{4.12}$$

gdzie w jest stawką wynagrodzenia za jednostkę czasu pracy. Jak widać, ogólny poziom konsumpcji może kształtować się wskutek zmian zarówno podaży pracy, jak i poziomu jednostkowych wynagrodzeń. Przy założeniu, że wynagrodzenia są stałe, zależność (4.12) ma charakter liniowy, a jej graficznym przedstawieniem na rysunku 4.7 jest linia budżetu B_1 o nachyleniu równym stawce płacy w.

Rysunek 4.7. Podatek od wynagrodzeń a konsumpcja i podaż pracy



Źródło: opracowanie własne.

Linia budżetu B_1 jest linią prostą i odzwierciedla związek między czasem pracy a konsumpcją, gdy wynagrodzenie za każdą jednostkę pracy jest stałe. Zakładamy też początkowo, że wynagrodzenia nie są opodatkowane.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Równowaga na rynku pracy ustali się wówczas w punkcie styczności linii budżetu B_1 i najdalej położonej krzywej użyteczności konsumpcji U_1 – jest to punkt *A*. W tym miejscu następuje zrównanie wynagrodzenia za jednostkę pracy i krańcowej użyteczności konsumpcji. Podaż pracy wynosi L_1, a poziom ogólnej konsumpcji K_1. Zgodnie z kształtem krzywej użyteczności, pracownikom opłaca się zwiększać podaż pracy aż do poziomu L_1, czyli do momentu, w którym otrzymywane wynagrodzenie jednostkowe *w* zrówna się z krańcową użytecznością czasu wolnego, mierzoną nachyleniem funkcji *U*. Zwiększanie podaży pracy powyżej poziomu L_1 staje się już dla pracowników nieopłacalne, ponieważ otrzymywane wynagrodzenie jest wtedy mniejsze niż krańcowa użyteczność czasu wolnego.

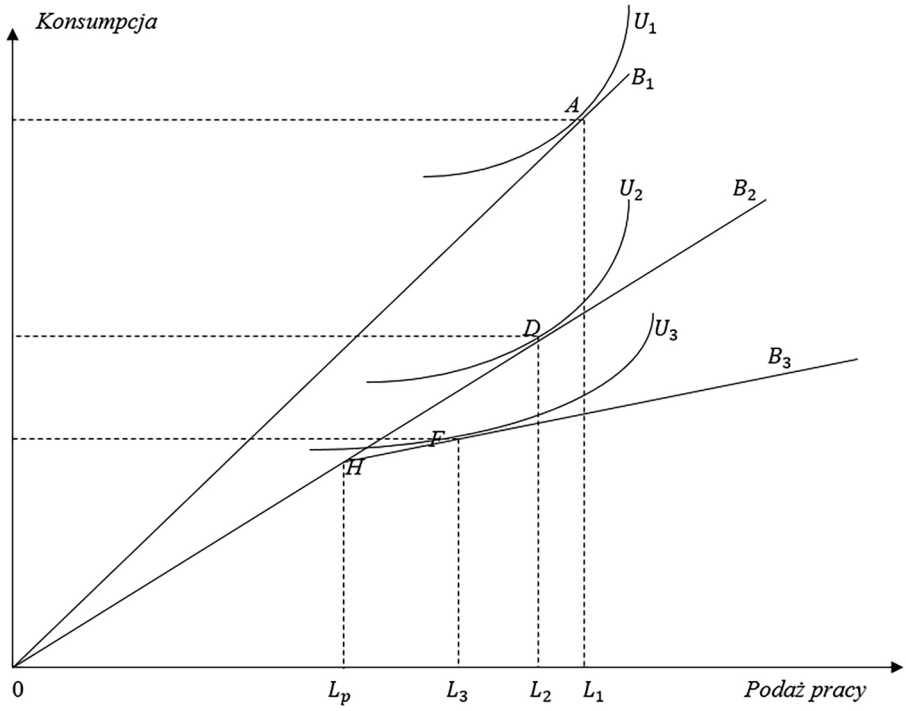
Załóżmy teraz, że władza gospodarcza nakłada na wynagrodzenia podatek proporcjonalny w wysokości 50%. Oznacza to, że siła nabywcza wynagrodzeń maleje o połowę i krzywa budżetu przesuwa się w dół, do położenia B_2. W tej sytuacji podaż pracy spada do poziomu L_2, konsumpcja maleje i osiąga wartość K_2, a nowym punktem równowagi staje się *D*, czyli punkt styczności nowej linii budżetu B_2 i najdalej położonej krzywej użyteczności (U_2). Spadek poziomu konsumpcji jest spowodowany głównie efektem dochodowym podatku (odcinek *AC*), ale mamy wówczas również do czynienia z efektem substytucji (odcinek *CD*). Ten ostatni skłania pracowników do ograniczenia podaży pracy na rzecz czasu wolnego, ponieważ krańcowa użyteczność czasu wolnego na odcinku *CD* krzywej U_2 jest wyższa od wynagrodzenia netto.

Zmiana podatku proporcjonalnego na podatek ryczałtowy w wysokości *EF* pozwala wyeliminować skutki efektu substytucji. Linia budżetu przesuwa sią równolegle do położenia B_3, podaż pracy wzrasta do poziomu L_3, a konsumpcja rośnie do wysokości punktu K_3. Łatwo zauważyć, że podatek ryczałtowy jest bezwzględnie lepszym rozwiązaniem niż podatek proporcjonalny, ponieważ zapewnia większe dochody do budżetu (odcinek *EF* jest dłuższy niż odcinek *ED*), nie pogarszając jednocześnie sytuacji konsumentów (konsumenci pozostają na niezmienionej krzywej użyteczności U_2). Jak pamiętamy z rozdziału pierwszego, uzyskana w ten sposób poprawa ogólnego dobrobytu nazywa się korzyścią w sensie Pareto. Alternatywnie można obniżyć podatek do poziomu *ED*, wówczas linia budżetu przesuwa się do B'_3, podaż pracy rośnie do L_4, konsumpcja wzrasta zaś do poziomu K_4. Nowym punktem równowagi staje się *G*, położony na korzystniejszej, wyżej zlokalizowanej krzywej użyteczności U_3. Również w tym przypadku podatek ryczałtowy okazuje się lepszym rozwiązaniem niż podatek proporcjonalny, ponieważ powoduje wzrost podaży pracy oraz wzrost użyteczności konsumpcji, nie zmniejszając przy tym dochodów budżetu. To drugie rozwiązanie jest także korzystne w sensie Pareto.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Na koniec zbadamy wpływ podatku progresywnego na podaż pracy i konsumpcję. Rysunek 4.8 jest kopią rysunku 4.7, przy czym główna różnica polega na tym, że do analizy wprowadza się progresję podatkową. Załóżmy, że dochody z podaży pracy do poziomu L_p są obciążone dotychczasową stawką podatku 50%, a powyżej tego poziomu stawka rośnie do 70%. Ograniczenie budżetowe reprezentuje obecnie linia łamana 0HB_3. Punktem równowagi staje się w tej sytuacji punkt *F*, gdzie nachylenie linii B_3 jest równe nachyleniu krzywej użyteczności U_3 w tym punkcie. Jak łatwo zauważyć, podatek progresywny w jeszcze większym stopniu ogranicza podaż pracy i konsumpcję niż podatek proporcjonalny – podaż pracy spada z poziomu L_2 do L_3, konsumpcja z K_2 do K_3, a punkt równowagi znajduje się na niżej położonej (gorszej) krzywej użyteczności U_3[10].

Rysunek 4.8. Podatek proporcjonalny a podatek progresywny



Źródło: opracowanie własne.

[10] W praktyce podatek progresywny wywiera podwójny efekt na ograniczenie podaży pracy. Po pierwsze, spada wówczas liczba godzin w przeliczeniu na jednego zatrudnionego, czyli zmniejsza się *intensywność* pracy, po drugie zaś zmniejsza się liczba zatrudnionych, a więc spada *ekstensywność* pracy.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Podsumowując oddziaływanie różnego rodzaju podatków na podaż pracy, dochód, konsumpcję oraz użyteczność, można stwierdzić, że najmniej szkodliwe skutki – w postaci ograniczenia pracy, dochodu i konsumpcji – niesie podatek ryczałtowy, bardziej dotkliwy jest podatek proporcjonalny (liniowy), a najbardziej niekorzystny okazuje się podatek progresywny. W świetle tych obserwacji warto więc zadać pytanie: dlaczego w praktyce tak rzadko stosuje się podatki ryczałtowe? Odpowiedź jest prosta. Gdyby wszyscy ludzie byli identyczni, nie trzeba byłoby różnicować podatków, i wówczas najlepszym podatkiem byłby podatek ryczałtowy. W rzeczywistości jednak ludzie nie są jednakowi, a ich pozycja w społeczeństwie jest zróżnicowana pod względem położenia materialnego, sytuacji rodzinnej, kosztów utrzymania, stanu zdrowia czy możliwości uzyskiwania dochodu. W związku z tym współczesne państwa z reguły dążą do uwzględniania tych różnic w ramach systemu podatkowego.

4.5.5. Efektywność ekonomiczna: podatek optymalny według Ramseya

Opodatkowanie optymalne to takie opodatkowanie, które jest efektywne w sensie Pareto, co oznacza, że nie istnieje inny system podatkowy, który zapewniłby poprawę sytuacji części osób bez jednoczesnego pogorszenia sytuacji pozostałych członków społeczeństwa. Innymi słowy, jest to taki system podatkowy, który minimalizuje zniekształcenia rynkowe wywołane funkcjonowaniem podatku (Stiglitz 2004, Mankiw i in. 2009). Przeprowadzona wyżej analiza wpływu podatku dochodowego na podaż pracy, dochód i konsumpcję pokazuje, że optymalnym podatkiem dochodowym jest podatek ryczałtowy. Jednocześnie wskazano jednak na ograniczenia, które sprawiają, że podatek ryczałtowy znajduje bardzo ograniczone zastosowanie w praktyce.

W ujęciu bardziej ogólnym optymalny podatek to taki, który maksymalizuje jakąś funkcję dobrobytu społecznego. Takich funkcji może być wiele i dlatego może też istnieć wiele optymalnych systemów podatkowych. Przykładowo, optymalny podatek będzie miał inną postać w przypadku funkcji użyteczności z dominującą rolą kryterium efektywności ekonomicznej, inną natomiast odnośnie funkcji użyteczności z dominującą rolą kryterium sprawiedliwości i równości.

Załóżmy, że władza gospodarcza musi zgromadzić określoną kwotę dochodów podatkowych, lecz ma do dyspozycji wyłącznie podatki pośrednie. Wybitny ekonomista brytyjski Frank Ramsey zasugerował, że takie podatki powinny być

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

nakładane na poszczególne towary w odwrotnej proporcji do elastyczności popytu na te towary. Innymi słowy, towary, na które popyt jest nieelastyczny, powinny być opodatkowane wyższymi stawkami, i odwrotnie. Ramsey wykazał, że optymalne opodatkowanie pośrednie wszystkich towarów powinno spełniać następujący warunek (Ramsey 1927):

$$\frac{MDW_i}{MR_i} = \lambda, \tag{4.13}$$

gdzie $MDW_i = \frac{\partial(DW_i)}{\partial(\tau_i)}$ jest krańcowym kosztem jałowym (*marginal dead-weight loss*) podatku τ_i na towarze *i*, $MR_i = \frac{\partial(R_i)}{\partial(\tau_i)}$ stanowi krańcowy dochód uzyskany z podniesienia podatku τ_i na towarze *i*, a λ jest swoistą „wartością" dodatkowego dochodu znajdującego się w dyspozycji rządu w stosunku do wartości uzyskanej z możliwie najlepszego wykorzystania tego dochodu przez sektor prywatny. Ogólnie relacja $\frac{MDW_i}{MR_i}$ mierzy krańcowy koszt jałowy na jednostkę dodatkowego dochodu podatkowego. Na rysunku 4.2 wykazano, że krańcowy koszt jałowy rośnie wraz ze stawką podatku, ponieważ straty społeczne są coraz większe (odcinek *CG* jest większy niż odcinek *EF*), oraz ustalono, że koszt krańcowy rośnie wraz z kwadratem stawki podatku. Jednocześnie krańcowy dochód z podniesienia podatku spada, ponieważ równolegle maleje poziom sprzedaży. Oznacza to, że relacja $\frac{MDW_i}{MR_i}$ rośnie wraz ze wzrostem stawki podatku. Gdy λ przyjmuje wysokie wartości, oznacza to, że dodatkowe dochody w rękach rządu mają dużą wartość w relacji do dochodów w rękach prywatnych, a więc relacja *MDW/MR* może być duża. Na przykład, przy ogólnie niskim poziomie wydatków publicznych na naukę, edukację czy ochronę zdrowia, wartość λ jest z reguły wysoka, co oznacza, że warto „odbierać" dochód od sektora prywatnego na rzecz sektora publicznego. Gdy z kolei wartość λ jest niska, dodatkowe dochody w rękach rządu mają niewielką wartość w relacji do dochodów w rękach prywatnych, a więc relacja *MDW/MR* powinna być niska.

Wzór (4.13) nazywa się regułą Ramseya. Wyznacza ona optymalny poziom opodatkowania konsumpcji dobra *i* przez sektor prywatny na rzecz sektora

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

publicznego. Gdy podaż jest doskonale elastyczna, reguła Ramseya implikuje, że optymalny podatek ma postać:

$$\tau_i^* = -\frac{\lambda}{\eta_D^i}, \tag{4.14}$$

gdzie t^* oznacza optymalną stawkę podatku, a η_D^i odpowiada elastyczności popytu na dobro i. Równanie (4.14) pokazuje, po pierwsze, że stawki podatku powinny być tym wyższe, im większy jest współczynnik λ, oraz tym niższe, im większa jest elastyczność popytu na dane dobro. Formalne wyprowadzenie reguły Ramseya zostało przedstawione w ramce 4.5.

Ramka 4.5. Podatki optymalne Ramseya*

Frank Ramsey poszukiwał odpowiedzi na pytanie o to, jaki system podatkowy pozwala minimalizować skutki zniekształceń rynkowych powodowanych przez podatki. Mając świadomość, że powszechne wykorzystanie podatku ryczałtowego jest mało realne, Ramsey starał się określić optymalne wartości stawek podatków pośrednich i bezpośrednich, maksymalizując funkcję użyteczności w ramach danych ograniczeń budżetowych.

Jak już wiadomo, szkodliwy wpływ podatku na efektywność ekonomiczną wynika przede wszystkim z powstawania efektu substytucji. Efekt dochodowy przesuwa jedynie zasoby z rąk prywatnych do budżetu i zasadniczo nie ma wpływu na poziom dobrobytu ogólnospołecznego[11]. Dlatego też Ramsey przyjął, że optymalne opodatkowanie towarów polega na ustaleniu takich stawek podatku, które zapewniają minimalizację efektu substytucji, czyli całkowitego kosztu jałowego podatku (DW_T), pod warunkiem uzyskania dochodów w wysokości co najmniej $\underline{R}$. Otrzymujemy więc wówczas:

$$DW_T = \sum_{i=1}^{i=n} DW_i = minimum. \tag{4.15}$$

Przy ograniczeniu:

$$R_1 + R_2 + \ldots R_n = \underline{R}, \tag{4.16}$$

[11] Przy założeniu, że krańcowa użyteczność dochodu z produkcji jest stała, a produktywność sektora publicznego jest taka sama jak w przypadku sektora prywatnego, co oczywiście nie musi być prawdą.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gdzie DW_i jest kosztem jałowym związanym z opodatkowaniem towaru i, a R_i oznacza dochód podatkowy z tytułu opodatkowania towaru i, przy czym $i = 1,2, \ldots n$. W celu znalezienia warunkowego minimum funkcji DW_T budujemy funkcję Lagrange'a o następującej postaci:

$$L = \left[DW_1 + DW_2 + \ldots + DW_n\right] - \lambda[R_1 + R_2 + \ldots R_n - \underline{R}], \tag{4.17}$$

gdzie λ jest mnożnikiem Lagrange'a, czyli ceną dualną ograniczenia dochodu (λ mierzy wpływ zmiany ograniczenia $\underline{R}$ o jednostkę na wartość funkcji DW_T). Różniczkując względem stawki podatku τ i przyrównując do zera, otrzymujemy warunki pierwszego rzędu:

$$\left.\begin{aligned} &\frac{\partial DW_1}{\partial \tau_1} = \lambda \frac{\partial R_1}{\partial \tau_1} \\ &\frac{\partial DW_2}{\partial \tau_2} = \lambda \frac{\partial R_1}{\partial \tau_2} \\ &\ldots \\ &\frac{\partial DW_n}{\partial \tau_n} = \lambda \frac{\partial R_n}{\partial \tau_n}, \end{aligned}\right\} \tag{4.18}$$

gdzie τ_i jest stawką podatku pośredniego nałożonego na towar i. W ten sposób uzyskuje się regułę Ramseya:

$$\frac{\partial(DW_i)}{\partial(\tau_i)} \Big/ \frac{\partial(R_i)}{\partial(\tau_i)} = MDW_i \big/ MR_i = \lambda \text{, dla wszystkich } i = 1,2, \ldots, n. \tag{4.19}$$

Równanie (4.19) jest identyczne jak równanie (4.13) i pokazuje, że funkcja DW_T osiąga minimum, gdy relacja krańcowego kosztu jałowego podatku i krańcowego dochodu z podatku wyrównuje się w przypadku wszystkich towarów ($i = 1,2, \ldots, n$).

Powracając do rysunku 4.2, należy przypomnieć, że wysokość kosztu jałowego odnośnie do podatku $\tau = EF$ mierzona jest powierzchnią trójkąta AEF. Można zatem zapisać, że:

$$DW = -\frac{\Delta Q \tau}{2}, \tag{4.20}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gdzie ΔQ stanowi wielkość (ujemną) spadku podaży w wyniku nałożenia podatku. Na podstawie wzorów na elastyczność podaży i popytu uzyskujemy:

$$\eta_S = \frac{\Delta Q}{\Delta P}\frac{P}{Q} \text{ oraz } \eta_D = \frac{\Delta Q}{(\Delta P + \tau)}\frac{P}{Q}, \tag{4.21}$$

gdzie $(\Delta P + \tau)$ jest zmianą ceny dla konsumenta. Przekształcając (4.21), otrzymujemy kolejno:

$$\Delta Q / Q = \eta_S \frac{\Delta P}{P} = \eta_D \frac{(\Delta P + \tau)}{P} \tag{4.22}$$

oraz

$$\Delta P = \frac{\eta_D}{\eta_S - \eta_D}\tau. \tag{4.23}$$

Podstawiając zaś (4.23) do (4.22), uzyskujemy:

$$\Delta Q = \frac{\eta_S \eta_D}{\eta_S - \eta_D}\tau\frac{Q}{P}. \tag{4.24}$$

Podstawiając następnie do (4.20), otrzymujemy rozwinięty wzór na koszt jałowy podatku:

$$DW = -\frac{\eta_S \eta_D}{2(\eta_S - \eta_D)}\tau^2\frac{Q}{P}. \tag{4.25}$$

Formułę (4.25) można uprościć, zakładając, że podaż jest doskonale elastyczna ($\eta_S = \infty$), oraz przyjmując, że $P = 1$. Mamy wówczas:

$$DW = -\frac{\eta_D \tau^2 Q}{2}. \tag{4.26}$$

Różniczkując, uzyskujemy wzór na krańcowy koszt jałowy względem podatku τ:

$$MDW = \partial DW / \partial\tau = -\eta_D Q\tau. \tag{4.27}$$

Wiedząc, że dochody podatkowe są równe $R = Q\tau$, można obliczyć, że krańcowy dochód z podatku, $MR = \partial R / \partial\tau$ jest równy Q. Wykorzystując regułę Ramseya (4.20), otrzymujemy ostatecznie wzór na optymalną stawkę podatku:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$\tau = -\frac{\lambda}{\eta_D}. \tag{4.28}$$

Równanie pokazuje, że optymalna stawka podatku pośredniego staje się tym wyższa, im niższa jest elastyczność popytu na dane dobro[12].

W bardzo podobny sposób można określić optymalną stawkę podatku dochodowego. Załóżmy, że konsumenci maksymalizują społeczną funkcję użyteczności V pod warunkiem, że dochody budżetowe mają być równe $\underline{R}$ tak, jak w równaniu (4.16). Mamy więc:

$$V = U_1 + U_2 + \ldots + U_n \tag{4.29}$$

oraz

$$R_1 + R_2 + \ldots R_n = \underline{R}. \tag{4.30}$$

Budując funkcję Lagrange'a, różniczkując i przyrównując do zera, otrzymujemy warunki pierwszego rzędu o postaci:

$$\frac{MU_i}{MR_i} = \lambda, \text{ dla } i = 1,2, \ldots, n, \tag{4.31}$$

gdzie $MU_i = \frac{\partial U_i}{\partial \tau_i}$ stanowi krańcową zmianę użyteczności konsumentów w wyniku nałożenia podatku na dobro i, a $MR_i = \frac{\partial R_i}{\partial \tau_i}$ jest krańcowym dochodem do budżetu z tytułu nałożenia podatku na dobro i.

Rysunek 4.9 ilustruje regułę Ramseya. Załóżmy, że podaż S jest doskonale elastyczna, co oznacza, że optymalna wysokość podatku zależy tylko od elastyczności popytu η_D (i od stałej λ). Gdy popyt jest mało elastyczny, krzywa popytu D' pozostaje stroma, a wartość η_D jest niska. W tej sytuacji stawka podatku τ może być odpowiednio duża (τ_1), ponieważ dla danego λ wyrażenie – $(1/\eta_D)$ osiąga wysoką

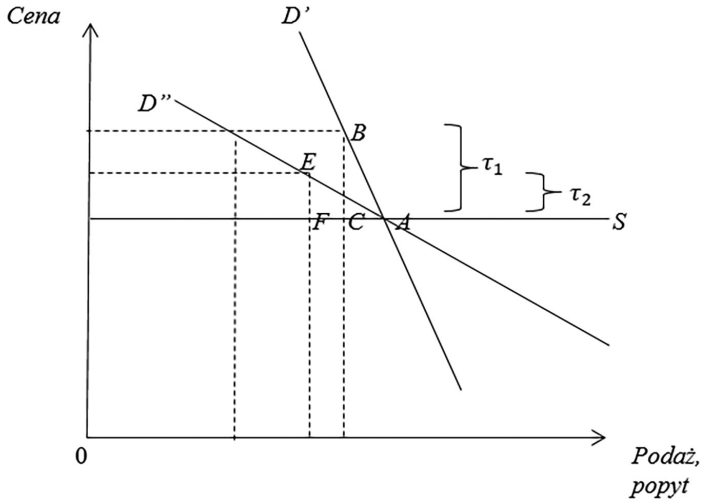
[12] Zgodnie z tą regułą, stawki podatku powinny być najwyższe w przypadku dóbr na które popyt jest sztywny. W praktyce oznacza to, że najwyżej opodatkowane będą artykuły pierwszej potrzeby, co może skutkować relatywnie większym obciążeniem podatkami grup najmniej zamożnych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wartość (co do modułu). Wówczas koszt jałowy podatku jest stosunkowo niewielki, równy powierzchni trójkąta *ABC*. Gdy popyt staje się bardzo elastyczny, wartość η_D jest duża (linia *D*" pozostaje płaska) i stawka podatku musi być odpowiednio niska (τ_2), ponieważ dla danego λ wyrażenie – (1/ η_D) osiąga niską wartość. Wówczas koszt jałowy podatku jest także niewielki, równy powierzchni trójkąta *AEF*. Można więc tak dobrać stawki podatku, aby w przypadku danej elastyczności popytu koszt jałowy był jak najmniejszy przy osiągnięciu wymaganego poziomu dochodów do budżetu $\underline{R}$.

Należy podkreślić, że pojęcie optymalnego podatku w ujęciu Ramseya odnosi się do podatku, który minimalizuje koszt jałowy, czyli straty w efektywności ekonomicznej spowodowane nałożeniem podatku. Definicja ta nie odnosi się natomiast do innych ważnych cech podatku, takich jak efektywność fiskalna czy pożądana skala redystrybucji dochodu w obrębie danego społeczeństwa (tj. zapewnienie sprawiedliwości społecznej). W szczególności pominięcie tego ostatniego aspektu systemu podatkowego oznacza, że podatek optymalny w sensie Ramseya nie uwzględnia zróżnicowanej sytuacji materialnej poszczególnych osób oraz ich zróżnicowanej zdolności do płacenia podatków.

Rysunek 4.9. Reguła Ramseya



Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Dlatego też w późniejszych pracach nad optymalnym podatkiem ekonomiści starali się uwzględnić zjawisko heterogeniczności podatników. Brytyjski ekonomista James Mirrlees zaproponował model, w którym bierze się pod uwagę zróżnicowanie podatników pod względem ich zdolności do uzyskiwania dochodu (Mirrlees 1971a, 1971b). Wiadomo, że zdolność ta zależy przede wszystkim od czynników takich jak talent i pracowitość, lecz obie te cechy nie są bezpośrednio obserwowalne. Zmienną obserwowalną stanowi bowiem jedynie dochód, który może być przedmiotem opodatkowania. Ze względu na malejącą krańcową użyteczność dochodu i konsumpcji, opodatkowanie zawsze zniechęca jednak do podejmowania dodatkowej pracy, co przekłada się na straty społeczne, choć skala tego zjawiska jest różna w zależności od poziomu dochodu. Skuteczne radzenie sobie z wymiennością między dążeniem do efektywności a dążeniem do równości staje się głównym wyzwaniem w zakresie polityki podatkowej każdego rządu.

4.6. Podatki a podział dochodu

4.6.1. Efektywność a równość

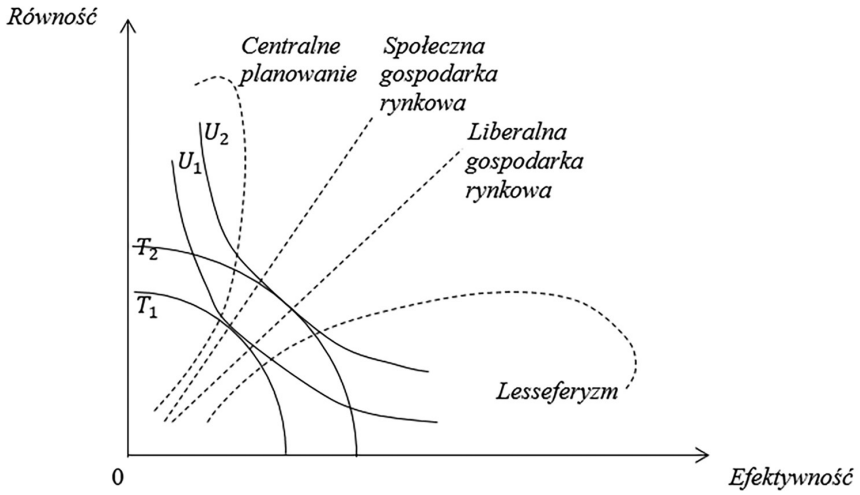
Wybór między efektywnością a równością (*fairness, equity*) stanowi jedno z centralnych zagadnień polityki gospodarczej. Jak już wiadomo, każdy podatek przyczynia się do obniżenia efektywności ekonomicznej, ponieważ ogranicza działalność obciążoną podatkiem. System podatkowy pozostaje jednak zarazem ważnym instrumentem redystrybucji dochodu od zamożniejszych do uboższych grup społecznych. Ponadto dochody podatkowe służą finansowaniu ważnych celów społecznych, takich jak powszechna edukacja, opieka zdrowotna czy wsparcie materialne dla osob niezdolnych do pracy. U podstaw podejmowanych decyzji leży przekonanie, że nadmierne rozpiętości dochodowe i majątkowe naruszają zasady sprawiedliwości społecznej oraz że współczesne społeczeństwa winny solidarnie pomagać tym osobom, które nie są w stanie samodzielnie zapewnić sobie godziwego bytu. Chodzi więc przede wszystkim o wsparcie dochodowe dla osób niezdolnych do pracy ze względu na podeszły wiek bądź też trwały lub przejściowy uszczerbek na zdrowiu, a także o pomoc dla dzieci i młodzieży. Mowa również o wsparciu dla osób, które wprawdzie są w stanie pracować, lecz ich dochody pozostają zbyt niskie, aby zaspokoić podstawowe potrzeby bytowe im oraz ich rodzinom,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zwłaszcza wielodzietnym, a także o pomoc dla osób, które utraciły pracę z powodów od nich niezależnych.

Większość ludzi akceptuje potrzebę wspierania osób niezdolnych do pracy – starych, słabych, chorych oraz dzieci. Ten rodzaj solidarności społecznej, funkcjonującej zarówno w wymiarze międzygrupowym, jak i międzypokoleniowym, stanowi jeden z fundamentów współczesnej cywilizacji. Wielu z nas może jednak wątpić w to, czy wsparcie dochodowe jest uzasadnione również dla osób uboższych, a jeśli tak, to w jakim zakresie. Część społeczeństwa uważa, że nierówności dochodowe są w gospodarce rynkowej czymś zupełnie normalnym i wynikają z indywidualnych różnic w zdolności i pracowitości poszczególnych osób. Zwolennicy tego poglądu hołdują zasadzie mówiącej o tym, że „każdy jest kowalem własnego losu". Inni wyrażają zaś przekonanie, że nierówności dochodowe są w znacznym stopniu niesprawiedliwe, ponieważ wynikają z różnic w sytuacji materialnej czy społecznej poszczególnych osób, co ogranicza dostęp do szans rozwojowych dla wielu z nich. Uważają oni również, że państwo powinno niwelować te różnice, zapewniać wszystkim możliwie równe szanse, a ewentualne braki wyrównywać za pomocą korygowania dochodów, ponieważ wszyscy mają podobne potrzeby i wszyscy zasługują na godne życie. W zależności od tego, który pogląd cieszy się większym poparciem społecznym, system podatkowy w danym kraju będzie w większym lub mniejszym stopniu spełniał funkcję redystrybucyjną.

Rysunek 4.10. Wymienność między efektywnością i równością



Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Jak już wiadomo, aby minimalizować straty na poziomie efektywności, podatki powinny być rozłożone równomiernie (lub neutralnie), ponieważ znika wówczas koszt jałowy. Jeśli podatki mają jednak pełnić rolę redystrybucyjną, muszą być z definicji zróżnicowane, to znaczy obciążać poszczególne osoby (lub poszczególne towary) w różnym stopniu. To z kolei generuje straty społeczne. Aby dana skala redystrybucji za pośrednictwem podatków znajdowała uzasadnienie, straty z tego tytułu muszą być zrównoważone korzyściami społecznymi, wynikającymi z bardziej wyrównanego rozkładu dochodów. Występuje więc wymienność między efektywnością i równością.

Nauka ekonomii nie jest w stanie ustalić, jak powinien wyglądać optymalny zakres redystrybucji. Nie potrafi też odpowiedzieć na pytanie o to, co jest, a co nie jest społecznie sprawiedliwe. Wynika to stąd, że ocena stopnia sprawiedliwości społecznej wymaga sądów wartościujących, aksjologicznych – a więc subiektywnych. W opinii różnych ludzi, w zależności od ich indywidualnej hierarchii wartości, poszczególne rozkłady dochodów mogą być bowiem bardziej lub mniej sprawiedliwe[13].

Fundamentalna wymienność między efektywnością a równością została przedstawiona na rysunku 4.10. Krzywe T_1 i T_2 są krzywymi transformacji systemowej, ilustrującymi możliwe warianty ustrojowe na danym poziomie rozwoju. Pozostają one wklęsłe do początku układu, co wskazuje na to, że najkorzystniejsze okazują się względnie zrównoważone kombinacje efektywności i sprawiedliwości, z kolei rozwiązania skrajne pociągają za sobą duże koszty w postaci utraconych poziomów efektywności lub równości. Krzywe U_1 i U_2 są krzywymi użyteczności społecznej, przedstawiającymi preferencje społeczne. Pozostają one wypukłe do początku układu, co oznacza, że krańcowa stopa substytucji między efektywnością i równością rośnie wraz ze wzrostem efektywności, i odwrotnie. W zestawieniu z przebiegiem krzywych T_1 i T_2 pokazuje to, że społeczeństwa ogólnie preferują systemy zrównoważone, a nie skrajne. Linie przerywane obrazują różne ścieżki rozwoju społeczno-gospodarczego, gdzie przywiązuje się różną wagę do efektywności względem równości. Systemy skrajne – takie jak centralne planowanie i wolnorynkowy

[13] Należy przywołać skrajny przykład, aby lepiej zrozumieć konflikt postaw i konflikt wartości. Część z nas może uznać, że bezrobotni zasługują na stały zasiłek, ponieważ utrata pracy była spowodowana czynnikami niezależnymi od nich, takimi jak kryzys gospodarczy, inni będą zaś zdania, że zasiłki im się nie należą, ponieważ zniechęcają bezrobotnych do poszukiwania pracy i skłaniają do odrzucania mniej atrakcyjnych ofert zatrudnienia. Zwolennicy pierwszego poglądu kierują się poczuciem zbiorowej solidarności, z kolei zwolennicy drugiego ze stanowisk poczuciem indywidualnej odpowiedzialności.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kapitalizm – ze względu na swą jednostronność prędzej czy później ustają w rozwoju i tracą pod względem obu cech. Na przykład system centralnego planowania po pewnym czasie załamał się pod ciężarem własnej niewydolności, przy czym w końcowej fazie tracił nie tylko na efektywności, ale także na równości (ze względu na nasilanie się różnych patologii, takich jak korupcja i woluntaryzm). Również system XIX-wiecznego kapitalizmu nie przetrwał ze względu na rosnące napięcia i protesty społeczne, i spowodowany nimi spadek efektywności. Stabilny rozwój zapewniają natomiast systemy względnie zrównoważone – mieszczące się w przedziale między społeczną gospodarką rynkową i liberalną gospodarką rynkową – które zmierzają do łączenia, co prawda w różnych proporcjach, cech efektywności i równości.

4.6.2. Zróżnicowanie dochodów: krzywa Lorenza i współczynnik Giniego

Dążenie do zapewnienia odpowiedniego zakresu równości – inaczej mówiąc sprawiedliwości – społecznej wymaga umiejętności mierzenia stopnia nierówności, w tym zwłaszcza nierówności dochodowych i majątkowych. Podstawowym sposobem opisu nierówności dochodowych w gospodarce, stosowanym przez ekonomistów, jest krzywa Lorentza, a podstawową miarę stopnia nierówności stanowi współczynnik Giniego. Obie miary przedstawione są na rysunku 4.11, gdzie na osi poziomej odłożono skumulowany odsetek ludności, a na osi pionowej skumulowany odsetek całkowitego dochodu w kraju odnośnie do danego udziału ludności. Linia *a* leżąca na przekątnej prostokąta odpowiada sytuacji pełnego egalitaryzmu, kiedy wszyscy członkowie społeczeństwa uzyskują dokładnie takie same dochody. Z kolei linia łamana składająca się z odcinków *AB* i *BC* odpowiada drugiej skrajności, gdy 100% całego dochodu przypada jednej osobie. Pomiędzy tymi skrajnościami zawierają się różne linie Lorentza, odzwierciedlające rozmaite stopnie zróżnicowania dochodów. Im bliżej przekątnej znajduje się krzywa Lorentza, tym mniej zróżnicowany jest rozkład dochodów. Łatwo zauważyć, że im bardziej wygięta w dół jest linia *b*, tym większe okazuje się zróżnicowanie dochodów. Inaczej mówiąc, im większe staje się pole powierzchni rozciągającej się między linią *a* i krzywą Lorentza, tym większe są rozpiętości dochodowe.

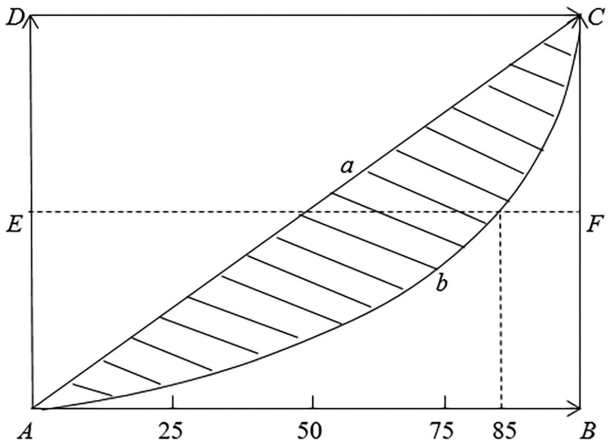
Załóżmy, że rozkład dochodów w danym społeczeństwie ilustruje krzywa Lorentza opisana linią *b*. Stopień zróżnicowania dochodów jest znaczny – połowa wszystkich środków (całość dochodów dzieli na pół linia *EF*) trafia do około 15%

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

społeczeństwa, podczas gdy druga połowa do pozostałych 85% ludności. Miarą stopnia nierówności może być w tej sytuacji stosunek pola powierzchni zawierającej się między liniami *a* i *b* (obszar zakreskowany na rysunku 4.11) oraz całą powierzchnią trójkąta znajdującego się poniżej linii *a* (trójkąt *ABC*). Tę relację nazywamy współczynnikiem Giniego.

Łatwo zauważyć, że wartość tego wskaźnika zawiera się w przedziale między zerem a jednością. Współczynnik jest równy zero, gdy linia *a* pokrywa się z linią *b* i rozkład dochodów jest wówczas idealnie równomierny. Wskaźnik ten znajduje się z kolei na poziomie równym jedności, gdy linia *b* pokrywa się z linią łamaną *ABC*. W praktyce skrajne wartości współczynnika Giniego nie występują. Jak pokazuje wykres 4.3, w krajach najbardziej rozwiniętych wartości te wahają się od 0,25 do 0,39, podczas gdy w krajach „wschodzących" i rozwijających się rozpiętości dochodowe są często wyższe[14].

Rysunek 4.11. Krzywa Lorentza i współczynnik Giniego (dane w %)



Źródło: opracowanie własne.

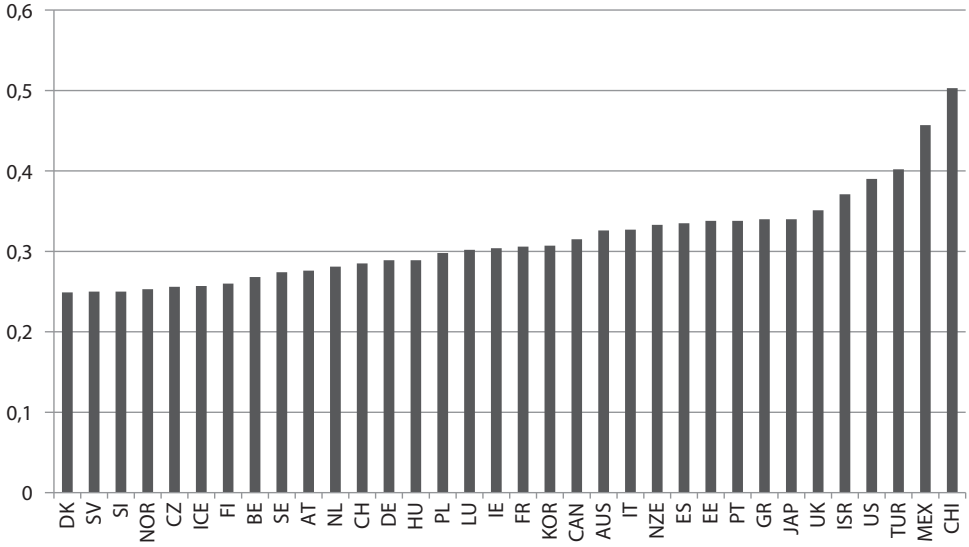
W Europie najmniejsze rozpiętości dochodowe występują tradycyjnie w krajach skandynawskich, a także w Słowacji, Słowenii i Czechach. W Polsce rozpiętości

[14] Według danych Banku Światowego, krajem o najwyższym zróżnicowaniu dochodów jest Republika Południowej Afryki (współczynnik Giniego wynosi tu 0,650). Bardzo wysokie wartości współczynnik ten osiąga również w krajach Ameryki Łacińskiej (np. w Brazylii – 0,527; Hondurasie – 0,535 czy Kolumbii – 0,574) – zob. http://data.worldbank.org/indicator/SI.POV.GINI (13.11.2015).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

dochodowe są natomiast zbliżone do średniej europejskiej – według danych OECD współczynnik Giniego w przypadku Polski wyniósł w 2012 roku 0,298, podczas gdy średnia wartość współczynnika odnośnie UE 27 sięgnęła 0,304 (i nie ulegała większym zmianom w okresie po 2000 roku). Najwyższe rozpiętości dochodowe w Europie występują w krajach bałkańskich i krajach bałtyckich. Z kolei wśród rozwiniętych państw pozaeuropejskich współczynnik Giniego jest generalnie nieco wyższy niż w Europie.

Wykres 4.3. Współczynnik Giniego[a] w krajach OECD w 2012 roku[b]



a) Dochody brutto, nieuwzględniające transferów pieniężnych.
b) Dane dotyczące Japonii pochodzą z 2009 roku, a w przypadku Kanady i Chile z 2011 roku.

Źródło: opracowanie własne na podstawie danych OECD, http://www.oecd.org/social/income-distribution-database.htm (22.07.2015).

Należy pamiętać o tym, że współczynnik Giniego może być liczony w przypadku dochodów brutto (przed opodatkowaniem i transferami pieniężnymi) lub w przypadku dochodów uwzględniających podatki i transfery. Oczywiście, podatki i transfery spłaszczają strukturę dochodów i obniżają wartości współczynnika.

Ekonomiści nie są zgodni co do tego, jaki poziom zróżnicowania dochodów jest optymalny. Często wskazuje się, że kraje skandynawskie – z generalnie wysokimi podatkami i szerokim zakresem usług społecznych – osiągnęły wysoki poziom trwałego rozwoju. Nie jest jednak do końca jasne, na ile taki model rozwoju pozostaje

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

uwarunkowany szczególnymi czynnikami kulturowymi i tradycją. Część ekonomistów podkreśla też negatywne skutki nadmiernego zróżnicowania dochodów w przypadku efektywności ekonomicznej i tempa wzrostu gospodarczego (Stiglitz 2015, Piketty 2015). Wskazuje się w tym kontekście na niestabilną sytuację społeczną i częste konflikty wewnętrzne w krajach o dużych rozpiętościach dochodowych i majątkowych (np. kraje Afryki i Ameryki Łacińskiej). Nie ma więc wprawdzie zgody co do tego, jakie zróżnicowanie należy uznać za „nadmierne", lecz większość ekonomistów jest zdania, że redystrybucja dochodów w pewnej skali staje się konieczna w celu zapewnienia stabilnego rozwoju.

Gdy zróżnicowanie dochodowe uznaje się za nadmierne, instrumentem służącym korekcie struktury dochodów są zwiększone podatki dochodowe, z reguły progresywne. Jak wykazano wyżej (rysunek 4.8), progresja podatkowa wiąże się jednak z silniejszym obniżeniem efektywności ekonomicznej. Jest to spowodowane zarówno osłabieniem bodźców do podejmowania pracy i prowadzenia działalności gospodarczej, nasileniem bodźców do ukrywania dochodów bądź przenoszenia ich do innych jurysdykcji podatkowych, jak i kosztem poboru podatków i dokonywania transferów. Ogólnie rzecz biorąc, straty społeczne w postaci spadku podaży pracy i poziomu konsumpcji są tym większe, im silniejsza staje się progresja podatkowa. Straty te trzeba porównać z korzyściami uzyskanymi dzięki bardziej równomiernemu rozłożeniu dochodów. Należy tu wziąć pod uwagę dwie kwestie. Po pierwsze, trzeba uwzględnić malejącą krańcową użyteczność dochodu – dodatkowa jednostka dochodu w rękach osoby uboższej ma większą społeczną wartość niż dodatkowa jednostka dochodu w rękach osoby bogatszej. Ten czynnik skłania do zwiększania skali redystrybucji. Po drugie, należy pamiętać o tym, że ogólna suma dochodów po redystrybucji będzie zawsze mniejsza niż przed redystrybucją z powodu strat efektywnościowych. To skłania z kolei do ograniczania skali redystrybucji. Okazuje się ona więc uzasadniona tylko wówczas, gdy bilans oddziaływania obu czynników pozostaje wciąż dodatni.

Poszukiwanie optymalnego rozwiązania może polegać na minimalizowaniu strat w efektywności ekonomicznej. Wymaga to przeniesienia części ciężaru redystrybucji z systemu podatkowego na stronę wydatków budżetowych. Przy lepszym rozdysponowaniu i dokładniejszym adresowaniu wydatków budżetowych można zapewnić praktycznie dowolną skalę redystrybucji, co pozwoliłoby utrzymać umiarkowane i neutralne – czyli najmniej szkodliwe – stawki podatkowe. System podatkowy podporządkowany byłby wówczas wymogom fiskalnym i efektywnościowym, z kolei cele równościowe oraz te związane z zapewnieniem sprawiedliwości

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

społecznej realizowano by za sprawą odpowiedniej polityki wydatków budżetowych (Rosati 2009, s. 60–61). Rozdzielenie tych funkcji pomiędzy stronę dochodowa i stronę wydatkową budżetu pozwoliłoby odciążyć system podatkowy od nadmiaru zadań i zwiększyłoby ogólna efektywność przy zapewnieniu pożądanej społecznie skali redystrybucji.

4.6.3. Ulgi i zwolnienia podatkowe

Redystrybucja dochodu w ramach systemu podatkowego może przyjmować postać nie tylko zróżnicowanych stawek podatkowych (progresja podatkowa), ale także ulg podatkowych i zwolnień od podatku. Ulga podatkowa (*tax credit*) to zmniejszenie należnego podatku (czyli inaczej odliczenie od podatku), zwolnienie podatkowe (*tax deduction*) oznacza zaś zmniejszenie dochodu do opodatkowania (czyli inaczej odliczenie od dochodu). Instrumenty te mają służyć korekcie początkowego rozkładu dochodów, najczęściej sprzyjając realizacji postulatu sprawiedliwości społecznej, choć nierzadko ich celem może być zachęcanie do inwestowania lub podejmowania działalności innowacyjnej. Obecnie przychodzi odpowiedzieć na pytanie, kiedy zastosowanie ulgi podatkowej ma sens oraz o ile lepsze stanowi ona rozwiązanie niż na przykład bezpośrednie subsydium dochodowe.

Rozpatrzmy najpierw efekty ulgi podatkowej na przykładzie działalności charytatywnej. Ponieważ publiczna działalność charytatywna może wypierać prywatną działalność charytatywną, stosowanie zwolnienia podatkowego w celu zachęcenia sektora prywatnego do przeznaczania środków na cele charytatywne pozostaje uzasadnione[15]. Załóżmy, że stawka podatku dochodowego wynosi $t = 50\%$ i że wydatki na cele charytatywne są odliczane od dochodu do opodatkowania. W takiej sytuacji, jeśli dana osoba przeznacza całość swoich dochodów netto na konsumpcję, to osiągnięcie poziomu konsumpcji w wysokości 100 PLN wymaga uzyskania dochodu w wysokości 100 PLN / $(1 - t)$, czyli 200 PLN Jeśli jednak dana osoba chce przeznaczyć 100 PLN na cele charytatywne (zamiast na konsumpcję), wystarczy, że uzyska dochód w wysokości tylko 100 PLN.

Zwolnienie podatkowe dotyczy jednak wszystkich osób dokonujących wydatków charytatywnych, w tym także i tych, które dokonałyby ich bez względu na ulgę

[15] Pomoc charytatywna, na przykład w formie wsparcia dla ubogich, jest dobrem publicznym. Sektor prywatny dostarcza z reguły dobra publiczne w niewystarczającej ilości z uwagi na problem „gapowicza". Dostarczanie danego dobra publicznego przez sektor publiczny zazwyczaj ogranicza więc podaż tego dobra w sektorze prywatnym.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

podatkową. W przypadku takich osób ulgi podatkowe są więc niepotrzebne, a ich stosowanie uszczupla dochody budżetowe, nie zwiększając zarazem skali działalności charytatywnej. Zwolnienie podatkowe ma zatem sens wówczas, gdy wywiera duży efekt krańcowy (*marginal impact*), czyli wpływa na zmianę zachowania, skłaniając nowych podatników do przeznaczania wydatków na cele charytatywne.

Załóżmy teraz, że przy stawce podatku dochodowego $t = 50\%$ prywatne wydatki charytatywne przeznaczane na pomoc dla ubogich wynoszą 1 mln PLN, co oznacza, że na ich sfinansowanie potrzebne jest uzyskanie dochodów równych 2 mln PLN. Rząd, chcąc je pobudzać, decyduje się zwolnić te wydatki z podatku. Przyjmijmy wobec tego, że dzięki zastosowaniu tej ulgi wydatki charytatywne rosną do 1,5 mln PLN. Jakie są tego skutki? Rząd uzyskał dodatkowe wydatki charytatywne z sektora prywatnego w wysokości 0,5 mln PLN, lecz stracił dochody budżetowe także w wysokości 0,5 mln PLN, ponieważ ci, którzy dokonywali do tej pory wydatków charytatywnych, uiszczając pełną kwotę podatków, teraz korzystają z ulgi w wysokości 0,5 mln PLN. Gdyby więc te utracone dochody przeznaczyć wcześniej bezpośrednio na pomoc dla ubogich, efekt całej tej operacji byłby zerowy (a nawet ujemny, jeśli weźmie się pod uwagę efekt wypierania oraz koszty administracyjne i transakcyjne). Tak więc zwolnienie podatkowe ma społeczny sens wtedy, gdy efekt krańcowy (*marginal impact*), czyli wzrost pożądanych wydatków charytatywnych, jest większy niż „efekt podkrańcowy" (*inframarginal impact*), tj. utrata dochodów budżetowych z tytułu przyznanego zwolnienia.

Warto zauważyć, że ulga podatkowa w istocie oznacza zwiększenie dochodu pozostającego do dyspozycji danej osoby. Jeśli więc celem jest poprawa sytuacji dochodowej danej osoby, powstaje pytanie, czy bardziej opłaca się stosować ulgi podatkowe, czy raczej bezpośrednie subsydiowanie dochodu? Jak już wspomniano, wspierając ubogich poprzez bezpośrednie subsydia, rząd wypiera część prywatnego dofinansowania, wspierając zaś ubogich pośrednio poprzez ulgę podatkową dla prywatnego sektora, rząd traci część dochodów budżetowych. Wynika z tego, że ulga podatkowa powinna być stosowana zamiast bezpośredniego subsydiowania dochodu, gdy wzrost prywatnych wydatków na cele charytatywne z każdej złotówki pochodzącej z ulgi podatkowej jest większy niż spadek publicznych wydatków charytatywnych przypadających na każdą złotówkę spadku wydatków rządowych. Gdy efekt krańcowy równa się utraconym dochodom i występuje efekt wypierania, ulga podatkowa na wydatki charytatywne będzie zawsze lepszym rozwiązaniem niż bezpośrednie subsydium (tę prawidłowość potwierdza większość badań empirycznych). Kiedy efekt krańcowy

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jest z kolei większy od utraconych dochodów, powstałą nadwyżkę należy porównać z efektem wypierania.

Jedną z najbardziej popularnych ulg podatkowych stanowi możliwość wspólnego rozliczania podatku dochodowego przez małżeństwa. Ulga ta ma głównie cel społeczny, ponieważ za jej pośrednictwem państwo wspiera finansowo małżeństwa, które ponoszą koszty wychowywania dzieci (zob. ramka 4.5). Inna popularną ulgą podatkową – również uzasadnioną względami społecznymi – jest zwolnienie z podatku części kosztów związanych z nabyciem mieszkania[16].

Ramka 4.6. Wspólne opodatkowanie małżeństw

W wielu krajach małżeństwa korzystają z możliwości wspólnego rozliczania podatków. Czy tego rodzaju ulga jest uzasadniona? A jeśli tak, to dlaczego? Załóżmy, że funkcjonuje system progresywnego podatku dochodowego: obowiązuje stawka 20% w przypadku dochodu nieprzekraczającego 50 000 PLN rocznie oraz stawka 30% powyżej tego poziomu dochodu. Załóżmy też, że mamy dwa małżeństwa A i B: w małżeństwie A mąż zarabia rocznie 80 000 PLN, a żona 20 000 PLN, a w małżeństwie B mąż i żona zarabiają po 50 000 PLN rocznie. Obie pary zarabiają więc tyle samo – 100 000 PLN rocznie.

Gdy nie ma ulgi małżeńskiej i podatki naliczane są indywidualnie, małżeństwo A płaci razem 19 000 PLN + 4000 PLN = 23 000 PLN, podczas gdy para B uiszcza podatek w wysokości 10000 PLN + 10000 PLN = 20 000 PLN. Oznacza to jednak, że pary osiągające taki sam łączny dochód płacą różne podatki – można uznać, że jest to niesprawiedliwe. Gdy nie stosuje się ulgi małżeńskiej, ale dochody

[16] Jeśli ktoś chce mieszkać samodzielnie, może albo wynająć mieszkanie (dom), albo je kupić. Jak wobec tego dylematu powinna kształtować się polityka podatkowa? Załóżmy, że w przypadku wynajmu mieszkania (domu) czynsz wynosi 1000 PLN miesięcznie, podczas gdy w sytuacji kupna koszt kredytu mieszkaniowego (odsetki) wynosi również 1000 PLN miesięcznie. Jeśli ponadto przyjmiemy, że funkcjonuje doskonały rynek kapitałowy, to oba rozwiązania są z punktu widzenia potencjalnego lokatora/nabywcy jednakowo korzystne – między wynajmem a kupnem mieszkania (domu) nie występuje żadna różnica. Lecz w praktyce obie te transakcje pod względem opodatkowania traktowane są odmiennie. W wielu krajach odsetki od kredytu mieszkaniowego odlicza się od podstawy opodatkowania w przeciwieństwie do kosztu czynszu, który nie jest traktowany preferencyjnie. Oznacza to, że rząd w ten sposób pośrednio subsydiuje prywatną własność nieruchomości. Dlaczego? Ponieważ posiadanie mieszkania (lub domu) na własność niesie dodatnie efekty zewnętrzne: nabycie mieszkania wiąże nabywców z danym otoczeniem i środowiskiem w sposób znacznie trwalszy niż wynajem mieszkania, stabilizuje więzi społeczne, zwiększa troskę o otoczenie i środowisko (co podnosi też wartość sąsiednich domów) oraz wzmaga aktywność obywatelską i polityczną.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

małżeństw są konsolidowane, pary A i B płacą tyle samo – po 25 000 PLN, ponieważ sumy ich dochodów są jednakowe. W tej sytuacji z punktu widzenia obciążeń podatkowych obie pary ponoszą więc takie same koszty, lecz w rzeczywistości każda z nich płaci więcej, niż gdyby poszczególni małżonkowie rozliczali się indywidualnie. Oznacza to, że osoby te są obciążone dodatkowym nieformalnym podatkiem tylko dlatego, że pozostają małżeństwem. Można tę dodatkową daninę nazwać podatkiem „małżeńskim" (*marriage tax* – zob. Gruber 2007). Dlatego też w Polsce stosuje się ulgę podatkową: małżeństwa sumują dochody, dzielą otrzymany wynik na pół, od każdej połowy oblicza się należny podatek, a następnie oba należne podatki sumuje się. W systemie ulgi małżeńskiej para A płaci ponosi więc koszty podatkowe w wysokości: 10 000 PLN + 10 000 PLN = 20 000 PLN, podczas gdy para B uiszcza również taką samą kwotę, tj. 10 000 PLN + 10 000 PL N = 20 000 PLN. Uzasadnieniem w przypadku zastosowania ulgi małżeńskiej jest więc chęć uniknięcia efektu „karania" małżeństw dodatkowym podatkiem „małżeńskim" oraz uznanie, że mogą one mieć niższą zdolność płacenia podatków ze względu na konieczność ponoszenia dodatkowych kosztów w związku z obowiązkiem wychowywania dzieci.

Literatura cytowana

Barro R.J., *A Cross-Country Study of Growth, Saving and Government*, "NBER Working Paper" 1989, no. 2855.

Burman L.E., Slemrod J., *Taxes in America. What Everyone Needs to Know*, Oxford University Press, Oxford 2012.

Feldstein M., *Effects of Taxes on Economic Behavior*, "NBER Working Paper" 2008, no. 13745.

Gruber J., *Public Finance and Public Policy*, Worth Publishers, New York 2007.

Harberger A., *The Incidence of the Corporation Income Tax*, "Journal of Political Economy" 1962, no. 70.

Hsing Y., *Estimating the Laffer curve and policy implications*, "Journal of Socio-Economics" 1996, no. 25(3).

Laffer A., *The Laffer Curve, Past, Present and Future*, Heritage Foundation, 2004, http://www.heritage.org/research/reports/2004/06/the-laffer-curve-past-present-and-future (1.06.2004).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Mankiw G., Weinzierl M., Yagan D., *Optimal Taxation in Theory*, "NBER Working Paper" 2009, no. 15071.

Mirrlees J., Diamond P., *Optimal Taxation and Public Production I: Production Efficiency*, "American Economic Review" 1971a, no. 61.

Mirrlees J., Diamond P., *Optimal Taxation and Public Production II: Tax Rules*, "American Economic Review" 1971b, no. 61.

Pecorino P., *Tax rates and tax revenues in a model of growth through human capital accumulation*, "Journal of Monetary Economics" 1995, no. 36(3).

Pigou A.C., *The Economics of Welfare*, Macmillan, London 1920.

Piketty T., *Kapitał w XXI wieku*, Wydawnictwo Krytyki Politycznej, Warszawa 2015.

Ramsey F., *A Contribution to the Theory of Taxation*, "Economic Journal" 1927, no. 37.

Rosati D.K., *Europejski Model Społeczny. Osiągnięcia, problemy, kierunki zmian*, w: *Europejski Model Społeczny. Doświadczenia i przyszłość*, red. D.K. Rosati, PWE, Warszawa 2009.

Slemrod J., *Optimal taxation and optimal tax systems*, "Journal of Economic Perspectives" 1990, no. 4(1).

Sosnowski M., *Znaczenie zasad podatkowych dla systemu podatkowego*, „Finanse, Rynki Finansowe, Ubezpieczenia" 2012, nr 52, t. 708.

Stiglitz J.E., *Ekonomia sektora publicznego*, Wydawnictwo Naukowe PWN, Warszawa 2004.

Stiglitz J.E., *Cena nierówności. W jaki sposób dzisiejsze podziały społeczne zagrażają naszej przyszłości?*, Wydawnictwo Krytyki Politycznej, Warszawa 2015.

Stuart C.E., *Swedish Tax Rates, Labor Supply and Tax Revenues*, "The Journal of Political Economy" 1981, no. 89(5).

Trabandt M., Uhlig H., *How Far Are We From The Slippery Slope? The Laffer Curve Revisited*, "NBER Working Paper" 2009, no. 15343.

Wagner A., *Finanzwissenschaft*, t. 1–4, Lipsk 1877–1889.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozdział 5

POLITYKA PIENIĘŻNA

5.1. Podstawowe pojęcia

5.1.1. Wprowadzenie

Polityka pieniężna polega na regulowaniu ilości pieniądza w obiegu. Ta krótka definicja może sugerować, że polityka pieniężna jest działaniem czysto mechanicznym, mało skomplikowanym. Tak jednak nie jest. Regulowanie ilości pieniądza w obiegu nie może odbywać się w sposób dowolny, ponieważ na rynku nie może się go znajdować ani za dużo, ani za mało. Gdy w gospodarce jest za dużo pieniądza, rosną wydatki na zakup dóbr, usług i aktywów, co przy danych zdolnościach wytwórczych wywołuje napięcia inflacyjne. Kiedy zaś pieniądza jest za mało, wydatki na zakup dóbr i usług spadają, produkcja w gospodarce zwalnia, rośnie bezrobocie i pojawia się zagrożenie recesją i deflacją. Pieniądza w obiegu powinno być „w sam raz", czyli tyle, ile wynosi bieżące zapotrzebowanie gospodarki, zapewniające jej równowagę. Posługując się terminologią ekonomiczną, stwierdzimy, że podaż pieniądza powinna równać się popytowi na pieniądz. Nie jest to zresztą żadna nowość. W rozdziale 2 zetknęliśmy się już z zagadnieniem relacji popytu i podaży na rynku pieniądza krajowego, kiedy omawiany był wzór dotyczący równowagi na rynku pieniądza w ramach modelu makroekonomicznego IS-LM.

Postulat zakładający, że podaż pieniądza powinna równać się popytowi jest intuicyjnie oczywisty, lecz, bynajmniej, nie rozwiązuje jeszcze kwestii odpowiedniego sposobu regulowania ilości pieniądza w obiegu. W tym celu trzeba bowiem znać bieżący poziom popytu na pieniądz w gospodarce. Popyt ten zgłaszają uczestnicy rynku, którym środki pieniężne są potrzebne do dokonywania transakcji zakupu dóbr i usług oraz do nabywania rozmaitych aktywów rzeczowych i finansowych. Precyzyjna odpowiedź na pytanie o wielkość tego popytu nie jest jednak łatwa,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ponieważ potrzeby transakcyjne uczestników rynku nie stanowią wielkości bezpośrednio obserwowalnej. Można je jedynie szacować z większym lub mniejszym prawdopodobieństwem na podstawie rozmaitych informacji i wskaźników. Co więcej, popyt ten ulega zmianom w czasie pod wpływem oddziaływania rozmaitych czynników. W rezultacie władza gospodarcza, która odpowiada za regulowanie ilości pieniądza w obiegu – czyli za prowadzenie polityki pieniężnej – staje przed trudnym zadaniem, związanym z ustawicznym dostosowywaniem podaży pieniądza do zmieniającego się, a zarazem bezpośrednio nieobserwowalnego popytu zgłaszanego przez uczestników rynku.

We współczesnych gospodarkach rynkowych władzą odpowiedzialną za politykę pieniężną jest bank centralny danego kraju (Wojtyna 1998, Kaźmierczak 2003, Kokoszczyński 2004). Wynika to przede wszystkim z tego, że pozostaje on jedyną instytucją, która ma prawo emitować pieniądz gotówkowy. Przyjmijmy więc, że bank ten potrafi w jakiś sensowny sposób oszacować popyt na pieniądz i będzie w związku z tym starał się tak regulować podaż pieniądza, aby była ona równa popytowi. Wówczas pojawią się jednak dalsze pytania. Pieniądz, który krąży w gospodarce nie jest, bynajmniej, jednorodny. Pozostaje nim oczywiście gotówka, czyli banknoty i monety będące oficjalnie w obiegu, lecz pieniądzem są także depozyty złożone na rachunkach w bankach komercyjnych. Znamy to z codziennego doświadczenia – dzięki kartom płatniczym możemy dokonywać transakcji zakupu dóbr i usług w sposób bezgotówkowy, oczywiście pod warunkiem, że na bieżącym rachunku bankowym znajdują się odpowiednie środki W praktyce, dzięki możliwości szybkiego transferowania środków na rachunki bieżące, można także korzystać z kapitału zgromadzonego na bankowym rachunku oszczędnościowym. Całkowita ilość pieniądza zasilająca w danym momencie gospodarkę obejmuje więc zarówno gotówkę w obiegu, jak i wkłady pieniężne znajdujące się na rachunkach bankowych. Kłopot z regulowaniem tak szeroko zdefiniowanej ilości pieniądza w obiegu polega na tym, że bank centralny kontroluje bezpośrednio (i to nie w pełni) wyłącznie podaż pieniądza gotówkowego, podczas gdy tworzenie depozytów – czyli podaż pieniądza depozytowego – zależy głównie od decyzji banków komercyjnych. Warto przy tym zaznaczyć, że pieniądza depozytowego jest z reguły znacznie więcej niż gotówki[1]. Aby więc kształtować ogólną podaż pieniądza na poziomie zapewniającym równowagę, bank centralny musi tak oddziaływać na banki komercyjne,

[1] Na przykład w Polsce pod koniec 2015 roku pieniądz gotówkowy w obiegu osiągnął wartość 149,7 mld PLN, podczas gdy całkowita suma środków zgromadzonych na rachunkach bankowych – tj. wszystkich depozytów – wyniosła 996,0 mld PLN – czyli ponad sześć i pół razy więcej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

aby ilość wytwarzanego przez nie pieniądza depozytowego była zgodna z oczekiwaniami i preferencjami banku centralnego. Przedmiotem niniejszego rozdziału będzie więc szczegółowe omówienie sposobów tego oddziaływania.

W pierwszej kolejności zostaną przypomniane takie zagadnienia jak funkcje i rodzaje pieniądza oraz główne miary obiegu pieniężnego. Następnie przyjdzie wyjaśnić pojęcie popytu na pieniądz i omówić czynniki, które go kształtują. Przedmiotem dalszej części rozdziału będzie podaż pieniądza. Omówimy kolejno cele oraz instrumenty polityki pieniężnej, kanały transmisji impulsów polityki pieniężnej oraz reguły prowadzenia polityki pieniężnej. Szczególną uwagę należy poświęcić w tym kontekście strategii bezpośredniego celu inflacyjnego. Na koniec przyjdzie też wskazać na nowe wyzwania, przed jakimi stanęła polityka pieniężna w okresie ostatniego kryzysu finansowego i gospodarczego lat 2008–2012.

5.1.2. Funkcje, rodzaje i miary pieniądza

Na bardzo wczesnym etapie rozwoju cywilizacyjnego człowiek wytwarzał podstawowe dobra wyłącznie na użytek własny i swoich najbliższych, a pieniądz jako środek wymiany nie był potrzebny. Gdy wraz z upływem czasu wykształciły się jednak początki gospodarki towarowej i ludzie zaczęli wymieniać się wytworzonymi produktami w ramach lokalnych wspólnot, a także miedzy nimi, wymiana naturalna typu „produkt za produkt" szybko okazała się niewystarczalna i zaczął występować problem tzw. podwójnej zbieżności (*double coincidence*). Polegał on na tym, że członek wspólnoty, posiadający nadmiar jakiegoś produktu, np. ziarna, i chcący wymienić tę nadwyżkę na inne produkty, których nie posiada, np. na skóry, musi znaleźć innego członka wspólnoty, który nie tylko ma nadwyżkę skór, lecz jednocześnie potrzebuje także więcej ziarna. Gdy posiadacz skór nie jest zainteresowany ziarnem, ale potrzebuje np. narzędzi, „podwójna zbieżność" nie wystąpi i nie dojdzie do bezpośredniej transakcji wymiany. Trzeba w takiej sytuacji znaleźć trzeciego członka wspólnoty, dysponującego nadmiarem narzędzi, a zarazem chcącego nabyć więcej ziarna. W takiej sytuacji wymiana naturalna staje się utrudniona, a w miarę wzrostu liczby wytwarzanych produktów komplikuje się jeszcze bardziej i okazuje praktycznie niemożliwa. W szczególności rośnie wykładniczo czas i nakład pracy związany ze zbieraniem informacji na temat dostępności poszczególnych produktów i możliwości ich wymiany.

W takich warunkach zachodzi potrzeba posługiwania się dobrem uniwersalnym szczególnego rodzaju, pozwalającym rozwiązać opisany problem podwójnej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zbieżności[2]. Dobro takie musi być łatwe w obrocie, transporcie i przechowywaniu, musi być także powszechnie akceptowalne. W ten właśnie sposób, na pewnym etapie rozwoju gospodarki towarowej pojawia się pieniądz funkcjonujący jako specyficzne dobro, które umożliwia nabywanie poszczególnych produktów w obrębie danej społeczności. W takich warunkach wspomniany wyżej posiadacz ziarna może sprzedać je w zamian za pieniądze każdemu, kto chce kupić ziarno, a następnie nabyć skóry od kogokolwiek, kto jest gotów je sprzedać. Pieniądz ułatwia więc dokonywanie transakcji kupna i sprzedaży. Co więcej, możliwość wymiany pieniądza na dowolne dobra i usługi sprawia, że reprezentuje on określoną siłę nabywczą, a zatem nabiera samoistnej wartości. Jeśli siła nabywcza pozostaje względnie stabilna w czasie, pieniądz może służyć także jako środek służący przechowywaniu zgromadzonego bogactwa.

W ten sposób ukształtowały się trzy główne funkcje pieniądza – jako środka wymiany, środka płatniczego i środka przechowywania wartości. W przypadku pierwszej z nich pieniądz jest wykorzystywany do nabywania dóbr i usług oraz do obejmowania w posiadanie innych aktywów finansowych i rzeczowych. Jest to więc funkcja transakcyjna. W ramach spełniania drugiej funkcji – płatniczej – pieniądz służy do regulowania zobowiązań finansowych (długów, podatków). Wreszcie, pełniąc trzecią funkcję, pieniądz umożliwia przechowywanie posiadanego bogactwa, a więc występuje jako zasób majątkowy. W tym ostatnim przypadku pieniądz może być traktowany jako szczególny rodzaj aktywu finansowego – jego charakterystyczną cechą jest to, że, choć pieniądz gotówkowy nie przynosi dochodu, zapewnia w zamian wygodę i bezpieczeństwo, ponieważ można go wymienić na określone produkty zawsze, gdy tylko zajdzie taka potrzeba. Tę ważną cechę pieniądza, polegającą na możliwości jego szybkiej i bezwarunkowej zamiany na towary i inne aktywa, nazywamy płynnością (*liquidity*).

Współcześnie pieniądz pełni jeszcze jedną istotną funkcję, występując w roli jednostki rozrachunkowej (tzw. „*numéraire*"). Pozwala bowiem wówczas porównywać i dodawać do siebie (agregować) wartości różnych niejednorodnych produktów. Dzięki temu jesteśmy w stanie obliczyć wartość całej produkcji wytworzonej w danym przedsiębiorstwie, sektorze, regionie bądź też w danym kraju, w określonym czasie. Na tej podstawie można także szacować koszty i zyski, co pozwala planować inwestycje i wyceniać ich opłacalność.

[2] Potrzebę „wynalezienia" pieniądza można więc uzasadniać niedoskonałą informacją o możliwościach wymiany naturalnej i rosnącym wykładniczo kosztem jej zbierania w miarę wzrostu liczby członków wspólnoty oraz wytwarzanych i wymienianych produktów.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wspomniano już wyżej o tym, że pieniądz w gospodarce przyjmuje różne formy. Występuje w postaci zarówno gotówki, jaki i depozytów zgromadzonych na rachunkach bankowych. Mechanizm tworzenia tych dwóch kategorii pieniądza pozostaje jednak różny. Dlatego też w pierwszej kolejności należy dokonać ważnego rozróżnienia: na pieniądz emitowany przez bank centralny oraz pieniądz emitowany przez banki komercyjne. Pierwszy z nich nazywany jest pieniądzem rezerwowym (*reserve money*), bazą monetarną (*monetary base*) lub pieniądzem „wielkiej mocy" (*high powered money*). Pochodzenie tych tajemniczo brzmiących terminów zostanie wyjaśnione w dalszej części rozdziału. Pieniądz rezerwowy emitowany przez bank centralny składa się z gotówki i płynnych rezerw pieniężnych banków komercyjnych, gromadzonych na rachunkach w banku centralnym. Banki komercyjne utrzymują te rezerwy ze względów „ostrożnościowych". Aby zapewnić bowiem niezakłócone dostarczanie usług bankowych, w tym zwłaszcza zdolność do bieżących wypłat gotówki i regulowania zapadających zobowiązań, banki muszą utrzymywać część swoich aktywów w postaci płynnych środków pieniężnych – albo jako środki gotówkowe w kasie banku (tzw. pogotowie kasowe), albo na specjalnym rachunku rezerwowym w banku centralnym. Rezerwy te mogą mieć charakter obowiązkowy, jeśli są wymagane prawem, lub dobrowolny. W większości krajów banki centralne ustalają wielkość obowiązkowych rezerw jako pewien odsetek od zgromadzonych depozytów. Na przykład w Polsce stopa rezerwy obowiązkowej wynosiła w ostatnich latach 3,5%, podczas gdy w większości pozostałych krajów Unii Europejskiej wskaźnik ten jest ustalony na poziomie 2,0%.

O ile pieniądz gotówkowy jest w miarę jednorodny (składa się z banknotów i monet pozostających w obiegu), o tyle pieniądz depozytowy emitowany przez banki komercyjne przyjmuje różne formy. Wśród depozytów (wkładów) można zasadniczo wyróżnić depozyty bieżące (zwane także depozytami na żądanie lub depozytami *à vista*), depozyty oszczędnościowe oraz depozyty terminowe. Różnią się one przede wszystkim płynnością i dochodowością. Depozyty bieżące są najbardziej płynne, ponieważ można nie tylko wypłacać z nich środki praktycznie bez ograniczeń, lecz także bezpośrednio dokonywać transakcji zakupu dóbr i usług w ciężar rachunku bieżącego przy pomocy debetowych kart płatniczych. Jednocześnie oprocentowanie wkładów znajdujących się na tych rachunkach pozostaje jednak minimalne lub wręcz zerowe. Najmniej płynne są z kolei depozyty terminowe, których rozwiązanie przed ustalonym terminem podlega wielu ograniczeniom i pociąga za sobą określone koszty. Zapewniają one jednak z reguły najwyższe oprocentowanie. Miejsce pośrednie między rachunkami bieżącymi i terminowymi

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zajmują rachunki oszczędnościowe, na których środki są oprocentowane, a wypłaty mogą być dokonywane dość swobodnie.

Istnienie różnych form pieniądza i odmiennych kategorii depozytów pieniężnych o zróżnicowanej płynności oraz dochodowości sprawia, że w praktyce, gdy mówi się o wielkości obiegu pieniężnego, trzeba sprecyzować, jakie kategorie depozytów ma się na myśli. W tym celu posługujemy się różnymi miarami ilości pieniądza – noszą one miano agregatów pieniężnych (*monetary aggregates*). Najwęższym z nich jest agregat M0, który najczęściej definiowany jest jako suma gotówki w obiegu (C), gotówki w kasach banków (*cash in vaults* – C_v) oraz rezerw obowiązkowych (R_o) i rezerw dobrowolnych (R_d) utrzymywanych przez banki komercyjne na rachunkach w banku centralnym:

$$M0 = C + C_v + \left(R_o + R_d\right). \tag{5.1}$$

Należy jednak pamiętać, że do podaży pieniądza zalicza się tylko pieniądz znajdujący się w obiegu, a więc gotówkę (C), natomiast rezerwy bankowe zgromadzone na rachunkach w banku centralnym, a także gotówka w kasach banków są wyłączone z obiegu i dlatego nie wlicza się ich do podaży pieniądza. Agregat M0 odpowiada definicji pieniądza rezerwowego (bazy monetarnej)[3]; nazywa się go także „wąskim pieniądzem" (*narrow money*). W Polsce, podobnie jak w większości pozostałych krajów UE, od pewnego już czasu nie stosuje się w statystyce pieniężnej wielkości agregatu M0, a zamiast tego wykorzystuje się jeszcze węższą kategorię pieniądza gotówkowego w obiegu.

Szerszym agregatem pieniężnym jest agregat M1, który obejmuje gotówkę znajdującą się w obiegu oraz depozyty (wkłady) bieżące (W_b):

$$M1 = C + W_b. \tag{5.2}$$

Agregat M1 odzwierciedla wielkość podaży pieniądza pełniącego funkcję środka wymiany. Depozyty bieżące są bowiem wykorzystywane w formie bezgotówkowej bezpośrednio do dokonywania płatności za nabywane dobra i usługi, dlatego z punktu widzenia funkcji środka wymiany depozyty te są traktowane jak pieniądz gotówkowy. Między innymi z tego właśnie powodu w większości krajów

[3] W niektórych krajach, na przykład w USA, agregat M0 obejmuje tylko gotówkę w obiegu, a więc stanowi jedynie część bazy monetarnej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

rachunki te nie są oprocentowane. Kilkadziesiąt lat temu dominującą formą płatności dokonywanych z rachunków bieżących było wystawianie czeków, co zresztą pozostaje ciągle bardzo popularne w Stanach Zjednoczonych. W Europie wykorzystuje się obecnie w tym celu głównie debetowe karty bankowe funkcjonujące w ramach elektronicznych systemów płatniczych.

Jeszcze szerszą miarą pieniądza jest agregat M2, który składa się z agregatu M1 oraz większości pozostałych depozytów, w tym oszczędnościowych (W_o), terminowych (W_t) oraz depozytów terminowych z okresem wypowiedzenia do 3 miesięcy ($W_{t<3}$):

$$M2 = M1 + \left(W_o + W_t\right) = C + \left(W_b + W_o + W_t + W_{t<3}\right). \tag{5.3}$$

Agregat M2 przedstawia więc wielkość podaży pieniądza pełniącego funkcję zarówno środka wymiany, jak i środka służącego gromadzeniu wartości. Depozyty oszczędnościowe, a zwłaszcza terminowe stanowią raczej formę przechowywania majątku, w mniejszym stopniu są zaś traktowane jako zasób pieniądza transakcyjnego. Dlatego też pozostają oprocentowane, a wypłaty z nich mogą podlegać pewnym ograniczeniom. Warto jednocześnie zaznaczyć, że poszczególne kraje różnią się między sobą co do sposobu szczegółowego definiowania agregatu M2. Różnice te zależą od tego, w jakim stopniu określone rodzaje rachunków bankowych są traktowane przez krajowe władze nadzorcze jako częściowe substytuty „wąskiego pieniądza" (M0), a w jakim stopniu uznaje się je za aktywa finansowe.

Najszerszą miarą ilości pieniądza w obiegu stosowaną przez większość państw jest agregat M3. Poza agregatem M2 obejmuje on także pozostałe zobowiązania sektora bankowego. W Polsce, podobnie jak wśród większości pozostałych krajów Unii Europejskiej, na agregat M3 składa się zdecydowana większość agregatu M2, a dodatkowo także zobowiązania systemu bankowego z tytułu pożyczek zaciągniętych w banku centralnym (operacje z przyrzeczeniem odkupu – tzw. operacje *repo* – O_r), z tytułu lokat w funduszach inwestycyjnych rynku pieniężnego (*money market funds* – I_{mmf}) oraz z tytułu wyemitowanych papierów wartościowych z okresem zapadalności do 2 lat (*covered bonds* – B_c). Jednocześnie z agregatu M3 wyłączone są depozyty terminowe z okresem zapadalności powyżej 2 lat, według klasyfikacji unijnej traktowane jako inwestycja finansowa, a nie jako zasób pieniądza[4]:

[4] Definicja agregatu M3 w USA jest nieco inna: obejmuje on całość M2, a także depozyty powyżej 100 000 USD, depozyty eurodolarowe oraz saldo operacji *repo*.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$M3 = M2 + \left(O_r + I_{mmf} + B_c\right). \tag{5.4}$$

Agregat M3 bywa także nazywany „szerokim pieniądzem" (*broad money*), choć w niektórych krajach w użyciu znajdują się jeszcze szersze miary pieniądza (M4, M5), obejmujące także inne instrumenty dłużne emitowane przez banki, a także rządy (np. w Stanach Zjednoczonych czy Wielkiej Brytanii).

Tabela 5.1. Składniki pieniądza rezerwowego (bazy monetarnej) oraz definicje agregatów pieniężnych stosowane aktualnie przez Europejski Bank Centralny oraz przez Narodowy Bank Polski

Zobowiązania monetarnych instytucji finansowych	Pieniądz rezerwowy (M0)	M1	M2	M3
Gotówka w obiegu	X	X	X	X
Gotówka w kasach banków	X			
Rezerwy banków na rachunkach w banku centralnym	X			
Depozyty bieżące		X	X	X
Depozyty oszczędnościowe			X	X
Depozyty terminowe z okresem zapadalności do 2 lat			X	X
Operacje z przyrzeczeniem odkupu (*repo*)				X
Udziały w funduszach rynku pieniężnego				X
Instrumenty dłużne z okresem zapadalności do 2 lat				X

Źródło: opracowanie własne na podstawie EBC (*The Monetary Policy of the ECB*, European Central Bank, Frankfurt am Main 2004, s. 37).

Zakres poszczególnych agregatów pieniężnych przedstawiono w tabeli 5.1. Ukazano w niej zobowiązania (pasywa) systemu bankowego jako składniki pieniądza rezerwowego (bazy monetarnej) oraz składniki szerszych agregatów pieniężnych, zgodnie z metodologią stosowaną aktualnie przez Europejski Bank Centralny i przez Narodowy Bank Polski. Z kolei w tabeli 5.2 znalazły się wielkości poszczególnych agregatów pieniężnych odnotowane w Polsce na koniec 2013, 2014 i 2015 roku.

Na koniec 2015 roku w obiegu gotówkowym w Polsce znajdowało się prawie 150 mld PLN, pozostała część wyemitowanego pieniądza gotówkowego w kwocie ponad 13 mld PLN pozostawała w kasach banków i była wyłączona z obiegu pieniężnego. Kwota zgromadzonych depozytów bieżących na rachunkach bankowych wynosiła około 542 mld PLN, co oznacza, że agregat pieniężny M1 był blisko czterokrotnie większy od zasobu gotówki znajdującej się w obiegu. Najszerszy

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

agregat pieniężny M3 osiągnął na koniec 2015 roku wartość około 1155 mld PLN – niemal osiem razy większą niż pieniądz gotówkowy w obiegu. Agregat M2 był tylko niewiele niższy, co pokazuje, że w warunkach polskich różnice między M2 i M3 są praktycznie nieistotne. Jednocześnie porównanie zmian w agregatach pieniężnych pozwala zauważyć, że w 2015 roku najszybciej rósł pieniądz gotówkowy w obiegu (o 15,1%), z kolei całkowita podaż pieniądza mierzona miarą M3 wzrastała wolniej (o 9,1%), podobnie jak suma depozytów terminowych (3,5%).

Tabela 5.2. Wielkości agregatów pieniężnych w Polsce w latach 2013–2015 (dane w mln PLN)

Agregat pieniężny	Stan na dzień 31.12.2013 roku	Stan na dzień 31.12.2014 roku	Stan na dzień 31.12.2015 roku	Zmiana w roku 2015 (dane w %)
Pieniądz gotówkowy w obiegu (poza kasami banków)	114 403	130 030	149 686	15,1
Pieniądz gotówkowy w kasach banków (wyłączony z obiegu)	11 481	12 631	13 265	5,0
Depozyty i inne zobowiązania bieżące	441 432	476 253	542 408	13,9
Agregat M1	555 835	606 283	692 094	14,2
Depozyty i inne zobowiązania do 2 lat włącznie	404 507	438 268	453 572	3,5
Depozyty z terminem wypowiedzenia do 3 miesięcy	2	2	1	–27,8
Agregat M2	960 345	1 044 553	1 145 667	9,7
Operacje z przyrzeczeniem odkupu (*repo*)	12 277	10 350	7 267	–29,8
Dłużne papiery wartościowe z terminem pierwotnym do 2 lat	6 286	4 264	2 467	–42,1
Jednostki uczestnictwa w funduszach rynku pieniężnego	0	0	0	–
Agregat M3	978 908	1 059 167	1 155 401	9,1

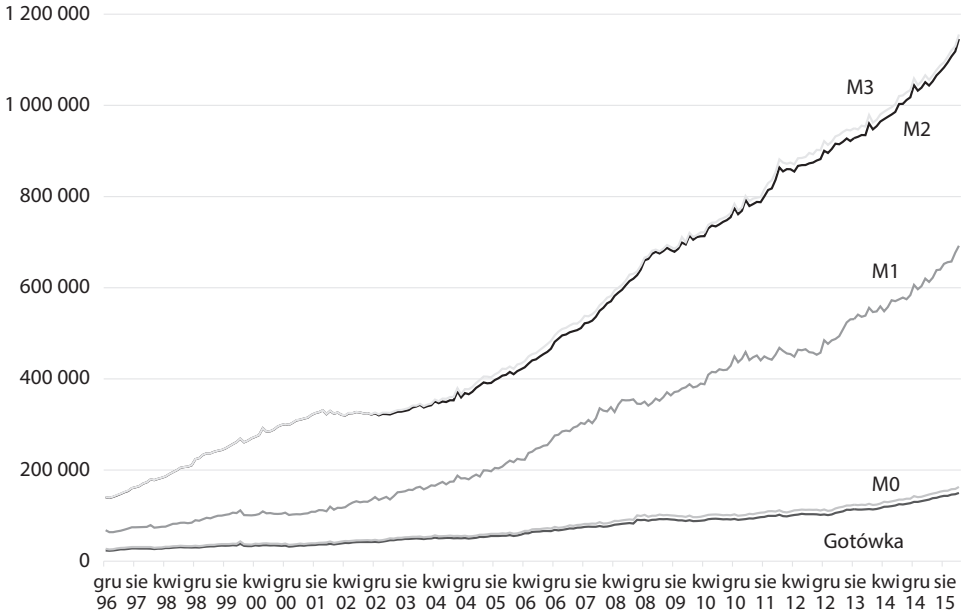
Źródło: NBP.

Wykres 5.1 przedstawia zmiany w zakresie głównych agregatów pieniężnych w Polsce w latach 1997–2015. Rosnąca masa pieniądza w obiegu jest zjawiskiem normalnym w warunkach rozwijającej się gospodarki. Przy czym tempo wzrostu podaży pieniądza pozostaje z reguły wyższe niż tempo wzrostu nominalnego PKB głównie dzięki szybszemu wzrostowi depozytów bankowych, co z kolei ma związek z rosnącą zamożnością społeczeństwa i rozwojem sektora finansowego

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w miarę postępu gospodarczego. Szybszy wzrost depozytów wywołuje szybszy wzrost szerokich agregatów pieniężnych w relacji do wzrostu gotówki w obiegu. W latach 1997–2015 średnie roczne tempo wzrostu agregatu M3 w Polsce wynosiło 12,2%, podczas gdy tempo wzrostu pieniądza gotówkowego w obiegu tylko 9,8%. Był to przejaw systematycznego wzrostu oszczędności krajowych, a także dowód na to, że klasyczne równanie ilościowe pieniądza nie jest w stanie wyjaśnić zmian dokonujących się w zakresie podaży pieniądza bez uchylenia założenia o stałej szybkości obiegu (zob. następny punkt). Tempo wzrostu podaży pieniądza szybsze niż tempo wzrostu nominalnego PKB pokazuje, że szybkość obiegu systematycznie spada, podczas gdy relacja podaży pieniądza do PKB regularnie wzrasta.

Wykres 5.1. Zmiany wielkości agregatów pieniężnych w Polsce od grudnia 1996 roku do grudnia 2015 roku (dane w mln PLN)



Źródło: opracowanie własne na podstawie danych NBP.

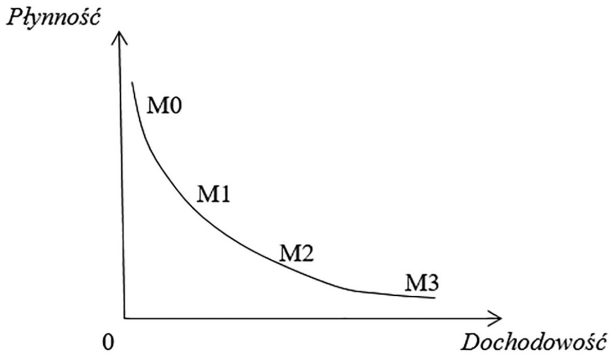
Miarą względną wielkości obiegu pieniężnego jest relacja sumy pieniądza w obiegu do wielkości PKB danego kraju. Relacja ta nazywana jest współczynnikiem monetyzacji gospodarki i stanowi wymierny wskaźnik poziomu rozwoju gospodarczego. Im wyższy poziom rozwoju państwa oraz im wyższy poziom zamożności społeczeństwa, tym z reguły wyższy jest także współczynnik monetyzacji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Na przykład w Polsce relacja M2 do PKB wzrosła z 38,2% w 2000 roku do 57,7% w 2014 roku. Lecz wśród większości krajów OECD współczynnik monetyzacji jest jeszcze wyższy i w państwach o największym obiegu pieniężnym znacznie przekracza 150%[5]. Współczynnik monetyzacji osiąga szczególnie wysokie wartości w krajach o znacznie rozwiniętym sektorze bankowym, gdzie przyjmuje się depozyty także od obywateli innych państw[6].

Jak wspomniano wcześniej, poszczególne składniki szerszych agregatów pieniężnych różnią się płynnością i dochodowością – między tymi cechami zachodzi wymienność (*trade off*), co ilustruje rysunek 5.1. Dzięki temu uczestnicy rynku mogą wybierać różne kombinacje posiadanych agregatów pieniężnych w zależności od indywidualnych preferencji wynikających z płynności i dochodowości zasobów pieniężnych.

Rysunek 5.1. Wymienność pomiędzy płynnością a dochodowością agregatów pieniężnych



Źródło: opracowanie własne.

Omówione wyżej agregaty pieniężne konstruowane są metodą prostego dodawania poszczególnych składników. Elementy te nie pozostają jednak w pełni homogeniczne i w niejednakowym stopniu pełnią funkcje pieniądza. Na przykład depozyty

[5] Na przykład stosunek M2 do PKB w 2012 roku we Francji wynosił 158%, w Wielkiej Brytanii 162% w Niemczech 173%, w Austrii 175%, w Szwajcarii 188%, w Hiszpanii 194%, w Holandii 238% i w Japonii 242%. Z kolei w USA współczynnik monetyzacji osiągnął wartość jedynie 87%, co wynika z mniejszego udziału depozytów bankowych w oszczędnościach społeczeństwa i znacznego zaangażowania instrumentów pozabankowych, takich jak akcje i obligacje – na podstawie danych Banku Światowego, http://data.worldbank.org/indicator/FM.LBL.MQMY.GD.ZS/countries?display=default (13.03.2016).

[6] Na przykład Cypr (264%), Hongkong (336%) czy Luksemburg (480%).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

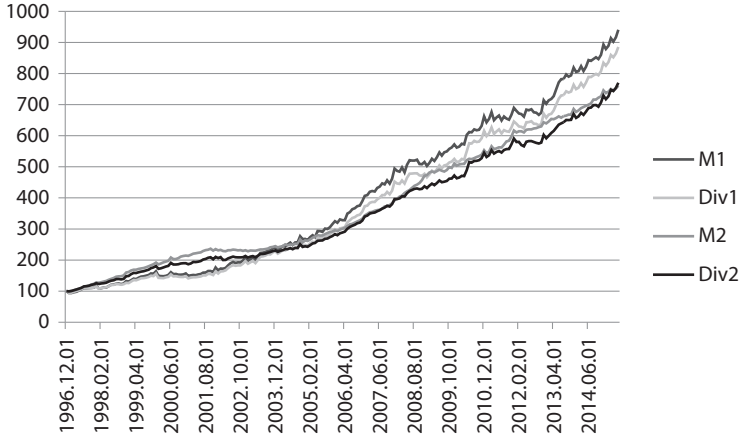
terminowe są mniej płynne niż gotówka i dlatego w mniejszym stopniu wykorzystuje się je w funkcji środka wymiany i środka płatniczego. Bierze się to stąd, że koszt ich wykorzystania do celów transakcyjnych jest wysoki ze względu na utratę części lub całości oprocentowania. Proste zsumowanie zasobu gotówki i wkładów pieniężnych nie zapewnia więc pełnego obrazu wielkości pieniądza znajdującego się w obiegu z punktu widzenia rzeczywistej wartości nabywczej. Dlatego też ekonomiści posługują się często innymi miarami pieniądza niż proste agregaty M0, M1, M2 i M3, aby uwzględnić różnice zakresie płynności i kosztu wykorzystania poszczególnych rodzajów pieniądza. Spośród tych miar najbardziej popularne są tzw. agregaty Divisia.

Agregaty Divisia (lub indeksy Divisia) różnicują aktywa pieniężne z pod względem relatywnych kosztów ich utrzymywania w porównaniu do pewnej wartości referencyjnej, zazwyczaj równej dochodowości aktywów pieniężnych o najniższej płynności (takich jak na przykład wkłady terminowe i wkłady z określonym terminem wypowiedzenia). Poszczególnym składnikom podaży pieniądza przypisywane są odpowiednie wagi odzwierciedlające wielkość dostarczanych przez nie usług pieniężnych. U podstaw konstrukcji takiego agregatu leży więc założenie, że uczestnicy rynku utrzymują pieniądz ze względu na „usługi", jakie on spełnia. Agregat Divisia stanowi więc sumę nominalnych wartości poszczególnych składników określonego agregatu (M1, M2, M3) skorygowanych wagami odzwierciedlającymi stopień ich faktycznej płynności oraz koszt wykorzystania. Tak zmodyfikowane agregaty dają bardziej precyzyjny obraz ilości pieniądza znajdującego się w obiegu z perspektywy pieniądza pełniącego funkcję środka wymiany i środka płatniczego (Barnett 2000, Cieśla 1999, Kluza 2001).

Ze względu na zastosowanie innej konstrukcji wskaźniki zmian nominalnych agregatów pieniężnych i agregatów Divisia różnią się między sobą. Obliczenia przeprowadzone w Narodowym Banku Polskim pokazują, że faktyczna dynamika zmian podaży pieniądza M1 w Polsce mierzona indeksami Divisia jest nieco niższa niż ta mierzona wielkościami nominalnymi. W okresie od grudnia 1996 roku do czerwca 2015 roku podaż pieniądza M1 wzrosła nominalnie o 841%, podczas gdy podaż tego samego agregatu mierzona indeksem Divisia (Div1) wzrosła tylko o 785%. Z kolei dynamika nominalna agregatu M2 była nieco niższa niż ta mierzona indeksem Div2 (odpowiednio wzrost o 661% i 670% – zob. wykres 5.2). Wskazuje to, że po uwzględnieniu różnic w płynności rzeczywista dynamika zmian płynnych środków pieniężnych różniła się od dynamiki nominalnej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 5.2. Dynamika zmian podaży pieniądza w Polsce w okresie od grudnia 1996 roku do czerwca 2015 roku: agregaty nominalne i agregaty Divisia



Źródło: opracowanie własne na podstawie danych NBP.

5.1.3. Popyt na pieniądz

Dlaczego ludzie chcą mieć pieniądze? Oczywiście nie chodzi tu wcale o posiadanie pieniędzy dla samych pieniędzy. Wyobraźmy sobie, że jesteśmy na bezludnej wyspie, gdzie nie ma żadnych sklepów. W takiej skrajnej sytuacji pieniądze – same w sobie – nie posiadają żadnej wartości. W normalnych warunkach ludzie gromadzą pieniądze, ponieważ w zamian za nie mogą nabyć potrzebne im dobra i usługi, a także aktywa finansowe i rzeczowe. To właśnie możliwość wymiany pieniędzy na wszelkie inne towary sprawia, że chcemy mieć pieniądze – gdyby taka sposobność nie istniała, gromadzenie pieniędzy byłoby bezcelowe. Pieniądz, za który nie można kupić niczego, co jest potrzebne do życia, byłby bezwartościowy[7]. Wynika stąd, że zapotrzebowanie na pieniądze zależy od tego, co i w jakiej ilości można za nie kupić – czyli od ich siły nabywczej. To spostrzeżenie pozwala zidentyfikować oraz zdefiniować czynniki określające popyt na pieniądz.

Wyróżniamy wśród nich przede wszystkim te zmienne, które określają potrzeby ludzi w zakresie nabywania dóbr i usług. Do czynników tych należy dochód

[7] Jak zauważył już ponad sto lat temu amerykański ekonomista Irving Fisher, sam pieniądz jako taki nie zaspokaja bezpośrednio żadnych rzeczywistych potrzeb ludzkich – natomiast pozwala nabywać dobra i usługi, które te potrzeby zaspokajają (Fisher 1911, s. 32).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

realny (y) oraz poziom cen (P). Im wyższym dysponujemy dochodem, tym większe są z reguły nasze wydatki na zakup dóbr i usług, choć zależność ta nie jest liniowo proporcjonalna (w przypadku wyższych poziomów dochodu udział wydatków bieżących w dochodzie pozostaje niższy). Podobnie im wyższy staje się poziom cen, tym większe nominalnie są nasze wydatki. Popyt na pieniądz rośnie zatem wraz ze wzrostem dochodu oraz poziomu cen. Ta kategoria popytu na pieniądz zależy od pożądanej sumy (poziomu) transakcji kupna dóbr i usług, w związku z czym jest kształtowana przez motywy transakcyjne.

Kolejnym czynnikiem kształtującym popyt na pieniądz jest nominalna stopa procentowa (i). Określa ona dochód z lokat bankowych, a zatem definiuje koszt alternatywny (*opportunity cost*) utrzymywania części majątku w postaci płynnych zasobów pieniężnych. Powszechnie wiadomo, że pieniądze znajdujące się w portfelu, a także środki zgromadzone na rachunku bieżącym w banku nie przynoszą dochodu. Dlaczego więc wszyscy utrzymują pewne rezerwy pieniądza gotówkowego? Czyni się tak ze względu na ważną cechę pieniądza, jaką jest płynność, czyli jego zdolność do praktycznie bezkosztowej zamiany na dobra i usługi. Rezygnujemy w ten sposób z dochodu odsetkowego, jaki moglibyśmy uzyskać, wkładając środki pieniężne do banku na lokatę oszczędnościową lub terminową, lecz czynimy to ze względu na wygodę i oszczędność czasu, którą podczas dokonywania bieżących transakcji zapewnia nam gotówka. Im wyższa staje się jednak nominalna stopa procentowa, tym wyższy będzie utracony dochód odsetkowy, czyli koszt alternatywny utrzymywania pieniądza w postaci gotówki i płynnych rezerw. Wobec tego im wyższa okazuje się stopa procentowa, tym niższy jest popyt na pieniądz, ponieważ bardziej opłaca się ulokować go w banku na rachunku przynoszącym dochód. Ta kategoria popytu na pieniądz kształtowana jest przez motyw, który umownie nazywa się motywem spekulacyjnym, ponieważ ludzie kierują się w tym przypadku dążeniem do uzyskania dochodu z zasobu majątku, jaki stanowi pieniądz.

Motywy spekulacyjne określają także tę część popytu na pieniądz, która związana jest z kosztem zamiany innych aktywów realnych i finansowych na pieniądz (c). Gdy w ramach własnego majątku posiada się jakieś papiery wartościowe (obligacje, akcje), można sprzedać je w zamian za środki pieniężne na rynku kapitałowym, lecz transakcja taka z reguły pociąga za sobą pewne koszty. Na przykład, gdy rynek papierów wartościowych jest płytki, szybka sprzedaż obligacji bądź akcji może wymagać znacznego obniżenia ich cen, co z punktu widzenia inwestora pociąga za sobą określone straty. Podobnie dzieje się w przypadku innych aktywów, takich jak waluty obce, złoto, antyki czy nieruchomości. Im wyższy staje się

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przeciętny koszt zamiany aktywów niepieniężnych na pieniądz, tym niższy będzie popyt na te aktywa, a tym wyższy popyt na pieniądz.

Oprócz motywów transakcyjnych i spekulacyjnych popyt na pieniądz regulują także czynniki psychologiczne. Ludzie różnią się między sobą pod względem skłonności do podejmowania ryzyka, odmiennie mogą też oceniać wygodę związaną z posiadaniem płynnych środków pieniężnych. Na przykład, jeśli system bankowy nie cieszy się zaufaniem i ulokowanie pieniędzy w banku jest postrzegane jako działanie obciążone ryzykiem, popyt na pieniądz wzrośnie, ponieważ zwiększy się skłonność do utrzymywania płynnych zasobów pieniężnych. W tym przypadku popyt na pieniądz jest więc kształtowany przez motyw ostrożnościowy (*precautionary*).

Rozpoznanie czynników określających popyt na pieniądz pozwala budować i szacować jego bardziej sformalizowane funkcje. Pierwsze próby podejmowano już w XVII i XVIII wieku, kiedy John Locke, a następnie David Hume zdefiniowali klasyczną ilościową teorię pieniądza (*quantity theory of money* – zob. Rothbard 1995a). Jej główny element stanowiło słynne równanie ilościowe o następującej postaci:

$$MV = PT, \tag{5.5}$$

gdzie M oznacza egzogeniczną podaż pieniądza, V – szybkość obiegu pieniądza na jednostkę czasu (np. na rok), P – poziom cen w gospodarce, a T – poziom realnych transakcji dokonywanych w gospodarce w danym okresie (rok). Wielkość T może być traktowana jako odpowiednik wielkości realnej produktu krajowego brutto y. Przy założeniu, że V i T są stałe, równanie (5.5) pokazuje, że poziom cen P zależy od wielkości podaży pieniądza M, w związku z czym lewa strona równania określa jego prawą stronę.

W warunkach równowagi popyt na pieniądz L jest więc równy podaży pieniądza:

$$L = M = \left(\frac{1}{V}\right)Py. \tag{5.6}$$

Równanie (5.6) pokazuje, że, zgodnie z teorią ilościową, popyt na pieniądz rośnie wraz ze wzrostem poziomu cen i poziomu produkcji (dochodu), a spada razem ze wzrostem szybkości obiegu. Wysoka szybkość obiegu pozwala obsłużyć potrzeby obrotu gospodarczego przy pomocy mniejszej ilości pieniądza.

Równanie ilościowe podlegało ewolucji. W pierwszych latach XX wieku grupa ekonomistów z uniwersytetu w Cambridge sformułowała inną jego wersję, w której

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

punktem wyjścia był popyt na pieniądz, a nie jego podaż. Tzw. wersja Cambridge zakłada, że pieniądz, oprócz pełnienia funkcji transakcyjnych, występuje też jako środek służący przechowywaniu wartości. Zgodnie z tą koncepcją pewna część podaży pieniądza nie zostanie wykorzystana podczas bieżących transakcji, lecz będzie utrzymywana przez uczestników rynku w postaci rezerw gotówkowych (*cash balances*) – w celu zapewnienia sobie wygody i bezpieczeństwa. Oznaczając relację pożądanych rezerw gotówkowych M do dochodu nominalnego (Py) jako k, otrzymujemy równanie popytu na pieniądz w wersji Cambridge:

$$L = M = kPy. \tag{5.7}$$

W powyższym równaniu zależność przyczynowo-skutkowa ma kierunek odwrotny – to prawa strona równania określa jego lewą stronę.

Równanie ilościowe, zarówno w postaci klasycznej, jak i w wersji Cambridge, miało szereg oczywistych mankamentów. Nie uwzględniało przede wszystkim zmian stóp procentowych oraz ich wpływu na popyt na pieniądz, a także zakładało stały charakter funkcji popytu na pieniądz, w tym zwłaszcza wielkości V i współczynnika k. Na niedostatki te wskazywał m.in. John Maynard Keynes. Współczesna postać funkcji popytu na pieniądz uwzględnia wpływ nominalnych stóp procentowych i kosztu wymiany innych aktywów na pieniądz, czyli zmiany popytu wynikające z dążenia do optymalizacji portfela aktywów utrzymywanych przez uczestników rynku. Współczesna funkcja popytu na pieniądz nominalny L^n przyjmuje najczęściej postać:

$$L = L^n(P, y, i, c). \tag{5.7a}$$

Podczas gdy funkcja popytu na pieniądz realny L^r przedstawia się następująco:

$$\frac{L}{P} = L^r(y, i, c). \tag{5.7b}$$

Keynes nazywał równanie (5.7a) funkcją preferencji płynności (*liquidity preference function*), ponieważ przedstawia ono jaką ilość płynnych środków pieniężnych chcą utrzymywać uczestnicy rynku w warunkach zmiennych stóp procentowych oraz zmiennych kosztów związanych z wymianą aktywów na pieniądz.

Utrzymanie równowagi na rynku pieniądza wymaga, aby jego podaż równała się popytowi. W dalszych częściach niniejszego rozdziału będziemy często

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

odwoływać się do funkcji preferencji płynności. Sięgając po wielkości nominalne i pomijając (w celu uproszczenia zapisu) współczynnik *c*, możemy zapisać warunek równowagi na rynku pieniądza w następujący sposób:

$$M = L(P, y, i). \tag{5.8}$$

Równanie (5.8) jest najczęściej wykorzystywaną w literaturze ekonomicznej ogólną postacią warunku równowagi pieniężnej.

5.1.4. Pieniądz w bilansie banku centralnego, w bilansach banków komercyjnych i w bilansie skonsolidowanym systemu bankowego

Jak wspomniano wyżej, pieniądz pozostaje dla większości uczestników rynku składnikiem majątku i dlatego może być traktowany jako szczególny rodzaj aktywów. Lecz w przypadku emitującej go instytucji stanowi on zobowiązanie. Pomijając wyjątek pieniądza kruszcowego, który posiadał wartość samoistną, wartość pieniądza współczesnego wywodzi się ze swoistej umowy, tj. formalnego zobowiązania instytucji emitującej pieniądz do zapewnienia możliwości wymiany środków pieniężnych na wszelkie inne dobra, usługi oraz pozostałe aktywa. Oznacza to, że obieg pieniężny we współczesnych gospodarkach opiera się na zaufaniu uczestników rynku do instytucji emitującej pieniądz –wynika to zarówno z przepisów prawa, jak i z wiarygodności tych instytucji.

W obiegu pieniężnym występują, o czym była już mowa, dwa rodzaje pieniądza – pieniądz gotówkowy, który jest emitowany przez bank centralny, oraz pieniądz depozytowy, emitowany przez banki komercyjne. W obu przypadkach pieniądz stanowi zobowiązanie instytucji monetarnych wobec pozostałych uczestników rynku, dlatego też w bilansach tych instytucji występuje zawsze po stronie pasywów. Odzwierciedla to bilans banku centralnego, zagregowany bilans banków komercyjnych oraz skonsolidowany bilans monetarnych instytucji finansowych, czyli całego sektora bankowego[8].

[8] Zgodnie ze standardami przyjętymi w ramach Europejskiego Systemu Banków Centralnych, sektor monetarnych instytucji finansowych (MIF) w Polsce obejmuje Narodowy Bank Polski, banki komercyjne działające w Polsce, rezydujące w naszym kraju oddziały instytucji kredytowych i oddziały banków zagranicznych oraz spółdzielcze kasy oszczędnościowo-kredytowe (SKOK). Sektor MIF nie zrzesza banków znajdujących się w stanie likwidacji bądź upadłości oraz pozostających w fazie organizacji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 5.3. Bilans Narodowego Banku Polskiego (ujęcie monetarne) – stan na dzień 31.12.2014 roku (dane w mln PLN)

AKTYWA		PASYWA	
1. Kredyty i pożyczki oraz inne należności od podmiotów krajowych	22,5	1. Pieniądz gotówkowy w obiegu ogółem	142 661,0
1.1. Pozostałe monetarne instytucje finansowe	0,0	1.1. Pieniądz gotówkowy w obiegu poza kasami banków	130 029,8
1.2. sektor instytucji rządowych i samorządowych	0,0	1.2. Gotówka w kasach banków	12 631,2
1.3. Pozostałe sektory krajowe[a)]	22,5	2. Depozyty podmiotów krajowych	69 990,5
2. Dłużne papiery wartościowe wyemitowane przez podmioty krajowe	0,0	2.1. Pozostałe monetarne instytucje finansowe	49 302,6
2.1. Pozostałe monetarne instytucje finansowe	0,0	2.2. Instytucje rządowe szczebla centralnego	19 431,9
2.2. Sektor instytucji rządowych i samorządowych	0,0	2.3. Pozostałe sektory krajowe[a)]	1 256,0
3. Papiery z prawem do kapitału i udziały wyemitowane przez podmioty krajowe	53,8	3. Emisja dłużnych papierów wartościowych	84 661,4
3.1. Pozostałe monetarne instytucje finansowe	0,0	4. Kapitał i rezerwy	12 851,1
3.2. Pozostałe sektory krajowe[a)]	53,8	4.1. Fundusze podstawowe	2 743,6
4. Aktywa zagraniczne	358 922,3	4.2. Rezerwy[b)]	10 107,5
5. Aktywa trwałe	2 224,6	5. Pasywa zagraniczne	26 825,8
6. Pozostałe aktywa	241,3	6. Pozostałe pasywa	24 474,7
Aktywa ogółem	**361 464,5**	**Pasywa ogółem**	**361 464,5**

a) Niemonetarne instytucje finansowe, instytucje samorządowe, fundusze ubezpieczeń społecznych oraz sektor niefinansowy.
b) Rezerwy celowe, umorzenia i odpisy aktualizujące.

Źródło: NBP.

W tabeli 5.3 przedstawiono bilans Narodowego Banku Polskiego zgodny ze stanem na koniec 2014 roku. Jak widać, bezwzględnie dominującą część aktywów polskiego banku centralnego – ponad 99% – stanowią aktywa zagraniczne, czyli oficjalne aktywa rezerwowe państwa. Są to rezerwy walutowe Polski, zarządzane przez NBP. W ich skład wchodzą zagraniczne papiery wartościowe, złoto monetarne oraz waluty obce. Pozostałe aktywa są wielokrotnie niższe, choć pozostają istotne z perspektywy funkcjonowania banku centralnego – jak na przykład aktywa trwałe o wartości księgowej ponad 2,2 mld PLN. Strona pasywów jest natomiast bardziej zróżnicowana. Dwie największe pozycje to wyemitowany pieniądz gotówkowy

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

znajdujący się w obiegu w postaci banknotów i monet na kwotę 142,7 mld PLN oraz wyemitowane przez NBP dłużne papiery wartościowe (bony pieniężne NBP) w kwocie ponad 84,7 mld PLN. Ponadto do zobowiązań polskiego banku centralnego należą także depozyty podmiotów krajowych, głównie banków komercyjnych, w kwocie 49,3 mld PLN oraz instytucji rządowych szczebla centralnego o łącznej wartości 19,4 mld PLN, a także zobowiązania wobec instytucji zagranicznych na kwotę 26,8 mld PLN. Po stronie pasywów znajdują się także rezerwy tworzone przez bank centralny na pokrycie ewentualnych strat spowodowanych zmianami w wycenie aktywów zagranicznych z tytułu zmian kursu walutowego.

Bilans banku centralnego pokazuje, że pieniądz gotówkowy w obiegu stanowi zobowiązanie banku centralnego wobec posiadaczy środków, czyli wobec przedsiębiorstw, gospodarstw domowych i instytucji publicznych. W wieku XIX oraz w pierwszej połowie wieku XX zobowiązanie to odnosiło się do bezwarunkowej gotowości banku centralnego do wymiany pieniądza gotówkowego (papierowego) na kruszec – złoto i srebro. Większość państw świata zrezygnowała jednak z systemu waluty złotej w latach 30. XX wieku. Współcześnie zobowiązanie to polega zaś na gwarantowaniu przez bank centralny możliwości bezwarunkowej wymiany pieniądza gotówkowego na wszelkie dobra, usługi oraz inne aktywa. Pokrycie tych zobowiązań stanowią aktywa banku centralnego, w tym przede wszystkim rezerwy walutowe oraz należności przekazywane do banku centralnego przez sektor publiczny i banki komercyjne z tytułu zaciągniętych kredytów, pożyczek i wyemitowanych papierów wartościowych. Dowodzi to, że współczesne systemy pieniężne opierają się na zaufaniu uczestników rynku do instytucji emitującej pieniądz.

W tabeli 5.4 zamieszczono zagregowany bilans banków komercyjnych w Polsce zgodny ze stanem na koniec 2014 roku. Pieniądz depozytowy emitowany przez banki komercyjne stanowi z ich strony zobowiązanie wobec pozostałych uczestników rynku – banków, przedsiębiorstw, gospodarstw domowych oraz instytucji sektora rządowego i samorządowego. Pokrycie tych zobowiązań stanowią głównie dwa rodzaje aktywów – udzielone kredyty i pożyczki, oraz zakupione przez banki dłużne i udziałowe papiery wartościowe, wyemitowane przez inne podmioty krajowe. Pozostałe aktywa, w tym zwłaszcza posiadane środki pieniężne, odgrywają zdecydowanie mniejszą rolę. Banki wykazują także zobowiązania wobec pozostałych uczestników rynku z tytułu wyemitowanych dłużnych papierów wartościowych. Środki uzyskane w ten sposób zwiększają zdolność banków do udzielania pożyczek oraz pozyskiwania innych aktywów.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 5.4. Bilans zagregowany banków – stan na dzień 31.12.2014 roku (dane w mln PLN)

AKTYWA		PASYWA	
1. Kredyty i pożyczki oraz inne należności	1 107 309,4	1. Depozyty i inne zobowiązania wobec podmiotów krajowych	1 031 300,5
1.1. Monetarne instytucje finansowe	100 071,9	1.1. Monetarne instytucje finansowe	51 218,8
1.2. Instytucje rządowe i samorządowe	79 184,2	1.2. Instytucje rządowe szczebla centralnego	35 952,3
1.3. Pozostałe sektory krajowe	928 053,3	1.3. Pozostałe sektory krajowe	944 129,4
2. Dłużne papiery wartościowe wyemitowane przez podmioty krajowe kraj	318 020,8	1.3.1. Depozyty bieżące	474 966,7
2.1. Monetarne instytucje finansowe	95 231,6	1.3.2. Depozyty terminowe	458 780,6
2.2. Instytucje rządowe i samorządowe	198 488,3	1.3.3. Depozyty z terminem wypowiedzenia	1,8
2.3. Pozostałe sektory krajowe	24 300,9	1.3.4. Operacje z przyrzeczeniem odkupu	10 350,3
3. Jednostki uczestnictwa w funduszach rynku pieniężnego	0,0	2. Jednostki uczestnictwa w funduszach rynku pieniężnego	0,0
4. Papiery z prawem do kapitału i udziały	15 083,7	3. Emisja dłużnych papierów wartościowych	55 255,8
4.1. Monetarne instytucje finansowe	3 761,0	4. Kapitał i rezerwy	227 160,5
4.2. Pozostałe sektory krajowe	11 322,7	4.1. Fundusze podstawowe	146 314,6
5. Aktywa zagraniczne	78 435,4	4.2. Fundusze uzupełniające	5 038,1
6. Aktywa trwałe	48 157,4	4.3. Rezerwy	72 807,8
7. Pozostałe aktywa	54 999,8	5. Pasywa zagraniczne	237 585,3
		6. Pozostałe pasywa	70 704,4
Aktywa ogółem	**1 622 006,5**	**Pasywa ogółem**	**1 622 006,5**

Źródło: NBP.

Zestawienie bilansów banków komercyjnych i bilansu banku centralnego w postaci skonsolidowanego bilansu monetarnych instytucji finansowych pozwala ukazać całość obiegu pieniężnego w gospodarce (tabela 5.5). Podaż pieniądza znajduje się po stronie pasywów. Po zsumowaniu wartości pieniądza gotówkowego w obiegu (poza kasami banków) i depozytów bieżących otrzymujemy agregat M1, a po dodaniu pozostałych depozytów (oszczędnościowych, terminowych do 2 lat i z terminem wypowiedzenia do 3 miesięcy) uzyskujemy agregat M2. Uwzględniając następnie dokonane przez banki operacje z przyrzeczeniem odkupu (*repo*),

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jednostki uczestnictwa w funduszach rynku pieniężnego (wynoszące zero) oraz dłużne papiery wartościowe z terminem zapadalności do 2 lat, otrzymujemy agregat M3. Bilans skonsolidowany ukazuje też źródła kreacji pieniądza w obiegu – są nimi kredyty i pożyczki udzielone podmiotom krajowym oraz aktywa zagraniczne netto (po potrąceniu pasywów zagranicznych) krajowego systemu bankowego.

Tabela 5.5. Bilans skonsolidowany monetarnych instytucji finansowych – stan na dzień 31.12.2014 roku (dane w mln PLN)

AKTYWA		PASYWA	
1. Kredyty i pożyczki oraz inne należności od podmiotów krajowych	1 007 260,0	1. Pieniądz gotówkowy w obiegu (poza kasami banków)	130 029,8
1.1. Sektor instytucji rządowych i samorządowych	79 184,2	2. Depozyty i inne zobowiązania wobec instytucji szczebla centralnego	55 384,2
1.2. Pozostałe sektory krajowe	928 075,8	3. Depozyty i inne zobowiązania wobec pozostałych sektorów krajowych	945 385,4
2. Dłużne papiery wartościowe wyemitowane przez podmioty krajowe	222 789,2	3.3.1. Depozyty bieżące	476 252,7
2.1. Sektor instytucji rządowych i samorządowych	198 488,3	3.3.2. Depozyty terminowe	458 780,6
2.2. Pozostałe sektory krajowe	24 300,9	3.3.3. Depozyty z terminem wypowiedzenia	1,8
3. Papiery z prawem do kapitału i udziały wyemitowane przez pozostałe sektory krajowe	11 376,5	3.3.4. Operacje z przyrzeczeniem odkupu	10 350,3
4. Aktywa zagraniczne	437 357,7	4. Jednostki uczestnictwa w funduszach rynku pieniężnego	0,0
5. Aktywa trwałe	50 382,0	5. Emisja dłużnych papierów wartościowych	44 685,6
6. Pozostałe aktywa	42 609,9	6. Kapitał i rezerwy	236 250,6
		7. Pasywa zagraniczne	262 411,1
		8. Pozostałe pasywa	95 178,8
		9. Saldo operacji wzajemnych monetarnych instytucji finansowych	449,8
Aktywa ogółem	**1 771 775,3**	**Pasywa ogółem**	**1 771 775,3**

Źródło: NBP.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.1.5. Dwuwalutowość

Mówiąc o podaży pieniądza, ma się najczęściej na myśli pieniądz krajowy emitowany przez narodowy bank centralny. Lecz w obiegu mogą znajdować się równolegle także pieniądze emitowane przez banki centralne innych państw. Choć tylko pieniądz krajowy pozostaje jedynym prawnym środkiem płatniczym, w praktyce uczestnicy rynku mogą posługiwać się także inną walutą (lub wieloma różnymi walutami). Sytuacja taka, nazywana dwuwalutowością (wielowalutowością) obiegu pieniężnego, może mieć miejsce w warunkach spadku zaufania do waluty krajowej, spowodowanego wysoką inflacją, niedoborami towarowymi w warunkach kontrolowanych cen lub obawami przed nieekwiwalentną wymianą pieniądza przez władze. Powodem może być także wzrost ryzyka politycznego. Występuje wówczas zjawisko ucieczki od pieniądza krajowego, tracącego wówczas szybko na wartości i sile nabywczej. W jego miejsce do obrotu pieniężnego wchodzą spontanicznie inne waluty, z reguły swobodnie wymienialne, o stabilnej wartości i cieszące się zaufaniem. Można więc powiedzieć, że mamy w takiej sytuacji do czynienia z prawem Kopernika-Greshama *à rebours*, gdzie pieniądz lepszy wypiera pieniądz gorszy[9].

W jaki sposób waluty obce stają się częścią obiegu pieniężnego w danym kraju? Czy należy je traktować jako część podaży pieniądza? Waluty obce napływają do krajowej gospodarki z zagranicy. Gdy ludzie gromadzą je w postaci gotówki, nie wlicza się ich do obiegu pieniężnego, ponieważ nie są wówczas wykorzystywane jako środki płatnicze, lecz stanowią raczej rodzaj aktywów finansowych. Gdy z kolei trafiają na rachunki w bankach jako depozyty walutowe, mogą być przeznaczane przez banki do udzielania kredytów (także walutowych) i dlatego stają się częścią obiegu pieniężnego.

W Polsce zjawisko dwuwalutowości występowało powszechnie w okresie wysokiej inflacji i masowych niedoborów towarowych u schyłku gospodarki centralnie planowanej pod koniec lat 80. Sprzyjała mu także stopniowa liberalizacja przepisów walutowych, w tym zwłaszcza możliwość prowadzenia rachunków bankowych w walutach obcych. Na przełomie 1989 i 1990 roku udział depozytów walutowych w całości podaży pieniądza (w zależności od przelicznika) sięgał 60–70% (Rosati

[9] Prawo Kopernika-Greshama głosi, że jeśli w obiegu funkcjonują jednocześnie dwa rodzaje pieniądza i jeden z nich jest postrzegany jako lepszy (np. o wyższej zawartości kruszcu), to wówczas gromadzony (tezauryzowany) będzie „lepszy" pieniądz, a w obiegu pozostanie głównie ten „gorszy". W rezultacie gorszy pieniądz wypiera lepszy.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

1998, s. 115). W strukturze walutowej dominowały dolar amerykański i marka zachodnioniemiecka – dwie waluty tradycyjnie cieszące się wśród Polaków największym zaufaniem. Reformy gospodarcze podjęte w 1990 roku i osiągnięta dzięki nim stabilizacja makroekonomiczna doprowadziły do stopniowej odbudowy zaufania wobec polskiego złotego, a w efekcie udział walut obcych w podaży pieniądza zaczął szybko spadać (także w wyniku silnej realnej aprecjacji złotego w latach 90.). Wykres 5.3 pozwala zauważyć, że proces ten uległ jednak zahamowaniu w czasie kryzysu finansowego z lat 2008–2009. Od tego momentu udział depozytów walutowych w całości agregatu M3 utrzymuje się na poziomie około 8%.

Wykres 5.3. Udział depozytów walutowych w podaży agregatu M3 w okresie od grudnia 1996 roku do grudnia 2015 (dane w %)

Źródło: opracowanie własne na podstawie danych NBP.

Wielowalutowość występuje także po stronie źródeł podaży pieniądza. Przyjmuje ona postać głównie kredytów i pożyczek udzielanych w walutach obcych. Operacje takie cieszą się szczególnym zainteresowaniem wówczas, gdy w warunkach pełnej swobody transgranicznego przepływu kapitału zagraniczne stopy procentowe są znacząco niższe od stóp krajowych. Koszt obsługi kredytu walutowego jest wówczas wyraźnie niższy niż koszt obsługi kredytu złotowego. Jednocześnie zaciągnięcie kredytu walutowego przez osoby lub firmy, których dochody są denominowane w walucie krajowej, prowadzi do powstania ryzyka kursowego. Nie zawsze jest ono jednak w pełni uwzględniane w decyzjach o zaciągnięciu kredytu, o czym świadczy masowy wzrost kredytów walutowych w Polsce w okresie poprzedzającym kryzys

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

finansowy lat 2008–2009. Wykres 5.4 pokazuje, że w latach 1996–2003 udział tych kredytów w ogólnej sumie należności w polskim sektorze bankowym wzrósł z 13% do 32%. Kolejny skok nastąpił w drugiej połowie 2008 roku i był spowodowany silnym osłabieniem kursu złotego wobec głównych walut obcych w związku z kryzysem finansowym. W ostatnich latach udział kredytów walutowych spadł poniżej 30%, głównie na skutek ograniczeń nałożonych na emisję nowych kredytów przez władze nadzoru bankowego. Ogólnie rzecz biorąc, mamy jednak w polskim systemie bankowym do czynienia z wyraźnym niedopasowaniem wielkości aktywów i pasywów walutowych – w końcu 2015 roku udział kredytów w walutach obcych w całkowitej należności sektora bankowego był ponad trzykrotnie wyższy niż udział depozytów walutowych w podaży pieniądza ogółem[10].

Wykres 5.4. Udział kredytów walutowych w całkowitej należności monetarnych instytucji finansowych w Polsce w okresie od grudnia 1996 do grudnia 2015 roku (dane w %)

Źródło: opracowanie własne na podstawie danych NBP.

Wielowalutowość w oczywisty sposób komplikuje prowadzenie polityki pieniężnej i w pewnych sytuacjach może zagrażać stabilności sektora finansowego. Duży udział depozytów walutowych oraz walut obcych w obiegu gotówkowym jest związany ze spadkiem popytu na pieniądz krajowy emitowany przez bank centralny.

[10] Tak duże otwarte pozycje walutowe po stronie kredytów wymagają od banków zabezpieczania się pożyczkami walutowymi i transakcjami swapowymi.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ogranicza to zdolność banku centralnego do regulowania ogólnej podaży pieniądza w obiegu, a także zmniejsza dochód z emisji. Duży udział kredytów walutowych powoduje wzrost ryzyka kursowego, które może się zmaterializować w okresie osłabienia waluty krajowej. Wywołany tym wzrost kosztów obsługi długu jest w stanie doprowadzić do trudności z obsługą kredytów i masowych bankructw, co skutkuje recesją i podważa stabilność sektora bankowego. Podobne efekty może przynieść odnotowany w bilansach banków brak równowagi między wielkością pasywów i aktywów w walutach obcych. W Polsce od 2005 roku występowała rosnąca nierównowaga między depozytami a kredytami walutowymi – w latach 2009–2011 suma udzielonych przez banki kredytów walutowych okazała się ponad czterokrotnie większa niż suma zgromadzonych depozytów. Banki zmuszone były finansować tę różnicę przy pomocy pożyczek zaciąganych na zagranicznych rynkach finansowych i operacji swapowych (*foreign exchange swap*), co zwiększało bieżące koszty.

5.1.6. Rezerwy bankowe

W systemie bankowym, funkcjonującym na zasadzie tzw. ułamkowej rezerwy (*fractional-reserve banking*), bank komercyjny utrzymuje rezerwy płynnych środków finansowych, stanowiące tylko niewielką część całej sumy depozytów złożonych w tym banku przez klientów. Pozostała część depozytów jest przez bank wypożyczana innym klientom w postaci kredytów lub przeznaczana na zakup innych aktywów.

Celem utrzymywania rezerw bankowych jest zapewnienie zdolności banku do dokonywania bieżących wypłat z prowadzonych rachunków. Ponieważ w normalnych warunkach nie zdarza się, aby wszyscy klienci chcieli wypłacić swoje depozyty w tym samym czasie, do obsługi bieżących wypłat wystarcza posiadanie płynnych środków, które stanowią zaledwie niewielką część zgromadzonych depozytów. Wielkość rezerw w relacji do posiadanych depozytów określa stopa rezerwy obowiązkowej, ustalana przez bank centralny. W niektórych krajach nie stosuje się tego zabezpieczenia, lecz banki same gromadzą rezerwy w takiej wysokości, którą uważają za konieczną, aby zapewnić niezakłócone funkcjonowanie bieżących wypłat (na przykład w Wielkiej Brytanii). Rezerwy bankowe są na ogół utrzymywane częściowo w postaci gotówkowej jako gotówka w kasach banku (tzw. pogotowie kasowe), a częściowo na rachunku w banku centralnym. Bank może utrzymywać rezerwy większe niż rezerwa obowiązkowa, a różnicę tych wartości nazywa się rezerwą dobrowolną. Gdy bank posiada rezerwy dobrowolne, może

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

udzielać dodatkowych kredytów, ponieważ ma środki na odprowadzenie wymaganej rezerwy obowiązkowej od depozytu utworzonego przy udzielaniu kredytu. Nowe depozyty zwiększają podaż szerszych agregatów pieniężnych dzięki mechanizmowi mnożnika, którego wielkość jest odwrotnie proporcjonalna do wielkości stopy rezerw (zob. następny punkt)[11].

Stopa rezerwy obowiązkowej w krajach rozwiniętych o ustabilizowanym systemie bankowym nie przekracza z reguły kilku procent. Na przykład w strefie euro wynosi ona 2%, a w Polsce 3,5% (choć we wczesnych latach 90. sięgała nawet 20%). Ogólnie rzecz biorąc, im niższa inflacja i im bardziej stabilna gospodarka, tym niższa może być stopa rezerwy obowiązkowej. Różnicuje się ją często w zależności od rodzaju depozytu jakiego dotyczy – mniejsza stawka rezerwy obowiązuje przy depozytach terminowych, a większa przy depozytach bieżących (*à vista*).

W większości krajów rezerwy bankowe są oprocentowane. Władza gospodarcza stara się w ten sposób zmniejszyć koszty związane z koniecznością utrzymywania rezerw obowiązkowych przez banki. Stopa oprocentowania pozostaje jednak na ogół dość niska. Na przykład w Polsce wysokość oprocentowania rezerw obowiązkowych wynosiła przez wiele lat 0,9 stopy redyskonta weksli, z kolei od października 2014 roku wynosi 0,9 stopy referencyjnej ustalanej przez NBP.

Zupełnie czymś innym są rezerwy walutowe (lub dewizowe). Kraje utrzymują rezerwy zagranicznych walut, aby zapewnić niezakłócony przebieg procesu międzynarodowych transakcji zakupu dóbr i usług (import) oraz zakupu aktywów zagranicznych. Niekiedy rezerwy walutowe służą utrzymaniu stałego kursu waluty krajowej, mogą też być czynnikiem podnoszącym wiarygodność i bezpieczeństwo ekonomiczne kraju.

Pożądana wysokość rezerw walutowych zależy od wielu czynników, w tym od wielkości obrotów handlu zagranicznego, obowiązującego systemu kursu walutowego i od ogólnej stabilności gospodarki. Ogólnie więc im wyższe będą rezerwy, tym większe stanie się bezpieczeństwo ekonomiczne kraju, choć jednocześnie wzrosną też koszty związane z ich utrzymywaniem. Ponieważ rezerwy utrzymywane są głównie w postaci zagranicznych rządowych papierów wartościowych

[11] System ułamkowej rezerwy wiąże się z ryzykiem utraty płynności przez bank w sytuacji paniki na rynkach. Dlatego też, zdaniem niektórych polityków i ekonomistów, należy z niego zrezygnować i zastąpić go systemem stuprocentowej rezerwy (zob. np. Rothbard 1995). W takiej sytuacji bank musiałby jednak inwestować środki wyłącznie w superpłynne aktywa, tak aby w każdej chwili mieć możliwość dokonania wypłaty depozytów. Oczywiście, w takim systemie udzielenie jakiegokolwiek kredytu byłoby praktycznie niemożliwe.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

emitowanych przez najbardziej wiarygodne kraje, ich dochodowość jest z reguły niewielka. Powszechnie uznaje się, że minimalną bezpieczną wielkość rezerw walutowych zapewnia utrzymywanie ich na poziomie odpowiadającym wartości 3–4 miesięcy importu dóbr i usług.

Tabela 5.6. Oficjalne aktywa rezerwowe NBP (dane w mln EUR)

Rok	Wartość	Rok	Wartość	Rok	Wartość
1993	3 819	2001	30 067	2009	55 221
1994	4 940	2002	28 450	2010	69 991
1995	11 679	2003	27 094	2011	75 722
1996	14 668	2004	26 966	2012	82 577
1997	19 376	2005	35 970	2013	77 144
1998	24 209	2006	36 833	2014	82 645
1999	27 179	2007	44 693	2015	86 894
2000	29 524	2008	44 139		

Źródło: NBP.

Tabela 5.7. Struktura rezerw walutowych NBP – stan na dzień 31.12.2015 roku

Rodzaj aktywów	Mln EUR	Udział w %
Złoto monetarne	3 218	3,70
Specjalne Prawa Ciągnienia (*Special Drawing Rights* – SDR)	1 202	1,38
Transza rezerwowa w MFW	606	0,70
Papiery wartościowe nominowane w walutach obcych i inne należności w walutach wymienialnych	81 868	94,22
RAZEM	86 894	100,00

Źródło: opracowanie własne na podstawie danych NBP.

Tabela 5.6 pozwala zauważyć, że poziom rezerw walutowych w Polsce w latach 1993–2015 zwiększył się niemal 23-krotnie. Szczególnie szybki wzrost nastąpił w latach 1993–1999, kiedy do Polski zaczął napływać kapitał zagraniczny, a NBP, broniąc kursu złotego przed niepożądaną aprecjacją, musiał interweniować, skupując waluty obce. W kolejnych latach tempo przyrostu rezerw walutowych osłabło – na przestrzeni piętnastu lat między rokiem 2000 a 2015 stan rezerw wzrósł zaledwie 3-krotnie. W 2014 roku oficjalne aktywa rezerwowe stanowiły ponad 49% rocznej wartości importu dóbr i usług, osiągając równowartość niemal półrocznego importu. Z kolei tabela 5.7 pokazuje, że głównym składnikiem rezerw

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

walutowych Polski były zagraniczne papiery wartościowe, które stanowiły ponad 94% wszystkich oficjalnych aktywów rezerwowych. Przy ustalaniu struktury portfela tych aktywów kluczowe znaczenie odgrywają kryteria bezpieczeństwa i płynności, dlatego przeważający udział w tych aktywach miały skarbowe papiery wartościowe emitowane przez rząd USA i rząd Republiki Federalnej Niemiec. Złoto monetarne stanowiło niecałe 4% wartości rezerw, a udziały Polski w Międzynarodowym Funduszu Walutowym niewiele ponad 2%.

5.2. Mnożnik pieniężny

5.2.1. Definicja mnożnika

Pisząc o rodzajach pieniądza, dokonaliśmy ważnego rozróżnienia miedzy pieniądzem rezerwowym, emitowanym przez bank centralny, i pieniądzem depozytowym, emitowanym przez banki komercyjne. Pieniądz rezerwowy – czyli baza monetarna – składa się z gotówki znajdującej się w obiegu i rezerw banków komercyjnych w banku centralnym. Z kolei całkowita podaż pieniądza w obiegu (agregat M3) obejmuje zarówno pieniądz gotówkowy, jak i depozyty bankowe, przy czym te ostatnie stanowią 85–90% obiegu pieniężnego. Bank centralny, regulując ilość pieniądza w obiegu, bezpośrednio kontroluje jedynie podaż gotówki, natomiast na podaż depozytów może wpływać tylko w sposób pośredni. Jak się dalej przekonamy, w normalnych warunkach gospodarczych istnieje określona, dość stabilna zależność wielkości tworzonych depozytów od zasobu gotówki znajdującej się w obiegu czy też szerzej – od zasobu pieniądza rezerwowego. Zależność tę ukazuje mnożnik pieniężny (*money multiplier*), który pozwala ustalić, ile jednostek szerszych agregatów M1, M2, M3 powstaje z jednostki bazy monetarnej M0 (gotówka plus rezerwy).

Znajomość wielkości mnożników pieniężnych (M1/M0, M2/M0, M3/M0) jest istotna z punktu widzenia skutecznego prowadzenia polityki pieniężnej. Ponieważ bank centralny bezpośrednio reguluje jedynie wielkość bazy monetarnej, to, aby kształtować wielkość szerszych agregatów pieniężnych (na przykład M2 lub M3) stosownie do wielkości popytu na pieniądz, bank centralny musi wiedzieć, o jaką kwotę zwiększy się podaż szerszych agregatów pieniężnych w wyniku wzrostu podaży pieniądza rezerwowego o określoną kwotę, na przykład o 1 mld PLN. Jeśli bank centralny jest w stanie w miarę precyzyjnie oszacować istniejącą zależność

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

między M0 i M2, będzie w stanie przewidzieć, ile jednostek M2 powstanie w wyniku zwiększenia podaży M0 o jednostkę. W ten sposób bank centralny może kontrolować całkowitą podaży pieniądza w gospodarce.

Jak wygląda więc mechanizm „powstawania" szerszych agregatów pieniężnych i jaki jest związek między nimi a podażą pieniądza rezerwowego? Zacznijmy od przypomnienia, że zawsze, gdy banki komercyjne udzielają kredytu, tworzą w ten sposób nowy depozyt, od którego muszą odprowadzić wymaganą rezerwę na rachunek w banku centralnym. Rezerwy (obowiązkowe, a także dobrowolne) mają postać płynnych środków na rachunku w banku centralnym. Kontrolując dostępność, cenę oraz zasób płynnych rezerw (bazy monetarnej) w gospodarce, bank centralny może więc wpływać na wielkość kredytu i depozytów, a w konsekwencji też na wielkość podaży pieniądza. W ramce 5.1 przedstawiono formalny związek między pieniądzem rezerwowym a szerokim agregatem pieniężnym M2, który pozostaje istotą mnożnika pieniężnego.

Ramka 5.1. Arytmetyka mnożnika pieniężnego

Zależność między rezerwami (*R*) a depozytami (*D*) określona jest przez stopę rezerw. Jeśli stopa rezerwy obowiązkowej wynosi $q_o = R_o / D$, a stopa rezerwy dobrowolnej wynosi $q_o = R_d / D$, to stosunek całej rezerwy do depozytów wynosi $q = q_o + q_d = \left(R_o + R_d\right) / D$. Rezerwy nie mogą być mniejsze niż iloczyn stopy rezerwy obowiązkowej i wielkości depozytów:

$$R \geq q_o D. \tag{5.9}$$

Oznaczmy bazę monetarną jako *M0*, szeroki agregat pieniężny (*broad money*) jako *M*2, a mnożnik szerokiego agregatu jako $M_2 = M2 / M0$. Z definicji agregatów pieniężnych otrzymujemy[12]:

$$M0 = C + R \tag{5.10}$$

$$M2 = C + D. \tag{5.11}$$

[12] Zob. poprzedni punkt.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Następnie oznaczmy udział gotówki w $M2$ jako $c = C/M2$, czyli $C = cM2$. Wówczas:

$$M0 = C + R = cM2 + q_o D \tag{5.12}$$

$$M2 = C + D = cM2 + D. \tag{5.13}$$

Ze wzoru (5.13) uzyskujemy $D = M2(1 - c)$. Podstawiając do równania (5.12), otrzymujemy:

$$M0 = M2\ [c + q_o(1 - c)]. \tag{5.14}$$

Dzieląc obustronnie przez $M0$ i porządkując, uzyskujemy wzór na mnożnik pieniężny M_2:

$$M_2 = M2/M0 = \frac{1}{c + q_0(1 - c)}. \tag{5.15}$$

Wielkość mnożnika zależy tylko od wartości c i q_o. Im wyższa wartość c dla danego q_o, czyli im większa część $M2$ przechowywana jest w gotówce, tym niższa wartość mnożnika. Dla danej stopy rezerwy obowiązkowej mnożnik osiąga maksimum, gdy wartość $c = 0$, a więc gdy nie ma gotówki w obiegu, a całość M2 składa się wyłącznie z depozytów, a wówczas $M_2 = \left(\frac{1}{q_o}\right)$. Mnożnik osiąga minimum, gdy $c = 1$, czyli całość M2 składa się wyłącznie z gotówki, a wówczas $M_2 = 1$. W skrajnym przypadku, dla $c = 0$ i $q_o = 0$, wartość mnożnika rośnie do nieskończoności (Burda i Wyplosz 2000).

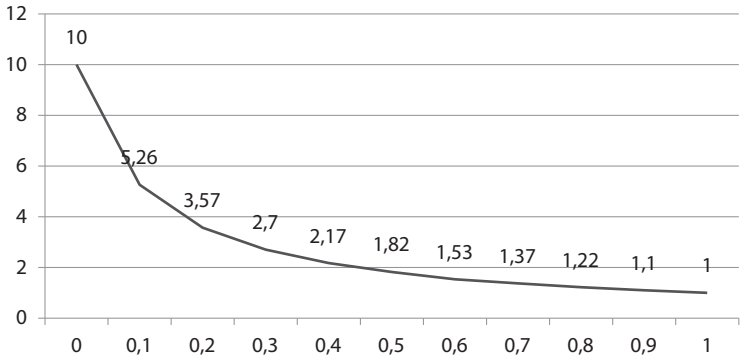
Wzór (5.15) pokazuje, że wielkość mnożnika zależy od parametrów c i q_o, z czego bank centralny bezpośrednio ustala jedynie wielkość stopy rezerwy obowiązkowej. Z kolei o wielkości c decydują uczestnicy rynku. Jeśli zechcą oni utrzymywać większą część swoich zasobów pieniężnych w postaci gotówki, a nie w formie wkładów na rachunkach bankowych, wielkość depozytów spadnie, zasób gotówki w obiegu wzrośnie, a banki stracą część posiadanych rezerw. Zmniejszy się też wówczas wielkość mnożnika, co oznacza, że wzrost udziału gotówki kosztem rezerw obniża zdolność banków do tworzenia kredytów i depozytów. W związku

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

z tym współczynnik *c* jest nazywany także współczynnikiem „wyciekania" pieniądza poza system bankowy.

Wnioski z powyższej analizy są następujące. Po pierwsze, wielkość mnożnika jest tym mniejsza, im większy jest współczynnik *c* udziału gotówki w M1, M2, M3, czyli im więcej pieniędzy „wycieka" poza system bankowy. Należy przy tym zauważyć, że mnożnik jest bardzo wrażliwy na zmiany tego współczynnika w przedziale małych wartości *c* – wykres 5.5 pokazuje, że gdy wartość *c* rośnie od zera do 10%, wartość mnożnika spada niemal o połowę. Po drugie, bank centralny nie ma wpływu na wielkość *c*, a zatem tylko pośrednio i w sposób niepełny oddziałuje na wielkość szerszych agregatów pieniężnych M1, M2, M3. Po trzecie, bank centralny ustala wprawdzie stopę rezerw obowiązkowych q_o, lecz nie ma wpływu na wielkość rezerw dobrowolnych R_d utrzymywanych przez banki komercyjne w banku centralnym. Im większe stają się rezerwy dobrowolne, tym większa będzie suma $(q_o + q_d)$ przy mniejszej wartości mnożnika.

Wykres 5.5. Wielkość mnożnika pieniężnego w zależności od wielkości współczynnika „wyciekania" *c* (dla *q* = 0,1)



Źródło: opracowanie własne.

5.2.2. Wielkość mnożnika bazy monetarnej

Wielkość mnożnika można wyliczyć dopiero wówczas, gdy mamy do dyspozycji dane statystyczne dotyczące obiegu pieniężnego. Bank centralny stara się jednak zawsze oszacować wielkości mnożników bazy monetarnej, aby wiedzieć, o ile wzrośnie całkowita podaż pieniądza na skutek zwiększenia bazy monetarnej o określoną kwotę. Dokonując tych obliczeń, bank centralny musi pamiętać, że

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

może on wprawdzie kontrolować cały agregat M0, lecz nie posiada wpływu na jego wewnętrzną strukturę, ponieważ udział gotówki *C* w obiegu zależy od preferencji ludzi, a o wielkości rezerw *R* decyduje skłonność banków do utrzymywania rezerw ponad poziom rezerwy obowiązkowej. Tabela 5.8 przedstawia funkcjonowanie mnożnika na przykładzie liczbowym, przy założeniu, że stopa rezerwy obowiązkowej wynosi 10%, a współczynnik „wycieku" *c* jest równy zeru.

Tabela 5.8. Przykład działania mechanizmu mnożnika (założenia: początkowy depozyt = 100, stopa rezerwy obowiązkowej q^o = 10%, współczynnik „wycieku" *c* = 0)

Krok 1		Krok 2	
Aktywa	Pasywa	Aktywa	Pasywa
10 – rezerwa 90 – kredyt (klient B) 100 – razem aktywa	100 – depozyt (klient A) 100 – razem pasywa	19 – rezerwa 90 – kredyt (klient B) 81 – kredyt (klient C) 190 – razem aktywa	100 – depozyt (klient A) 90 – depozyt (klient B) 190 – razem pasywa
Krok 3		**Krok 4**	
Aktywa	Pasywa	Aktywa	Pasywa
27,1 – rezerwa 90 – kredyt (klient B) 81 – kredyt (klient C) 72,9 – kredyt (klient D) 271 – razem aktywa	100 – depozyt (klient A) 90 – depozyt (klient B) 81 – depozyt (klient C) 271 – razem pasywa	34,39 – rezerwa 90 – kredyt (klient B) 81 – kredyt (klient C) 72,9 – kredyt (klient D) 65,61 – kredyt (klient E) 343,9 – razem aktywa	100 – depozyt (klient A) 90 – depozyt (klient B) 81 – depozyt (klient C) 72,9 – depozyt (klient D) 343,9 – razem pasywa

Źródło: opracowanie własne.

Załóżmy, ze pewien bank komercyjny ma wielu klientów i że klient A deponuje na założonym tam rachunku kwotę 100 PLN. Bank musi od tego depozytu odprowadzić 10% – czyli 10 PLN – w postaci rezerwy. Pozostałe 90 PLN bank przeznacza na udzielenie kredytu klientowi B. Stan bilansu banku po wykonaniu tych operacji opisuje krok 1 w tabeli 5.8. Po stronie aktywów mamy dwie pozycje – rezerwę obowiązkową na rachunku w banku centralnym oraz kwotę kredytu udzielonego klientowi B, po stronie pasywów figuruje zaś depozyt klienta A. W następnym kroku, na skutek decyzji o udzieleniu kredytu klientowi B, na koncie tego klienta pojawia się depozyt w wysokości 90 PLN. Bank musi zatem odprowadzić rezerwę obowiązkową od tego depozytu w wysokości 9 PLN. Pozostałe 81 PLN bank przeznacza na udzielenie kredytu klientowi C. Stan bilansu banku po wykonaniu tych operacji ilustruje krok 2. W następnych krokach sytuacja się powtarza – bank udziela kolejnych kredytów, tworząc depozyty i odprowadzając rezerwę obowiązkową.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W tabeli 5.8 przedstawiono jeszcze krok 3 i 4 w celu zilustrowania powtarzalności tych decyzji. W wyniku dokonanych operacji suma aktywów banku wzrosła ze 100 PLN na początku do 343,9 PLN po przeprowadzeniu czterech kolejnych kroków i – podobnie – suma depozytów utworzonych w banku dzięki początkowej wpłacie gotówkowej 100 PLN również osiągnęła wartość 343,9 PLN. Jak można jednak zauważyć, kolejne przyrosty kredytów i depozytów są coraz mniejsze – wynika to z tego, że na każdym etapie bank musi odprowadzić część posiadanych środków gotówkowych jako rezerwę obowiązkową, co każdorazowo zmniejsza kwotę udzielanego kredytu. Proces ten będzie trwał aż do momentu, kiedy całość początkowej wpłaty klienta A zostanie przekazana na rachunek rezerwy obowiązkowej.

Ile wyniesie łączna kwota stworzonych depozytów, a więc jaka jest maksymalna wartość mnożnika, przy założeniu, że proces udzielania kredytów i tworzenia depozytów będzie kontynuowany aż do całkowitego wyczerpania posiadanego zasobu pieniądza rezerwowego?

Można to łatwo obliczyć, wykorzystując wzór na sumę ciągu. Zakładając nadal, że $c = 0$, a D jest początkowym depozytem, suma ta ma postać ogólną:

$$\Sigma(D) = D(1-q) + D(1-q)(1-q) + D(1-q)(1-q)(1-q) + \ldots = D(1/q). \qquad (5.16)$$

Suma depozytów stworzonych w procesie mnożnikowym wyniesie zatem:

$$100 + 90 + 81 + 72{,}9 + 65{,}61 + \ldots = D(1/q) = 1000.$$

W związku z tym wartość mnożnika w naszym przykładzie osiągnie poziom:

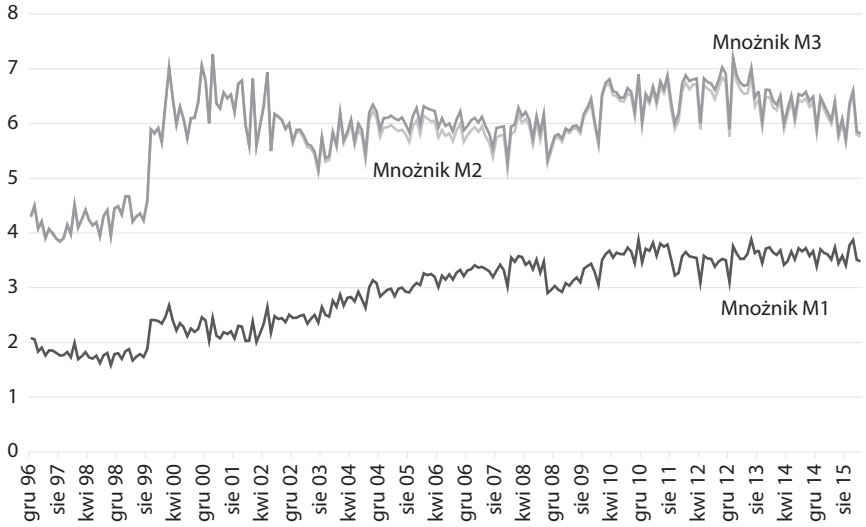
$$M_2 = \frac{D\left(\frac{1}{q}\right)}{D} = \frac{1}{q} = \frac{1}{0{,}1} = 10.$$

Interpretując otrzymane wyniki, należy pamiętać o przyjętych założeniach. Po pierwsze, w rzeczywistości współczynnik „wyciekania" jest z reguły większy od zera, co powoduje spadek mnożnika. Im większą część depozytów ludzie wypłacają w gotówce, tym więcej pieniędzy „wycieka" z systemu bankowego, tym mniejsze stają się depozyty i tym mniejszy będzie mnożnik. Na przykład dla $c = 0{,}5$, suma wszystkich stworzonych depozytów wynosi już tylko 181,18, a wartość mnożnika spada do 1,82 (zob. wykres 5.5). Po drugie, banki rzadko przeznaczają wszystkie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wolne środki na nowe kredyty – z reguły kupują także inne aktywa finansowe, na przykład obligacje skarbowe lub korporacyjne. Im mniejszy więc będzie udział kredytów w aktywach, tym niższa stanie się wartość mnożnika. Po trzecie, część posiadanych przez ludzi pieniędzy trafia do niemonetarnych instytucji finansowych pozbawionych możliwości kreowania pieniądza, takich jak kasy oszczędnościowe, fundusze inwestycyjne, emerytalne i ubezpieczeniowe. Muszą one utrzymywać pełne pokrycie powierzonych im środków, co oznacza *de facto* wymóg utrzymania 100% rezerwy obowiązkowej[13]. Wreszcie banki mogą utrzymywać rezerwy na wyższym poziomie niż wymagany stopą rezerwy obowiązkowej. Na przykład w sytuacjach kryzysowych banki niechętnie udzielają kredytów i często wolą lokować wolne środki na rachunku rezerwy w banku centralnym (jako rezerwy dobrowolne), czekając na lepsze czasy.

Wykres 5.6. Mnożniki pieniężne w Polsce w okresie od grudnia 1996 roku do grudnia 2015 roku



Uwaga: mnożniki wyliczone jako stosunek M1 i M3 do bazy monetarnej.

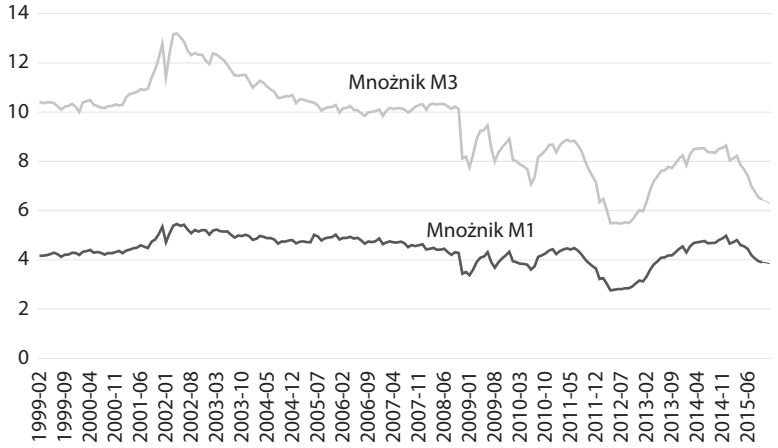
Źródło: opracowanie własne na podstawie danych NBP.

[13] Szczególnym rodzajem niemonetarnej instytucji finansowej są tak zwane parabanki, czyli firmy udzielające na rynku pożyczek bez wymaganego prawem zezwolenia na działalność bankową.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykresy 5.6 i 5.7 prezentują wartości mnożników M_1, M_2, M_3 dla Polski oraz strefy euro w latach 1996–2015. W Polsce występuje tendencja do wzrostu wartości mnożników, choć tempo zmian jest niewielkie. Odpowiednie mnożniki w przypadku strefy euro w okresie przed kryzysem finansowym były przeciętnie dwukrotnie wyższe niż w Polsce, co potwierdza wcześniejsze spostrzeżenie, że w krajach gospodarczo bardziej rozwiniętych wartości mnożników są z reguły wyższe. Po 2007 roku wartość mnożników w strefie euro wyraźnie spadła, co miało związek z kryzysem finansowym i wzrostem awersji do ryzyka w sektorze bankowym.

Wykres 5.7. Mnożniki pieniężne w strefie euro w okresie od lutego 1999 roku do grudnia 2015 roku



Uwaga: mnożniki wyliczone jako stosunek M1 i M3 do bazy monetarnej. Ze względu na brak niektórych danych miesięcznych w przypadku 2015 roku (luty, maj, sierpień, listopad) brakujące obserwacje oszacowano metodą interpolacji.

Źródło: opracowanie własne na podstawie danych EBC.

5.3. Cele polityki pieniężnej

5.3.1. Tradycyjne cele polityki pieniężnej

W najbardziej ogólnym ujęciu celem polityki pieniężnej jest dostarczanie środków płatniczych do gospodarki, tak aby pokryć potrzeby uczestników obrotu gospodarczego związane z zakupem dóbr, usług i aktywów. Taka definicja nie jest jednak dostatecznie precyzyjna i jej wykorzystanie w praktyce mogłoby napotykać

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

trudności. Na wstępie niniejszego rozdziału zauważyliśmy, że polityka pieniężna polega na regulowaniu ilości pieniądza w obiegu w taki sposób, aby nie było go ani za dużo, ani za mało. Gdy podaż pieniądza jest zbyt wysoka w stosunku do popytu, wystąpi tendencja do inflacyjnego wzrostu cen. Kiedy podaż pieniądza jest zaś zbyt niska, pojawia się tendencja do deflacyjnego spadku cen i/lub spadku produkcji oraz zatrudnienia. Można więc stwierdzić, że celem polityki pieniężnej powinno być dostarczanie pieniądza w takiej ilości, która pozwoli zapewnić stabilność cen. Wówczas podaż pieniądza będzie równa jego popytowi[14].

Takie podejście stało się bardzo popularne w okresie ostatnich 30 lat. Wcześniej w większości krajów stawiano przed polityką pieniężną także inne cele. W okresie przed I wojną światową głównym zadaniem polityki pieniężnej była z reguły obrona mocnego pieniądza, polegająca na utrzymywaniu wysokiego kursu waluty krajowej lub stałego parytetu wobec złota. Z kolei po wielkim kryzysie lat 1929–1933, pod wpływem upowszechniania się keynesowskich idei polityki stabilizacyjnej, przed polityką pieniężną postawiono dodatkowo zadanie dostarczania taniego kredytu dla gospodarki w celu niwelowania skutków cyklu koniunkturalnego, pobudzania wzrostu gospodarczego i utrzymywania pełnego zatrudnienia. W miarę upływu czasu okazało się jednak, że mnogość celów – często alternatywnych – utrudnia prowadzenie polityki pieniężnej. Na dłuższą metę polityka realizowania wielu celów nie była skuteczna, a na dodatek niosła za sobą szereg negatywnych skutków ubocznych. Na przykład forsowanie mocnego kursu waluty osłabiało konkurencyjność eksportu i przyczyniało się do stagnacji gospodarczej. Z kolei polityka „taniego pieniądza", nakierowana na stymulowanie wzrostu, powodowała narastanie napięć inflacyjnych, wzrost zadłużenia i utratę konkurencyjności międzynarodowej.

Pod wpływem tych doświadczeń w latach 60. i 70. XX wieku podjęto rewizję dotychczas stosowanych modeli polityki pieniężnej. W szczególności keynesowskie recepty pieniężnego pobudzania gospodarki zostały poddane krytyce ze strony ekonomistów reprezentujących nurt monetarystyczny w ekonomii, którego czołowym przedstawicielem był Milton Friedman. Zarówno on, jak i inni monetaryści (np. H.G. Johnson, K. Brunner, A.H. Meltzer) negowali zdolność polityki pieniężnej do wywierania trwałego wpływu na produkcję i zatrudnienie, twierdząc, że uczestnicy rynku są w stanie przewidzieć, iż wzrost podaży pieniądza wywoła wzrost inflacji i spadek realnych wynagrodzeń. Skłoni to pracowników do wycofania części podaży pracy z rynku, i w efekcie poziom aktywności gospodarczej

[14] Zob. rozdział 2.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

nie zmieni się. Ramy nowego podejścia do polityki pieniężnej wyznaczały trzy główne myśli: friedmanowskie przekonanie, że inflacja jest „zawsze i wszędzie zjawiskiem pieniężnym", założenie o istnieniu racjonalnych oczekiwań dotyczących inflacji oraz założenie o neutralności pieniądza, czyli braku relacji w długim okresie między zmiennymi nominalnymi (takimi jak inflacja, nominalne stopy procentowe, nominalny kurs walutowy) a zmiennymi realnymi (takimi jak produkcja i zatrudnienie)[15].

5.3.2. Stabilność cen jako cel polityki pieniężnej

Konsekwencją zastosowania nowego podejścia było upowszechnienie się w latach 80. przekonania o tym, że polityka pieniężna nie jest w stanie skutecznie realizować w tym samym czasie wielu różnych celów, a w szczególności nie jest w stanie w sposób trwały wpływać na zmienne realne, takie jak produkcja czy zatrudnienie. Jednocześnie to właśnie nadmierna podaż pieniądza jest „zawsze i wszędzie" źródłem inflacji. Uznano, że trzeba stawiać przed polityką pieniężną skromniejsze zadania, lecz zarazem takie, które dzięki niej można realizować skuteczniej niż przy wykorzystaniu innych narzędzi polityki gospodarczej. W większości krajów rozwiniętych do zadań takich należy w pierwszej kolejności utrzymanie stabilności cen.

Na czym polega stabilność cen? W dosłownym rozumieniu jest to brak zmian cen, a więc stopa inflacji równa zeru. Oznacza to, że wartość realna pieniądza nie ulega obniżeniu, a zatem jego siła nabywcza, mierzona ilością dóbr i usług, jakie można nabyć za określoną kwotę, nie spada. Ochrona siły nabywczej pieniądza jest z pewnością ważna z punktu widzenia interesów konsumenta. Lecz żaden bank centralny na świecie nie dąży do zapewnienia stabilności cen interpretowanej w tak dosłowny sposób. Inflacja na poziomie zerowym nie jest bowiem korzystna dla gospodarki z trzech zasadniczych powodów.

Po pierwsze, jeśli *średnia* stopa inflacji w całej gospodarce miałaby się równać zeru, to ceny przynajmniej części dóbr i usług musiałyby spadać. Rodziłoby to niebezpieczeństwo deflacji w niektórych sektorach, przede wszystkim w sektorze produkcji przemysłowej. Jak pokazują dane statystyczne, ceny usług z reguły rosną bowiem szybciej niż ceny produktów przemysłowych (ze względu na wolniejsze tempo wzrostu wydajności pracy). Oznacza to, że jeśli średnia stopa inflacji

[15] Zob. zwłaszcza Friedman (1970, 1977).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

miałaby być równa zeru, ceny produktów przemysłowych musiałyby spadać. Taka deflacja może zaś mieć groźne następstwa w odniesieniu do gospodarki, o czym przekonamy się w dalszej części rozdziału.

Po drugie, zerowa średnia stopa inflacji oznacza, że zmiany cen relatywnych poszczególnych produktów, wynikające ze zmian technologicznych lub zmian warunków rynkowych, wymagałyby w określonych sektorach absolutnego obniżenia kosztów wytwarzania, w tym wynagrodzeń. Lecz obniżanie nominalnych wynagrodzeń pozostaje bardzo trudne ze względu na konieczność zmian kontraktów płacowych, co spotyka się zawsze z silnym oporem pracowników. Zjawisko sztywności płac w dół (*downward rigidity of wages*) praktycznie uniemożliwia elastyczne dostosowania płac. Niezbędne dostosowania cen relatywnych dokonują się znacznie łatwiej w warunkach umiarkowanie rosnących cen, bo nie zachodzi wówczas konieczność obniżania wynagrodzeń. Dlatego stabilność cen jest powszechnie definiowana jako pewna większa od zera, ale niska i stabilna stopa inflacji[16].

Po trzecie, panuje powszechne przekonanie, że standardowe mierniki inflacji, takie jak CPI czy PPI, zawyżają jej rzeczywisty. Jest to spowodowane stosowaniem okresowo stałej struktury koszyka dóbr i usług, co uniemożliwia w krótkim okresie uchwycenie zmian w strukturze konsumpcji (produkcji), a także zmianami jakościowymi, które wywołują wzrost cen tych samych (choć już nie takich samych) produktów (zob. też ramka 5.2). Skala zawyżenia inflacji z tego tytułu może być dosyć znacząca, na przykład odnośnie do gospodarki amerykańskiej oceniana jest na 0,5–2,0 punktów procentowych (Barro 1995, Boskin i in. 1996). Gdyby więc w tych warunkach przyjęto cel inflacyjny na poziomie bliskim zeru, taka polityka groziłaby zepchnięciem gospodarki w deflację[17].

Szczególna skuteczność polityki pieniężnej w stabilizowaniu inflacji wynika ze wspomnianego wyżej założenia mówiącego o tym, że trwała inflacja nie może mieć miejsca bez stałego zwiększania podaży pieniądza, co czyni z polityki pieniężnej naturalny instrument spowalniania inflacji. Nawet jeśli w ostatnich dwóch dekadach średniookresowy związek między podażą pieniądza a inflacją uległ osłabieniu na skutek liberalizacji przepływów kapitałowych oraz innowacji finansowych, związek

[16] Przyjęcie celu inflacyjnego na poziomie równym zeru mogłoby być uzasadnione tylko wobec bardzo restrykcyjnych i mało realistycznych założeń (Woodford 2003, Buiter 2006).

[17] Często wysuwanym argumentem przeciwko ustalaniu celu inflacyjnego na zbyt niskim poziomie jest związana z tym konieczność utrzymywania także nominalnych stóp procentowych na niskim poziomie, co ogranicza możliwości ich dalszego obniżania w warunkach dekoniunktury (nominalne stopy procentowe teoretycznie nie mogą spaść poniżej zera – zob. Summers 1991).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ten pozostaje na tyle wyraźny, że polityka pieniężna jest powszechnie uznawana za instrument pierwszego wyboru, jeśli chodzi o kontrolowanie inflacji. Ponieważ regulowaniem ilości pieniądza w obiegu zajmuje się bank centralny, to właśnie on powinien być również instytucją odpowiedzialną za utrzymanie stabilności cen.

Od momentu, w którym większość banków centralnych przyjęła jako główny cel polityki pieniężnej utrzymanie stabilności cen, inflacja spadła. Według danych Międzynarodowego Funduszu Walutowego, w latach 1980–2008 udział państw notujących inflację poniżej 5% rocznie wzrósł z 10% do 30%, a udział państw z inflacją przekraczającą 15% rocznie spadł z ponad 40% do 10%. Sukces ten miał niewątpliwie związek ze zmianami instytucjonalnymi: banki centralne w wielu krajach stały się niezależne od władz politycznych, a cele, które miały obowiązek realizować – czyli tzw. mandat – zostały zapisane w prawie krajowym na poziomie konstytucyjnym lub co najmniej ustawowym. Lecz warto dodać, że spadek inflacji wynikał także z korzystnej sytuacji w gospodarce światowej w latach 90. i pierwszej dekadzie XXI wieku. W wyniku globalizacji na rynkach międzynarodowych pojawili się nowi, konkurencyjni dostawcy z krajów rozwijających się (Chiny, Indie), gdzie koszty wytwarzania dóbr były niższe. Wywołało to presję na spadek cen produktów będących przedmiotem wymiany międzynarodowej i stało się ważnym czynnikiem dezinflacji w krajach rozwiniętych oraz w odniesieniu do całej gospodarki światowej.

5.3.3. Pozostałe cele współczesnej polityki pieniężnej: stabilizacja makroekonomiczna, stabilność kursu walutowego i stabilność finansowa

Jakkolwiek w większości krajów rozwiniętych nadrzędnym celem banków centralnych jest utrzymanie stabilności cen, to ich mandat obejmuje także inne zadania, choć z reguły podporządkowane celowi nadrzędnemu. Do tych dodatkowych zadań należy przede wszystkim działanie na rzecz stabilizacji makroekonomicznej. Większość ekonomistów uznaje bowiem, że choć polityka pieniężna nie ma wpływu na zmienne realne w długim okresie, to w krótkim okresie jest w stanie oddziaływać na produkcję i zatrudnienie. Ma to miejsce tak długo, jak długo zwiększanie podaży pieniądza nie wywołuje wzrostu oczekiwań inflacyjnych, co umożliwia wykorzystanie tej polityki w celu stabilizowania gospodarki w obrębie cyklu koniunkturalnego, w tym zwłaszcza zapobieganie głębszym załamaniom koniunktury.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wzrost podaży pieniądza przy danym poziomie popytu na pieniądz pociąga za sobą wzrost zagregowanego popytu w gospodarce. Dzieje się tak dlatego, że w warunkach częściowo sztywnych cen i płac (czyli powolnej oraz niepełnej reakcji cen i płac na zmiany popytu), wzrost podaży pieniądza wywołuje spadek realnych stóp procentowych, a to powoduje wzrost popytu, przede wszystkim inwestycyjnego, a także konsumpcyjnego. Dodatkowo, w systemie płynnego kursu walutowego, spadek stóp procentowych wywołuje deprecjację waluty krajowej i wzrost popytu zagranicznego. Z kolei wzrost zagregowanego popytu przekłada się częściowo na zwiększenie produkcji, a częściowo na wzrost cen, w proporcjach, które zależą głównie od stopnia wykorzystania mocy wytwórczych, a także od elastyczności popytu na dobra i usługi względem stopy procentowej oraz elastyczności popytu na pieniądz względem stopy procentowej. Jak pokazaliśmy w rozdziale 2, im niższy jest stopień wykorzystania zdolności wytwórczych, im bardziej elastyczny staje się popyt na towary oraz im mniej elastyczny okaże się popyt na pieniądz względem stopy procentowej, tym skuteczniejsza będzie polityka pieniężna w pobudzaniu produkcji i zatrudnienia. Stymulowanie gospodarki przy pomocy polityki pieniężnej jest zatem szczególnie zalecane wówczas, gdy gospodarka znajduje się w stanie recesji, ponieważ utrzymują się wtedy duże niewykorzystane zdolności produkcyjne, w tym z reguły wysokie bezrobocie. W takich warunkach ekspansywna polityka pieniężna staje się skuteczna. Z kolei gdy gospodarka znajduje się w fazie silnego wzrostu, zdolności wytwórcze są w pełni wykorzystane, wobec czego wzrost podaży pieniądza spowodowałby wówczas głównie wzrost cen. W takich warunkach polityka pieniężna powinna pozostawać restrykcyjna. Prowadzenie antycyklicznej polityki pieniężnej, choć pożądane, nie jest jednak zadaniem łatwym, głównie ze względu na nieuchronne opóźnienia w transmisji impulsów pieniężnych.

Celem polityki pieniężnej może być także zapewnienie stabilnego kursu walutowego. Było to bardzo popularne w okresie przed 1990 rokiem, kiedy większość państw stosowała stałe kursy. Gdy w danym kraju obowiązuje system sztywnego (stałego) kursu walutowego, bank centralny ma prawny obowiązek kupować i sprzedawać każdą ilość walut obcych po ustalonym kursie. Wówczas podaż pieniądza krajowego staje się zmienną endogeniczną, zależną od skali napływu i odpływu walut obcych, a polityka pieniężna jest nastawiona wyłącznie na utrzymanie wartości waluty krajowej względem walut obcych. Zaletą stałego kursu staje się jego przewidywalność oraz możliwość obniżenia kosztów transakcyjnych dla krajowych wytwórców, wadą jest natomiast brak automatycznego mechanizmu przywracania równowagi i duża podatność na kryzysy. W okresie ostatnich dwóch dekad wiele

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

krajów stosujących różne odmiany systemu stałego kursu przeszło przez załamania walutowe i w rezultacie system ten stracił na atrakcyjności.

Kryzys finansowy lat 2007–2010 spowodował powrót do dyskusji o celach polityki pieniężnej i postawił w centrum uwagi kwestię odpowiedzialności banków centralnych za utrzymanie stabilności systemu finansowego, zwłaszcza bankowego. Okazało się, że, działając w ramach dotychczasowego modelu polityki pieniężnej, banki centralne nie były w stanie zapobiec narastaniu nierównowagi i napięć na rynkach finansowych, a w szczególności samo dążenie do zapewnienia stabilności cen nie gwarantowało stabilności systemu finansowego. Dlatego też coraz częściej proponuje się uzupełnienie podstawowego celu, jakim jest stabilność cen, drugim równorzędnym celem, jaki stanowi stabilność systemu finansowego (zob. Blanchard, Dell'Ariccia i Mauro 2010). Instrumentem, który ma zapewniać stabilność całego systemu finansowego, są regulacje makroostrożnościowe hamujące narastanie ryzyka systemowego (de Larosière 2009, Szpunar 2012, Tymoczko 2012, Rosati 2014).

Stabilność systemu finansowego zależy głównie od sposobu postępowania instytucji finansowych (banki, firmy ubezpieczeniowe, fundusze inwestycyjne). Gdy podejmują one nadmiernie ryzykowne decyzje, na przykład inwestując w niepewne aktywa bądź udzielając kredytów niewiarygodnym klientom i bez odpowiednich zabezpieczeń, w sektorze finansowym wrasta ryzyko systemowe, które może prowadzić do bankructwa i zaprzestania dostarczania do gospodarki ważnych produktów finansowych (takich jak kredyty czy ubezpieczenia). Regulacje i politykę makroostrożnościową (*macroprudential regulations and policy*) definiuje się najczęściej jako całościowe działanie władzy gospodarczej, koncentrujące się na identyfikacji i ograniczaniu ryzyka systemowego oraz zapewnieniu dzięki temu stabilności sektora finansowego jako całości (Crockett 2000, Brunnermeier i in. 2009, Billio i in. 2010, Clement 2010, IMF 2011a, Rosati 2014). Kontrola ryzyka systemowego poprzez regulacje makroostrożnościowe polega na ograniczaniu realnych kosztów gospodarczych wywoływanych zaburzeniami, jakie mogą powstawać w procesie dostarczania ważnych produktów finansowych do gospodarki. Do kosztów tych należy m.in. spadek produkcji wywołany ograniczeniem akcji kredytowej, spadek zatrudnienia i wzrost bezrobocia, bankructwo firm i banków, wzrost długu publicznego (Elliott, Feldberg i Lehnert 2013; IMF 2011b).

Większość ekonomistów jest zdania, że odpowiedzialność za regulacje i nadzór makroostrożnościowy powinna spoczywać na banku centralnym. Przemawia za tym fakt, że bank centralny gromadzi informacje sprawozdawcze o bankach

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

komercyjnych, posiada wyłączne kompetencje w zakresie polityki pieniężnej i w obszarze stabilności finansowej oraz pełni funkcję pożyczkodawcy ostatniej szansy dla banków. Dlatego też niezależny bank centralny posiada zarówno podstawy, jak i silną motywację do podejmowania działań w sytuacji rosnącego ryzyka. Ponadto przeważający udział banku centralnego w polityce makroostrożnościowej ułatwia niezbędną koordynację z polityką pieniężną, wpływającą na zagregowane poziomy popytu i kredytu (szerzej na ten temat piszą m.in. Goodhart 2011 i Tymoczko 2012)[18].

5.3.4. Cele polityki pieniężnej w wybranych krajach

Choć cele polityki pieniężnej w większości państw są podobne, występują między nimi pewne różnice. Przyjrzyjmy się temu, jak zostały zdefiniowane oficjalne cele polityki pieniężnej w wybranych krajach.

Cele polityki pieniężnej w Unii Gospodarczej i Walutowej (w strefie euro) zapisane są w Traktacie o Funkcjonowaniu Unii Europejskiej (TFUE). Art. 127.1 TFUE określa główny cel funkcjonowania Europejskiego Systemu Banków Centralnych (ESBC): „głównym celem ESBC jest utrzymanie stabilności cen. Bez uszczerbku dla celu stabilności cen, ESBC wspiera polityki gospodarcze w Unii, mając na względzie przyczynianie się do osiągania celów Unii ustanowionych w art. 3 TUE”[19].

Z zapisu tego jasno wynika, że głównym celem polityki pieniężnej w strefie euro jest utrzymanie stabilności cen, z kolei zadaniem podporządkowanym temu celowi jest wspieranie polityki gospodarczej w państwach Unii, lecz tylko wtedy, gdy nie narusza to celu głównego. Mamy tu do czynienia z wyraźną hierarchizacją celów, na czele z dążeniem do zapewnienia stabilności cen.

Cele polityki pieniężnej w Polsce określa ustawa o Narodowym Banku Polskim. W art. 3 znajduje się zapis: „podstawowym celem działalności NBP jest utrzymanie stabilnego poziomu cen, przy jednoczesnym wspieraniu polityki gospodarczej rządu, o ile nie ogranicza to podstawowego celu NBP” (ustawa o Narodowym Banku Polskim z dnia 29 sierpnia 1997 roku, art. 3, ust. 1).

Łatwo zauważyć, że cele polityki pieniężnej w Unii Gospodarczej i Walutowej oraz w Polsce pozostają w istocie identyczne. W obu przypadkach celem

[18] Polityka makroostrożnościowa została omówiona w rozdziale 7.

[19] Tekst ten jest następnie w dosłownej formie powtórzony w art. 2 Statutu Europejskiego Banku Centralnego (protokół nr 4 w sprawie Statutu Europejskiego Systemu Banków Centralnych i Europejskiego Banku Centralnego, dołączony do Traktatu o Funkcjonowaniu Unii Europejskiej).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

nadrzędnym jest utrzymanie stabilności cen, a zadaniem podporządkowanym jest wspieranie polityki gospodarczej. Zbieżność ta wynika z dążenia do zapewnienia pełnej harmonizacji polskich przepisów dotyczących funkcjonowania banku centralnego i polityki pieniężnej z przepisami obowiązującymi w Unii Europejskiej. Już w 1997 roku, kiedy przygotowywano obowiązującą obecnie ustawę o Narodowym Banku Polskim, przyjęto rozwiązania zgodne z ustawodawstwem unijnym, ponieważ już wówczas polskie władze planowały akcesję Polski do Unii Europejskiej. Obecnie Polska, jako państwo członkowskie UE, ma obowiązek uczestnictwa w Unii Gospodarczej i Walutowej, a to wymaga ujednolicenia prawodawstwa w obszarze polityki pieniężnej.

Bardzo zbliżone – czy wręcz identyczne – rozwiązania zostały także wprowadzone w innych krajach europejskich. Na przykład w Wielkiej Brytanii celem polityki pieniężnej jest zapewnienie stabilności cen (rozumianej jako niska inflacja) oraz, bez uszczerbku dla celu stabilności cen, wspieranie celów gospodarczych rządu, w tym dążenia do wywołania wzrostu gospodarczego i zatrudnienia[20]. Podobnie w Czechach mamy do czynienia z celem nadrzędnym, jakim jest stabilność cen, oraz celem mu podporządkowanym, jaki stanowi wsparcie dla ogólnej polityki gospodarczej rządu Czech i Unii Europejskiej[21]. Dodatkowym celem ma zaś być zapewnienie stabilności finansowej i bezpieczeństwa działania systemu finansowego. Stabilność i bezpieczeństwo systemu finansowego jest również dodatkowym celem działania – obok zapewnienia stabilności cen – banku centralnego Szwecji[22]. Również w Szwajcarii celem polityki pieniężnej jest zapewnienie stabilności cen, przy uwzględnieniu ekonomicznych warunków rozwoju[23]. Zapewnienie stabilności

[20] W ustawie o Banku Anglii (Bank of England Act, 1988), w części II, sekcji 11, czytamy:

In relations to monetary policy, the objectives of the Bank of England shall be:

a) to maintain price stability, and
b) subject to that, to support the economic policy of Her Majesty's Government, including its objectives for growth and employment.

[21] W ustawie o Narodowym Banku Czech z 1993 roku, w art. 2 znajduje się zapis:

The primary objective of the Czech National Bank shall be to maintain price stability. In addition, the Czech National Bank shall work to ensure financial stability and the safe and sound operation of the financial system in the Czech Republic. Without prejudice to its primary objective, the Czech National Bank shall support the general economic policies of the Government leading to sustainable economic growth and the general economic policies in the European Union with a view to contributing to the achievement of the objectives of the European Union (art. 2, Act no. 6/1993 Coll., on the Czech National Bank).

[22] W ustawie o szwedzkim Riksbanku z 1988 roku czytamy:

The objective of the Riksbank's activities shall be to maintain price stability. The Riksbank shall also promote a safe and efficient payments system (The Sveriges Riksbank Act, 1988, rozdział 1, art. 2).

[23] National Bank Act, art. 5.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

cen jest także głównym celem polityki pieniężnej w wielu krajach pozaeuropejskich, m.in. w Japonii, Kanadzie i Nowej Zelandii.

Inaczej z kolei wygląda mandat do prowadzenia polityki pieniężnej sprawowany przez bank centralny Stanów Zjednoczonych. Zgodnie z ustawą powołującą do życia System Rezerwy Federalnej (Federal Reserve Act 1913), bank centralny USA prowadzi politykę pieniężną w taki sposób, aby osiągać cele maksymalnego zatrudnienia, stabilnych cen i umiarkowanych długookresowych stóp procentowych[24]. Mandat Systemu Rezerwy Federalnej obejmuje więc trzy równorzędne cele, bez jednoznacznego wskazania na nadrzędność któregokolwiek z nich. Zbliżone podejście stosuje także Australia, gdzie bank centralny realizuje dwa równorzędne cele: stabilność cen i pełne zatrudnienie.

Rozwiązanie przyjęte w USA bywa niekiedy krytykowane, ponieważ narusza zasadę Tinbergena, głoszącą, że liczba celów polityki nie powinna być większa niż liczba niezależnych instrumentów[25]. Jak wspomniano wyżej, polityka pieniężna polega na regulowaniu ilości pieniądza w obiegu. Ilość pieniądza jest więc podstawowym instrumentem, przy pomocy którego polityka pieniężna realizuje swoje cele. Jak się przekonamy, regulowanie ilości pieniądza w obiegu może odbywać się przy wykorzystaniu różnych narzędzi technicznych, takich jak zmiany poziomu stóp procentowych, zmiany stopy rezerwy obowiązkowej, określanie wielkości emisji pieniądza rezerwowego czy nakładanie limitów kredytowych na banki komercyjne. Narzędzia te nie są jednak niezależnymi instrumentami, ponieważ ich wspólnym przeznaczeniem jest regulowanie ilości pieniądza. Dlatego też uznaje się, że bank centralny dysponuje w istocie tylko jednym instrumentem, jakim jest podaż pieniądza. W tej sytuacji skuteczne realizowanie większej liczby celów może okazać się niewykonalne – w każdym momencie można przy pomocy jednego instrumentu realizować tylko jeden cel, podczas gdy pozostałe zadania kształtują się wynikowo, jako zmienne rezydualne. Czy sposób zapisu celów polityki pieniężnej w USA jest zatem rzeczywiście sprzeczny z zasadą Tinbergena?

---

[24] W sekcji 2 ustawy Federal Reserve Act (po poprawkach wprowadzonych w 1978 roku ustawą Humphreya-Hawkinsa) znajduje się zapis:

The Board of Governors of the Federal Reserve System and the Federal Open Market Committee shall maintain long run growth of the monetary and credit aggregates commensurate with the economy's long run potential to increase production, so as to promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates (Federal Reserve Act, section 2A, Monetary Policy Objectives).

[25] Zasada Tinbergena bywa także nazywana „złotą zasadą polityki gospodarczej" (zob. rozdział 1, punkt 1.3.4).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Okazuje się, że niekoniecznie. Ustawa określająca mandat Rezerwy Federalnej oznacza, że w praktyce amerykański bank centralny – tak zwany Fed – realizuje *de facto* w danym momencie tylko jeden cel spośród trzech zapisanych w mandacie, przy czym zawsze realizowany jest ten, który w danym momencie uchodzi za najważniejszy. Fed każdorazowo ocenia więc stopień odchylenia rzeczywistych wartości inflacji, zatrudnienia i stóp procentowych od ustalonych celów oraz identyfikuje obszar, w którym odchylenia te są największe. Następnie luzuje albo zaostrza politykę pieniężną w zależności od tego, czego wymaga realizacja celu, w przypadku którego odchylenie w danym momencie jest największe.

W tabeli 5.9 przedstawiono cele polityki pieniężnej w wybranych, najważniejszych krajach świata, prowadzących autonomiczną politykę pieniężną i stosujących płynne kursy walutowe. Nie uwzględniono państw stosujących kontrolowane kursy walutowe ze względu na to, że poziom inflacji w tych krajach jest zmienną podporządkowaną celowi stałego kursu. Dane zawarte w tabeli wskazują, że stabilność cen jest głównym lub jedynym celem polityki pieniężnej wśród większości państw. Wyjątkami od tej zasady są regulacje przyjęte w Stanach Zjednoczonych oraz w Australii, gdzie poza dążeniem do osiągniecia stabilności cen banki centralne realizują również inne, równorzędne cele, związane przede wszystkim z zapewnieniem pełnego zatrudnienia. Tabela pozwala także zauważyć, że wśród większości krajów stabilność cen definiowana jest jako stopa inflacji utrzymująca się w granicach 2–3% rocznie.

Tabela 5.9. Cele polityki pieniężnej i cele inflacyjne w wybranych krajach – stan na dzień 31.12.2014 roku

Kraj	Cel główny polityki pieniężnej	Cele dodatkowe polityki pieniężnej	Cel inflacyjny (w skali rocznej)
UGW	stabilność cen	wspieranie polityki gospodarczej Unii	stopa inflacji utrzymująca się poniżej, lecz blisko 2%
USA	stabilność cen, pełne zatrudnienie, umiarkowane długookresowe stopy procentowe	brak	stopa inflacji około 2%
Wielka Brytania	stabilność cen	wspieranie polityki gospodarczej rządu	stopa inflacji 2%
Szwajcaria	stabilność cen	uwzględnianie ekonomicznych warunków rozwoju	stopa inflacji poniżej 2%
Szwecja	stabilność cen	bezpieczny i efektywny system płatniczy	stopa inflacji około 2%

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Kraj	Cel główny polityki pieniężnej	Cele dodatkowe polityki pieniężnej	Cel inflacyjny (w skali rocznej)
Czechy	stabilność cen	wspieranie polityki gospodarczej rządu	stopa inflacji 2%
Norwegia	stabilność cen	brak	stopa inflacji około 2,5%
Japonia	stabilność cen	brak	stopa inflacji 2%
Kanada	stabilność cen	brak	stopa inflacji 2% +/– 1 punkt procentowy
Australia	stabilność cen, pełne zatrudnienie, ekonomiczny dobrobyt ludności	brak	stopa inflacji zawierająca się w przedziale między 2% a 3%
Nowa Zelandia	stabilność cen	brak	stopa inflacji około 2% +/– 1 punkt proc.
Polska	stabilność cen	wspieranie polityki gospodarczej rządu	stopa inflacji 2,5% +/– 1 punkt procentowy

Źródło: opracowanie własne.

5.3.5. Mierniki inflacji

Jeśli celem polityki pieniężnej ma być utrzymanie stabilnych cen, bank centralny musi najpierw określić miernik inflacji, przy pomocy którego będzie dokonywał jej pomiaru, a następnie przyjąć i ogłosić definicję stabilnych cen, tj. ustalić pożądaną stopę inflacji, która jego zdaniem okaże się najkorzystniejsza dla gospodarki. Co do pomiaru inflacji w praktyce najczęściej stosuje się mierniki odnoszące się do określonych szerokich grup towarów. Do najbardziej popularnych należą: wskaźnik cen konsumpcyjnych (*Consumer Price Index* – CPI), wskaźnik cen detalicznych (*Retail Price Index* – RPI), wskaźnik cen hurtowych (*Wholesale Price Index* – WPI), wskaźnik cen producentów (*Producer Price Index* – PPI) oraz deflator PKB (*GDP deflator*). Banki centralne stosują także rozmaite wskaźniki tzw. inflacji bazowej (*core inflation*), z których wyłączone są zmiany cen nie wynikające z działania krajowej polityki pieniężnej (chodzi m.in. o ceny ustalane administracyjnie, ceny importowanych surowców i paliw bądź ceny, w których duży udział mają podatki pośrednie, których wysokość zależy od decyzji administracyjnych). Każdy z wymienionych wskaźników dostarcza nieco innych informacji i dlatego wszystkie są z reguły przedmiotem analizy. Bank centralny musi jednak wybrać jeden wskaźnik wiodący, który stanie się głównym miernikiem inflacji (*headline inflation index*) i będzie wykorzystywany do określenia stabilności cen oraz komunikowania się z uczestnikami rynku. Najczęściej stosowanym w tym przypadku

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

miernikiem inflacji jest wskaźnik cen konsumpcyjnych CPI. Jego zaletą jest prosta i zrozumiała konstrukcja oraz bezpośredni związek z nastrojami konsumentów, wadę stanowią natomiast pewne obciążenia statystyczne (zob. ramka 5.2).

Ramka 5.2. Wskaźnik CPI

Wskaźnik CPI jest średnią ważoną cen dóbr i usług konsumpcyjnych nabywanych przez przeciętne gospodarstwo domowe, przy czym wagami są udziały poszczególnych towarów w całkowitej konsumpcji. Przy obliczaniu wskaźnika bierze się pod uwagę produkty wchodzące w skład tzw. statystycznego koszyka zakupów (wyłączone są towary inwestycyjne, nieprzetworzone produkty rolne, surowce przemysłowe). W Polsce obliczaniem wskaźnika CPI zajmuje się Główny Urząd Statystyczny (GUS). Na podstawie systematycznych badań budżetów gospodarstw domowych GUS określa statystyczny koszyk zakupów i co roku ustala, co i w jakich ilościach kupuje „uśrednione" gospodarstwo domowe, oraz jaki jest udział poszczególnych produktów i usług w łącznych wydatkach. Na podstawie obserwacji zmian w zakresie cen tak dobranych dóbr i usług (ich lista liczy łącznie około 2 tysięcy pozycji, a ceny sprawdza się raz lub dwa razy w miesiącu) GUS oblicza wskaźnik inflacji.

Z konstrukcji wskaźnika CPI wynika, że jest on nie tyle miarą ogólnej inflacji (nie obejmuje cen surowców i dóbr inwestycyjnych), ile raczej wskaźnikiem zmian w zakresie kosztów utrzymania przeciętnego gospodarstwa domowego. W praktyce występują jednak różnice między koszykiem „codziennym", rzeczywistym, a koszykiem „statystycznym". Ceny niektórych dóbr czy usług rosną w ciągu roku szybciej, niż dzieje się to w przypadku innych produktów, co skłania konsumentów do zastępowania ich towarami tańszymi (tzw. efekt substytucji). Mamy też do czynienia ze wzrostem cen poszczególnych dóbr, wynikającym z poprawy ich jakości. Wskazywałoby to, że wskaźnik CPI może przeszacowywać rzeczywistą inflację (tę tendencję nazywamy zjawiskiem „obciążenia" wskaźnika CPI). Występują też tzw. „nierówności inflacyjne", tj. różnice w strukturze towarów i usług w obrębie struktury wydatków różnych grup społecznych, które rzutują na ceny. Większy udział gospodarstw zamożniejszych w całkowitej wartości wydatków konsumpcyjnych sprawia, że gospodarstwa o niskich dochodach są we wskaźniku CPI niedoreprezentowane. W związku z tym mówi się niekiedy, że wskaźnik CPI staje się wskaźnikiem „plutokratycznym" – czyli silniej odzwierciedlającym wpływ bogatszych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

grup społecznych. Co więcej, zmiany w strukturze spożycia wywołane pojawianiem się nowych produktów i postępem technologicznym dostarczają trudności przy porównywaniu wskaźnika CPI w dłuższym okresie[26]. Podobnymi wadami są jednak obciążone także pozostałe indeksy cen.

Definiując stabilność cen, bank centralny określa pożądany poziom inflacji (mierzony wskaźnikiem wiodącym), który z jednej strony zapewnia inflację na tyle niską, aby nie wywoływała negatywnych skutków dla gospodarki, z drugiej zaś na tyle większą od zera, aby procesy dostosowawcze w gospodarce przebiegały płynnie i bez zakłóceń. W rozwiniętych, zrównoważonych gospodarkach rynkowych stabilność cen jest z reguły definiowana jako stopa inflacji na poziomie 2–3% w skali rocznej (zob. punkt 5.1.2).

5.3.6. Dlaczego (wysoka) inflacja jest szkodliwa?

Inflacja oznacza wzrost ogólnego poziomu cen w gospodarce. Nie chodzi więc o zmiany w zakresie cen poszczególnych dóbr czy usług: ceny te mogą rosnąć szybciej lub wolniej, niektóre z nich będą nawet spadać, lecz jeśli występuje wzrost ogólnego wskaźnika cen, mamy wówczas do czynienia z inflacją. Wskaźnik taki musi obejmować ceny wszystkich lub większości towarów znajdujących się na rynku, a jego konstrukcja powinna uwzględniać udział poszczególnych towarów w całości sprzedaży w gospodarce. Takie „ważone" wskaźniki cen są obliczane i publikowane przez krajowe i międzynarodowe instytucje statystyczne. Im szybszy jest wzrost ogólnego poziomu cen, tym wyższa staje się inflacja w gospodarce.

Ogólnie rzecz biorąc, inflacja pojawia się wtedy, gdy podaż pieniądza jest większa od jego popytu. Stan ten może więc wynikać z nadmiernego wzrostu masy pieniądza w obiegu, któremu nie towarzyszy odpowiedni wzrost podaży towarów – dóbr i usług, ponieważ wtedy popyt (transakcyjny) na pieniądz jest odpowiednio mniejszy. Mamy wówczas do czynienia ze wzrostem cen. Inflacja może też wynikać ze spadku popytu na pieniądz przy danej podaży. Ma to miejsce w okresach zwiększonej niepewności na rynkach, kiedy ludzie pozbywają się pieniądza, nabywając większe ilości towarów oraz inne aktywa, takie jak złoto, nieruchomości czy waluty obce.

[26] Zob. Hałka i Leszczyńska (2011).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Inflacja, zwłaszcza wysoka, jest zjawiskiem gospodarczo i społecznie szkodliwym. Wiemy wszyscy z własnego doświadczenia, że jeśli określone towary drożeją, zmuszeni jesteśmy ograniczyć ich zakup i spożycie. Poziom życia ludzi ulega wówczas obniżeniu. Zasadniczo wzrost cen towarów, jeśli nie towarzyszy mu taki sam wzrost dochodów nominalnych, wywołuje także spadek dochodów realnych i zmniejsza siłę nabywczą ludności. Obywatele narzekają wówczas na drożyznę i mówią, że inflacja „zjada" ich dochody.

Wysoka inflacja hamuje również wzrost gospodarczy, i to z kilku powodów. Po pierwsze, wyższa inflacja oznacza z reguły większą zmienność cen: ich poszczególne wartości zmieniają się proporcjonalnie w większym stopniu – nie o kilka, a o kilkanaście czy kilkadziesiąt procent. Zwiększona zmienność – statystycznie mierzona wzrostem wariancji – powoduje wzrost niepewności co do przyszłych wartości cen, a to z kolei komplikuje planowanie zakupów i utrudnia wycenę opłacalności projektów inwestycyjnych. W takiej sytuacji część inwestorów powstrzyma się od realizacji inwestycji, co przełoży się na spowolnienie wzrostu gospodarczego. Po drugie, inflacja zaburza rynkowy mechanizm sygnałów cenowych. W normalnych warunkach wzrost ceny danego produktu jest sygnałem wzrostu popytu na ten produkt. Skłania to producentów do zwiększenia podaży danego produktu, najpierw poprzez wzrost produkcji w ramach istniejących mocy wytwórczych, a następnie poprzez podejmowanie nowych inwestycji. W ten sposób mechanizm rynkowy przywraca naruszoną równowagę miedzy podażą a popytem. Lecz w warunkach inflacji część producentów nie będzie w stanie ocenić, czy obserwowany wzrost cen wynika rzeczywiście ze wzrostu popytu, czy jest jedynie skutkiem ogólnej inflacji, i z w związku z tą niepewnością powstrzymają się od działań zwiększających produkcję. Z kolei inni przedsiębiorcy potraktują inflacyjny wzrost cen jako przejaw przewagi popytu nad podażą i zwiększą produkcję, co jednak może okazać się decyzją chybioną, jeśli wzrost cen miał wyłącznie charakter inflacyjny. Inflacja zniekształca więc sygnały rynkowe i sprawia, że rynkowy mechanizm przywracania równowagi ulega zakłóceniu. W taki sam sposób wysoka inflacja zaburza informacyjną wartość kształtowania się cen relatywnych. W efekcie dochodzi do obniżenia inwestycji, nieoptymalnej alokacji zasobów i spadku tempa wzrostu. Wreszcie, wysoka inflacja dezorganizuje rynek i pogłębia chaos w gospodarce. Ludzie coraz więcej czasu i pieniędzy poświęcają na poszukiwanie tańszych źródeł zakupów (tę stratę społeczną nazywa się kolokwialnie „kosztem zdartych zelówek"), a firmy wkładają coraz większy wysiłek w rozpoznanie sytuacji rynkowej, której rzeczywisty stan jest zniekształcony przez inflację. Rosną więc koszty transakcyjne, co przyczynia

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

się do dalszego obniżenia tempa wzrostu gospodarczego. Wszystko to może prowadzić do stagnacji, a nawet przewlekłej recesji[27].

Inflacja pociąga za sobą też spontaniczną i nieplanowaną redystrybucję dochodów na niekorzyść najuboższych warstw społeczeństwa (Easterly i Fischer 2000). Mechanizm tego ukrytego transferu jest prosty. Ponieważ nominalne dochody z pracy (wynagrodzenia) i świadczenia społeczne (emerytury, renty, stypendia, zasiłki) są z reguły określone kwotowo, to ich realna wartość maleje na skutek inflacji. Nawet jeśli dochody te podlegają indeksacji w związku z inflacją (jak na przykład emerytury w Polsce), ma ona z reguły miejsce *ex post*, co sprawia, że realna siła nabywcza dochodów nominalnych jest systematycznie zaniżona. W przeciwieństwie do tego w dochodach zamożniejszych warstw społecznych większy – często dominujący – udział mają dochody z kapitału, takie jak udziały w zyskach przedsiębiorstw lub funduszy inwestycyjnych, czy dochody z aktywów finansowych, w postaci akcji lub obligacji. Dochody te rosną wraz z inflacją, co oznacza, że ich realna wartość jest w praktyce chroniona. Ten inflacyjny mechanizm „dzikiej" redystrybucji dochodów od ubogich do bogatych warstw społecznych pogłębia istniejące nierówności dochodowe i majątkowe oraz przyczynia się do dalszej destabilizacji społecznej i politycznej w kraju[28].

Do innych negatywnych skutków gospodarczych inflacji trzeba też zaliczyć wpływ na budżet państwa. Wprawdzie umiarkowana inflacja może obniżyć deficyt budżetowy, ponieważ bieżące dochody, w tym zwłaszcza podatki, rosną szybciej niż ustalone w ustawie budżetowej wydatki, lecz wówczas, gdy inflacja przekracza pewien poziom, sytuacja ulega odwróceniu. W warunkach bardzo wysokiej inflacji uruchamiane są z reguły mechanizmy indeksacyjne powodujące wzrost bieżących wydatków w trakcie roku budżetowego, z kolei opóźnienia w spływie dochodów podatkowych wywołują spadek ich realnej wartości. W efekcie dochody nie nadążają za wydatkami i pojawia się tendencja do wzrostu deficytu budżetowego.

Im wyższa staje się inflacja, tym niższe będą też dochody z emisji pieniądza krajowego (seniorat). Bardzo wysoka inflacja oznacza bowiem ucieczkę od krajowego

[27] Na przykład Stanley Fischer szacuje, że wzrost inflacji o 1 punkt procentowy wywołuje obniżenie tempa wzrostu gospodarczego o ponad 0,1 punktu procentowego (Fischer 1993).

[28] William Easterly i Stanley Fisher (2000) na podstawie badania ankietowego obejmującego ponad 31 tysięcy gospodarstw domowych w 38 krajach dowodzą, że dla osób uboższych inflacja stanowi większe zagrożenie niż dla osób zamożniejszych. Z przeprowadzonego badania wynika, że inflacja jest negatywnie skorelowana z takimi miernikami sytuacji materialnej osób uboższych jak udział dochodów tej grupy w dochodzie narodowym ogółem, zmiany realnej płacy minimalnej i poprawa wskaźnika ubóstwa.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pieniądza oraz zastępowanie go innymi walutami. Spadek popytu na pieniądz krajowy zmniejsza jego obieg i powoduje spadek senioratu (zob. ramka 5.3).

Ramka 5.3. Seniorat i podatek inflacyjny

Senioratem nazywamy różnicę między nominalną wartością pieniądza a kosztem jego wyemitowania. Seniorat stanowi więc dochód z emisji pieniądza. Nazwa pochodzi od starofrancuskiego słowa *seigneurage*, oznaczającego prawo władcy (*seigneur*) do bicia własnej monety. W tradycyjnym znaczeniu pojęcie senioratu odnosiło się do różnicy między nominalną wartością pieniądza kruszcowego, nadaną mu przez władcę, a rzeczywistą – niższą od nominału – zawartością kruszcu (złota, srebra) w bitych monetach. Inaczej mówiąc, był to swoisty podatek pobierany od bicia monet kruszcowych. Na przykład w XIX wieku brytyjski funt szterling zawierał 92,5% czystego srebra, resztę uzupełniano miedzią. Ta różnica, pomniejszona o koszt produkcji monet, stanowiła dochód z emisji pieniądza, czyli seniorat, uzyskiwany przez rządowe mennice.

W rozumieniu współczesnym senioratem nazywamy dochód z emisji pieniądza gotówkowego (banknotów i monet). Bank centralny kupuje na rynku papiery wartościowe (krajowe i zagraniczne), płacąc za nie gotówką. Emisja pieniądza gotówkowego, a zwłaszcza papierowego, nie pociąga za sobą praktycznie żadnych kosztów, bo koszt produkcji pieniądza papierowego jest bardzo niski w stosunku do wartości nominalnej banknotów, a posiadacze gotówki (przedsiębiorstwa, gospodarstwa domowe) nie otrzymują z tytułu jej posiadania żadnego dochodu. Emisję gotówki można więc traktować jako nieoprocentowaną pożyczkę udzielaną bankowi centralnemu przez posiadaczy gotówki. „Pożyczka" ta jest nie tylko nieoprocentowana, ale także bezterminowa, ponieważ banknoty nie mają terminu zapadalności. Gdy banknot ulega zużyciu, jest wymieniany przez bank centralny na nowy banknot o tej samej wartości nominalnej.

Z kolei papiery wartościowe kupione za wyemitowaną gotówkę przez bank centralny przynoszą dochód w postaci odsetek. Różnica między wartością dochodu odsetkowego a kosztem druku banknotów jest dochodem z emisji, który uzyskuje bank centralny. Na przykład w 2015 roku w obiegu gotówkowym w Polsce znajdowało się średnio około 150 mld PLN w postaci banknotów i monet wyemitowanych przez Narodowy Bank Polski. Odpowiednikiem tych zobowiązań NBP po stronie aktywów były oficjalne aktywa rezerwowe, głównie w postaci zagranicznych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

papierów wartościowych. Zakładając, że średnia stopa oprocentowania tych papierów wynosiła około 3%, można oszacować, że dochód z emisji uzyskany przez NBP w 2015 roku wyniósł około 4,5 mld PLN.

Jak łatwo zauważyć, wielkość senioratu zależy od wielkości obiegu gotówkowego. Im więcej gotówki skłonni są przechowywać w portfelach uczestnicy rynku, tym większy dochód z emisji otrzymuje bank centralny. To z kolei zależy od zaufania, jakim cieszy się dany pieniądz, oraz od tego, jak daleko sięga jego obieg. Najwyższe dochody z emisji uzyskują banki centralne krajów, których waluta cieszy się największym zaufaniem w kraju i za granicą. Tradycyjnie taką walutę stanowi dolar amerykański, który – zwłaszcza za granicą – jest traktowany jako środek przechowywania majątku. Według oficjalnych szacunków, w latach 2000–2010 około 36% wyemitowanych banknotów dolarowych znajdowało się poza granicami USA. Urząd Rezerwy Federalnej uzyskiwał w ten sposób dochód z emisji rzędu 25 USD dolarów rocznie (około 0,2% PKB), z czego od posiadaczy zagranicznych wpływało około 7–8 mld USD[29].

Źródłem dochodów z senioratu jest głównie emisja banknotów, ponieważ to w ich przypadku występuje największa różnica między nominalną wartością a kosztem produkcji. W przypadku monet seniorat jest znacznie mniejszy, a często staje się nawet ujemny, gdy koszt produkcji monet przekracza ich wartość nominalną. Na przykład w 2013 roku koszt wybicia monety jednocentowej w USA wynosił 2 centy. Duży dochód z emisji banknotów skłania niektóre kraje do wypuszczania do obrotu banknotów o bardzo dużych nominałach, które są wykorzystywane jako środek przechowywania wartości. Taką rolę pełnią m.in. banknoty o nominale 500 euro emitowane przez EBC i banknoty o nominale 1000 franków szwajcarskich emitowane przez bank centralny Szwajcarii. Trzeba jednak pamiętać, że banknoty o dużych nominałach ułatwiają nierejestrowaną i nielegalną działalność gospodarczą[30], a także zachęcają do fałszowania pieniędzy.

Jak łatwo zauważyć, wzrost emisji pieniądza przekłada się na wzrost senioratu. Dlatego też w przeszłości wiele rządów drukowało pieniądz w nadmiernej ilości, doprowadzając w ten sposób do inflacji. Gdy trudno jest uczestnikom rynku zastąpić krajowy pieniądz alternatywnym środkiem płatniczym (na przykład inną

[29] Zob. Goldberg (2010) oraz Feige (2011).

[30] Dlatego na przykład w USA największym nominałem jest banknot 100-dolarowy, a rząd tego państwa nie zdecydował się na ponowne wprowadzenie do obiegu banknotów o nominale 500 USD.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

walutą), zmuszeni są oni ponosić koszty utraty wartości waluty krajowej[31]. Realna strata na posiadanych zasobach gotówkowych w wyniku inflacji nazywana bywa podatkiem inflacyjnym (*inflation tax*). Gdy pojawiają się jednak możliwości posługiwania się pieniądzem alternatywnym, popyt na pieniądz krajowy drastycznie spada, a wraz z nim zmniejsza się także dochód z emisji.

5.3.7. Epizody hiperinflacji na świecie

Jak wspomnieliśmy wcześniej, inflacja jest szczególnie szkodliwa, gdy ma charakter galopujący. Takie epizody bardzo wysokiej inflacji ekonomiści nazywają hiperinflacją. Umowną granicą, powyżej której inflacja przeradza się w hiperinflację, jest tempo wzrostu cen powyżej 50% w skali miesięcznej[32]. Gdy inflacja osiąga i przekracza ten poziom, spada produkcja, pojawiają się braki towarów, rośnie bezrobocie, władza państwowa traci kontrolę nad obiegiem pieniądza, a gospodarka pogrąża się w chaosie. Hiperinflacja na ogół nie trwa długo. Po kilku miesiącach, a co najwyżej po roku lub dwóch, galopująca inflacja dezorganizuje życie gospodarcze i zmusza władze do przeprowadzenia niezbędnych reform, obejmujących drastyczne oszczędności połączone z reguły z reformą monetarną, czyli z wymianą pieniądza.

Tabela 5.10, zaczerpnięta z opracowania Hanke i Krus (2012), przedstawia listę krajów, w których wystąpiły epizody hiperinflacji, wraz z danymi dotyczącymi różnych wskaźników tego zjawiska. Jak widać, hiperinflacje były dość częstym zjawiskiem w XX wieku, z kolei w bieżącym stuleciu odnotowano, jak dotychczas, tylko jeden przypadek hiperinflacji (Zimbabwe). Najwyższą inflację w historii zaobserwowano na Węgrzech: w lipcu 1946 roku osiągnęła astronomiczny (i niewyobrażalny) poziom 40 tysięcy bilionów procent w skali miesięcznej! Ceny wzrastały w tempie przekraczającym 200% dziennie, co oznaczało, że poziom cen ulegał podwojeniu co 15 godzin. Słynna inflacja niemiecka w latach 20. była wielokrotnie niższa, choć w jej szczytowym momencie ceny nadal rosły w tempie 29 500% miesięcznie i 21% dziennie.

[31] Według niektórych źródeł, podatek inflacyjny stanowił połowę dochodów rządu Zimbabwe w okresie galopującej inflacji w latach 2007–2008 (Gerson 2008).

[32] Autorem tej definicji jest amerykański ekonomista Philip Cagan (Cagan 1956).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 5.10. Epizody hiperinflacji na świecie

	Kraj	Okres trwania		Miesiąc szczytowy	Najwyższa inflacja Miesięczna (dane w %)	Najwyższa inflacja dzienna (dane w %)	Czas podwojenia cen	Waluta	Miernik inflacji[a]
		od	do						
1.	Węgry	08.1945	07.1946	07.1946	$4{,}19 \times 10^{16}$	207,0	15,0 godz.	pengö	CPI
2.	Zimbabwe[b]	03.2007	11.2008	11.2008	$7{,}96 \times 10^{10}$	98,0	24,7 godz.	dollar	kurs walut
3.	Jugosławia	04.1992	01.1994	01.1994	$313{,}0 \times 10^{6}$	64,6	1,41 dni	dinar	CPI
4.	Republika Serbska[c]	04.1992	01.1994	01.1994	$297{,}0 \times 10^{6}$	64,3	1,41 dni	dinar	CPI
5.	Niemcy	08.1922	12.1923	10.1923	29 500	20,9	3,70 dni	marka	WPI
6.	Grecja	05.1941	12.1945	10.1944	13 800	17,9	4,27 dni	drachma	kurs walut
7.	Chiny	10.1947	04.1949	04.1949	5 070	14,1	5,34 dni	yuan	WPI
8.	Wolne Miasto Gdańsk	08.1922	10.1923	09.1923	2 440	11,4	6,52 dni	marka niemiecka	kurs walut
9.	Armenia	10.1993	12.1994	11.1993	438	5,77	12,5 dni	dram i rubel ros.	CPI
10.	Turkmenistan	01.1992	11.1993	11.1993	429	5,71	12,7 dni	manat	CPI
11.	Taiwan	08.1945	09.1945	08.1945	399	5,50	13,1 dni	yen	CPI
12.	Peru	07.1990	08.1990	08.1990	397	5,49	13,1 dni	inti	CPI
13.	Bośnia i Hercegowina	04.1992	06.1993	06.1992	322	4,92	14,6 dni	dinar	CPI
14.	Francja	05.1795	11.1796	08.1796	304	4,77	15,1 dni	mandat	kurs walut
15.	Chiny	07.1943	08.1945	06.1945	302	4,75	15,2 dni	yuan	WPI
16.	Ukraina	01.1992	11.1994	01.1992	285	4,60	15,6 dni	rubel ros.	CPI
17.	Polska	01.1923	01.1924	10.1923	275	4,50	16,0 dni	marka	WPI
18.	Nikaragua	06.1986	03.1991	03.1991	261	4,37	16,4 dni	córdoba	CPI
19.	Kongo (Zair)	11.1993	09.1994	11.1993	250	4,26	16,8 dni	zaïre	CPI

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

cd. tab. 5.10

	Kraj	Okres trwania		Miesiąc szczytowy	Najwyższa inflacja Miesięczna (dane w %)	Najwyższa inflacja dzienna (dane w %)	Czas podwojenia cen	Waluta	Miernik inflacji[a]
		od	do						
20.	Rosja	01.1992	01.1992	01.1992	245	4,22	17,0 dni	rubel	CPI
21.	Bułgaria	02.1997	02.1997	02.1997	242	4,19	17,1 dni	lew	CPI
22.	Mołdawia	01.1992	12.1993	01.1992	240	4,16	17,2 dni	rubel ros.	CPI
23.	Rosja/ZSSR	01.1922	02.1924	02.1924	212	3,86	18,5 dni	rubel	CPI
24.	Gruzja	09.1993	09.1994	09.1994	211	3,86	18,6 dni	kupon	CPI
25.	Tadżykistan	01.1992	10.1993	01.1992	201	3,74	19,1 dni	rubel ros.	CPI
26.	Gruzja	03.1992	04.1992	03.1992	198	3,70	19,3 dni	rubel ros.	CPI
27.	Argentyna	05.1989	03.1990	07.1989	197	3,69	19,4 dni	austral	CPI
28.	Boliwia	04.1984	09.1985	02.1985	183	3,53	20,3 dni	boliviano	CPI
29.	Białoruś	01.1992	02.1992	01.1992	159	3,22	22,2 dni	rubel ros.	CPI
30.	Kirgistan	01.1992	01.1992	01.1992	157	3,20	22,3 dni	rubel ros.	CPI
31.	Kazachstan	01.1992	01.1992	01.1992	141	2,97	24,0 dni	rubel ros.	CPI
32.	Austria	10.1921	10.1922	08.1922	129	2,80	25,5 dni	korona	CPI
33.	Bułgaria	02.1991	03.1991	02.1991	123	2,71	26,3 dni	lew	CPI
34.	Uzbekistan	01.1992	02.1992	01.1992	118	2,64	27,0 dni	rubel ros.	CPI
35.	Azerbejdżan	01.1992	12.1994	01.1992	118	2,63	27,0 dni	rubel ros	CPI
36.	Kongo (Zair)	10.1991	09.1992	11.1991	114	2,57	27,7 dni	zaïre	CPI
37.	Peru	09.1988	09.1988	09.1988	114	2,57	27,7 dni	inti	CPI
38.	Taiwan	10.1948	05.1949	10.1948	108	2,46	28,9 dni	taipi	WPI
39.	Węgry	03.1923	02.1924	07.1923	97,9	2,30	30,9 dni	korona	CPI
40.	Chile	10.1973	10.1973	10.1973	87,6	2,12	33,5 dni	escudo	CPI
41.	Estonia	01.1992	02.1992	01.1992	87,2	2,11	33,6 dni	rubel ros.	CPI
42.	Angola	12.1994	01.1997	05.1997	84,1	2,06	34,5 dni	kwanza	CPI

Plik zabezpieczony watermarkiem jawnym i niejawnym: 80981 23A35373237

cd. tab. 5.10

	Kraj	Okres trwania		Miesiąc szczytowy	Najwyższa inflacja Miesięczna (dane w %)	Najwyższa inflacja dzienna (dane w %)	Czas podwojenia cen	Waluta	Miernik inflacji[a]
		od	do						
43.	Brazylia	12.1989	03.1990	03.1990	82,4	2,02	35,1 dni	cruzado i cruzeiro	CPI
44.	Dem. Rep. Kongo	08.1998	08.1998	08.1998	78,5	1,95	36,4 dni	franc	CPI
45.	Polska	10.1989	01.1990	01.1990	77,3	1,93	36,8 dni	złoty	CPI
46.	Armenia	01.1992	02.1992	01.1992	73,1	1,85	38,4 dni	rubel ros.	WPI
47.	Tadżykistan	10.1995	11.1995	11.1995	65,2	1,69	42,0 dni	rubel	WPI
48.	Łotwa	01.1992	01.1992	01.1992	64,4	1,67	42,4 dni	rubel ros.	CPI
49.	Turkmenistan	11.1995	01.1996	01.1996	62,5	1,63	43,4 dni	manat	CPI
50.	Filipiny	01.1944	12.1944	01.1944	60,0	1,58	44,9 dni	JWN	CPI
51.	Jugosławia	09.1989	12.1989	12.1989	59,7	1,57	45,1 dni	dinar	CPI
52.	Niemcy	01.1920	01.1920	01.1920	56,9	1,51	46,8 dni	marka	WPI
53.	Kazachstan	11.1993	11.1993	11.1993	55,5	1,48	47,8 dni	tenge i rubel ros.	CPI
54.	Litwa	01.1992	01.1992	01.1992	54,0	1,45	48,8 dni	rubel ros.	CPI
55.	Białoruś	08.1994	08.1994	08.1994	53,4	1,44	49,3 dni	rubel	CPI
56.	Taiwan	02.1947	02.1947	02.1947	50,8	1,38	51,4 dni	taipi	WPI

Uwaga: co najmniej jeden przypadek stwierdzonej hiperinflacji nie został uwzględniony w tabeli ze względu na brak dostępu do systematycznie gromadzonej bazy danych. Dotyczy on Koreańskiej Republiki Ludowo-Demokratycznej, gdzie, zgodnie z szacunkami autorów, w okresie od grudnia 2009 roku do stycznia 2011 roku inflacja wynosiła 496% w skali miesięcznej (6,13% w ciągu dnia).

a) CPI–Consumer Price Index, WPI – Wholesale Price Index.

b) Inflacja w przypadku Zimbabwe obliczona w sposób pośredni, na podstawie zmian cen akcji notowanych zarówno na giełdzie w Harare, jak i w Londynie.

c) Republika Serbska emitowała własną walutę, dinar, w latach 1992–1994.

Źródło: Hanke i Krus 2012.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 80981 23A35373237

Największa liczba zarejestrowanych przypadków hiperinflacji (dwadzieścia jeden) dotyczy krajów byłego Związku Radzieckiego w okresie bezpośrednio następującym po jego rozpadzie w 1991 roku. Ponieważ w momencie likwidacji ZSRR państwa te posługiwały się rublem rosyjskim, odpowiedzialność za występowanie w tych krajach hiperinflacji ponosił bank centralny Federacji Rosyjskiej, który teoretycznie kontrolował podaż rubli (w praktyce o podaży pieniądza rezerwowego decydowały potrzeby finansowania wydatków rządowych). Siedem kolejnych przypadków odnosi się do krajów europejskich z okresu po I wojnie światowej, a pięć epizodów odnotowano w czasie II wojny światowej i bezpośrednio po niej. Zdecydowana większość zarejestrowanych przypadków hiperinflacji (trzydzieści osiem) miała miejsce w krajach europejskich i krajach byłego Związku Radzieckiego, sześć przypadków wystąpiło w Azji, pięć w krajach afrykańskich oraz siedem w krajach Ameryki Środkowej i Południowej. W Polsce hiperinflacja wystąpiła dwukrotnie – w 1923 roku oraz w początkowym okresie transformacji ustrojowej na przełomie lat 1989–1990, przy czym ten pierwszy epizod był znacznie poważniejszy.

5.3.8. Dlaczego deflacja jest szkodliwa?

Czy spadek cen może być szkodliwy? Przecież jeśli towary tanieją, możemy nabywać ich więcej za daną kwotę pieniędzy, a to oznacza wzrost konsumpcji i poprawę dobrobytu. Tak rzeczywiście sytuacja może wyglądać z punktu widzenia pojedynczego konsumenta. Lecz z perspektywy całej gospodarki deflacja, czyli spadek ogólnego poziomu cen, jest zjawiskiem bardzo groźnym i – w przeciwieństwie do inflacji – znacznie trudniejszym do opanowania.

Typową reakcją przedsiębiorstw na spadek cen jest początkowo wstrzymanie nowych inwestycji, następnie ograniczanie produkcji i zamówień, a wreszcie zwalnianie pracowników. Pociąga to za sobą spadek zagregowanego popytu w gospodarce i przyczynia się do dalszego spadku cen. Z kolei gospodarstwa domowe w takich warunkach ograniczają swoje wydatki, zwłaszcza te dotyczące dóbr trwałego użytku, ponieważ liczą na dalszy spadek cen („jutro będzie taniej"). Również rosnące bezrobocie skłania do zmniejszania wydatków – zarówno ze względu na spadek bieżących dochodów, jak i na zwiększone ryzyko utraty pracy w przyszłości. Powoduje to, podobnie jak w przypadku przedsiębiorstw, spadek zagregowanego popytu i dalszy spadek cen. Dodatkowym czynnikiem hamującym popyt jest wzrost realnych stóp procentowych wywołany spadkiem cen – im głębsza będzie deflacja, tym wyższe staną się realne stopy procentowe przy danych nominalnych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

poziomach stóp procentowych. Spadek popytu wywołuje dalszy spadek produkcji, a to z kolei pociąga za sobą dalszy spadek popytu. Amerykański ekonomista Paul Krugman twierdził, że gospodarka nie jest w stanie wyjść o własnych siłach z tej swoistej „spirali" zwijania się, i przyrównał taką sytuację do astrofizycznego zjawiska „czarnej dziury" (*black hole*).

Deflacja staje się szczególnie niebezpieczna w przypadki gospodarek o wysokim poziomie zadłużenia, zwłaszcza w zakresie sektora prywatnego. Spadek cen i dochodów powoduje wzrost realnego obciążenia długiem, ponieważ relacja długu do dochodu rośnie. W warunkach rosnącego udziału obsługi długu w dochodzie następuje dalsze ograniczenie wydatków, które w konsekwencji zmniejsza produkcję i zatrudnienie. W ten sposób zadłużenie działa procyklicznie – w czasie recesji skłania do obniżania wydatków, co oczywiście pogłębia kryzys. Zjawisko to z wielką siłą wystąpiło w czasie Wielkiego Kryzysu w latach 30., gdy ciężar zadłużenia zmuszał dłużników do drastycznego zaciskania pasa w czasie deflacji. Wybitny amerykański ekonomista Irving Fisher nazwał ten mechanizm „deflacją długu" (*debt deflation*).

Dynamika deflacji jest niższa niż w przypadku inflacji, lecz przezwyciężyć jest ją o wiele trudniej niż inflację (*vide* stracone dwie dekady w Japonii). O ile poskromienie inflacji wymaga ograniczenia podaży pieniądza, co można przeprowadzić właściwie z dnia na dzień (jeśli tylko pojawi się wola polityczna, aby to uczynić), to pokonanie deflacji wymaga zmiany oczekiwań i zachowań wśród uczestników rynku, co wymaga czasu i zastosowania wiarygodnej polityki. Szczególnie trudna staje się sytuacja banków centralnych. Nie są one w stanie obniżyć realnych stóp procentowych do poziomu zapewniającego równowagę popytu i podaży, ponieważ w warunkach spadających cen nawet obniżenie nominalnych stóp do zera pozostawia realne stopy wciąż na dodatnim poziomie i nie gwarantuje przywrócenia równowagi[33].

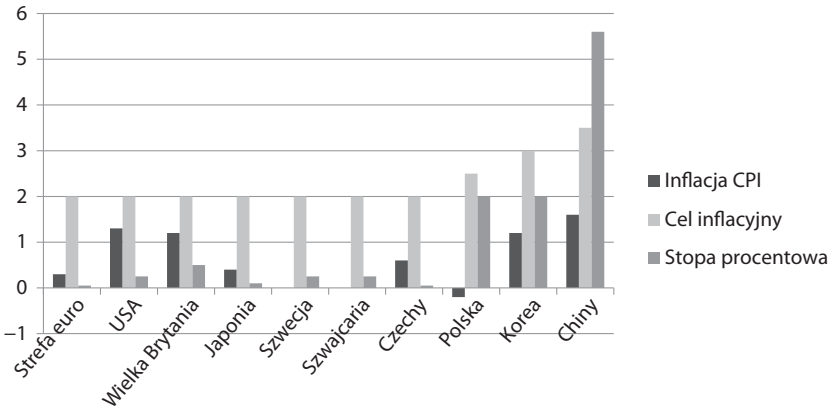
Zbliżone skutki niesie także bardzo niska inflacja, wyraźnie niższa od celu inflacyjnego, jaki stawia sobie bank centralny. Tak zwana *lowflation* jest niebezpieczna, ponieważ może szybko przerodzić się w regularną deflację. Ponadto w sytuacji *lowflation* mamy z reguły do czynienia ze spadkiem cen produktów przemysłowych, przy niewielkim wzroście ceny usług, co wywołuje tendencje recesyjne w przemyśle i wzrost bezrobocia. Zjawisko *lowflation* pojawiło się w 2014 roku, kiedy w wielu

[33] Nie wszyscy ekonomiści obawiają się deflacji. Np. Cochrane (2014) uważa, że umiarkowana deflacja rzędu 0–2% nie jest szkodliwa, a „spirala deflacyjna" nigdy w praktyce nie wystąpiła nawet w Japonii.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

krajach rozwiniętych główne wskaźniki inflacji znajdowały się wyraźnie poniżej przyjętych celów inflacyjnych (zob. wykres 5.8). Jednocześnie wśród większości tych państw stopy procentowe banków centralnych były już obniżone do poziomu bliskiego zeru.

Wykres 5.8. Zjawisko *lowflation*: inflacja i cel inflacyjny w wybranych krajach – stan z października 2014 roku (dane w %) (r/r)



Źródło: opracowanie własne na podstawie danych Eurostatu i banków centralnych poszczególnych krajów.

5.3.9. Cel nadrzędny (ostateczny) i cele pośrednie

Jak już wiadomo, nadrzędnym celem polityki pieniężnej pozostaje najczęściej utrzymanie stabilności cen. Lecz w warunkach gospodarki rynkowej bank centralny nie wyznacza poziomu cen – kształtują się one pod wpływem zmian podaży i popytu na poszczególnych rynkach. Bank centralny staje więc przed niełatwym zadaniem – ma zapewnić stabilność cen, mimo że nie ma na nie bezpośredniego wpływu. Musi więc realizować swój główny cel za pośrednictwem innych zmiennych – takich, które może bezpośrednio kontrolować, a które zarazem mają wpływ na inflację. Najbardziej naturalnym sposobem wpływania na inflację za pośrednictwem innej zmiennej jest regulowanie ilości pieniądza w obiegu.

Przywoływane już wcześniej stwierdzenie Miltona Friedmana, że „inflacja jest zawsze i wszędzie zjawiskiem pieniężnym" (Friedman 1970, s. 42), nie wyjaśnia wprawdzie ekonomicznych przyczyn inflacji, lecz wskazuje na jej kluczową cechę – że mianowicie trwały wzrost cen może mieć miejsce jedynie w warunkach rosnącej podaży pieniądza. W związku z tym utrzymanie stabilności cen wymaga

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

odpowiedniego kształtowania wielkości pieniądza w obiegu. Bank centralny nie ustala bezpośrednio cen, lecz jest w stanie kontrolować podaż pieniądza. Oznacza to, że utrzymanie tempa wzrostu podaży pieniądza na określonym poziomie może być celem pośrednim, którego realizacja pozwala bankowi centralnemu osiągać swój cel nadrzędny, ostateczny (*ultimate goal*), jakim jest „stabilność cen". Warunkiem skuteczności prowadzenia takiej polityki jest rozpoznanie związku, jaki zachodzi między celem pośrednim (podaż pieniądza) a celem ostatecznym (poziom inflacji). Bank centralny sięga do historycznych danych statystycznych, aby rozpoznać i określić rodzaj oraz charakter zależności występujących między zmianami podaży pieniądza a zmianami inflacji, a następnie wybiera jako cel pośredni zmiany dokonujące się w agregacie pieniężnym, najsilniej skorelowane ze zmianami poziomu cen. Najczęściej banki centralne ustalają jako cel pośredni zmianę najszerszego agregatu M3, rzadziej zmiany dokonujące się w zakresie pieniądza rezerwowego M0.

Z kolei, aby utrzymać określone tempo wybranego agregatu pieniężnego, spójne z celem stabilności cen, bank centralny stosuje różne instrumenty, takie jak krótkookresowa stopa procentowa określająca „cenę" pieniądza na rynku pieniężnym (regulacja cenowa) lub określona skala operacji zasilających banki w płynność (regulacja ilościowa). Wówczas pożądany poziom krótkookresowych stóp procentowych lub pożądana wielkość operacji otwartego rynku, które zapewniają utrzymanie podaży pieniądza na określonym poziomie, mogą stać się kolejnymi celami pośrednimi w polityce pieniężnej.

W praktyce wybór między regulacją ilościową a regulacją cenową zależy głównie od tego, która ze zmiennych – ilość pieniądza czy stopa procentowa – daje się łatwiej mierzyć i kontrolować. Jeśli chodzi o ilość pieniądza, w danym momencie bank centralny może jedynie szacować wielkość szerszych agregatów na podstawie posiadanych danych o wielkości wąskiego agregatu M0 oraz odpowiednich mnożników. Niedokładne szacunki oznaczają, że wielkość M2 lub M3 w stosunku do popytu może być za duża lub za mała, a to może wywołać wahania rynkowych stóp procentowych i w konsekwencji niepożądane zmiany PKB i inflacji[34]. Z kolei regulacja cenowa wymaga dobrego pomiaru realnej stopy procentowej, co wymaga zaś dobrego oszacowania oczekiwanej inflacji. Gdy oczekiwania inflacyjne zmieniają się, spotyka to także realne stopy procentowe, co wywołuje niepożądane wahania

[34] Wahania M2, M3 wywołują zmiany położenia krzywej LM, co przy danym położeniu krzywej IS powoduje wahania stóp procentowych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

PKB i pociąga za sobą zmiany w zakresie inflacji[35]. Wybór zależy więc od skali możliwego błędu: jeśli szacunki ilości pieniądza w obiegu są zmienne i niepewne, lepiej stosować regulację cenową (ustalanie stóp procentowych), z kolei jeśli realne stopy procentowe ulegają dużym wahaniom na skutek zmiennych oczekiwań inflacyjnych, lepiej stosować regulację ilościową (ustalanie podaży pieniądza)[36].

Wielkość podaży pieniądza można kształtować również za pomocą ustalania wielkości kredytu nominalnego w gospodarce. Wówczas celem pośrednim polityki pieniężnej staje się tempo wzrostu kredytu dla gospodarki. W niektórych systemach polityki pieniężnej, zwłaszcza w małych otwartych gospodarkach, celem pośrednim jest utrzymanie określonego poziomu kursu walutowego. W warunkach stałego kursu waluty krajowej wobec stabilnej waluty międzynarodowej nabywcy krajowi zawsze mają możliwość zwiększenia zakupów z importu, jeśli dostawcy krajowi podniosą ceny (arbitraż cenowy). W tych okolicznościach wystąpienie w kraju inflacji znacząco wyższej od inflacji za granicą jest po prostu niemożliwe tak długo, jak kurs pozostaje stały. Wówczas stały kurs staje się instrumentem pozwalającym osiągnąć cel nadrzędny, czyli stabilność cen. Ogólnie rzecz biorąc, taki sposób prowadzenia polityki pieniężnej, który polega na wpływaniu na inflację za pośrednictwem zmiennych, kontrolowanych przez bank centralny w sposób bezpośredni, nazywa się strategią (lub regułą) celu pośredniego. Krótkie podsumowanie celów pośrednich stosowanych w polityce pieniężnej zawiera tabela 5.11.

Wykorzystanie celów pośrednich w polityce pieniężnej zakłada – po pierwsze – wysoką zdolność banku centralnego do bieżącego kontrolowania zmiennej będącej celem pośrednim (podaż pieniądza, krótkookresowa stopa procentowa, kurs walutowy), a po drugie, znajomość związku pomiędzy wybranym celem pośrednimi a celem ostatecznym. Jeśli celem pośrednim jest jeden z szerszych agregatów pieniężnych (M1, M2, M3), bank centralny musi być w stanie kształtować odpowiednio zmiany tego agregatu przy pomocy zmian podaży pieniądza rezerwowego (M0). Znajomość wartości i ocena stabilności mnożników pieniężnych okazuje się na tym etapie kluczowa. Z kolei związek między zmianami agregatów pieniężnych

[35] Wahania realnej stopy procentowej wywołują zmiany położenia krzywej IS, co przy danym położeniu krzywej LM powoduje wahania poziomu produkcji.

[36] Polityka pieniężna Rezerwy Federalnej w USA w drugiej połowie XX wieku dostarcza przykładów zmian w celach pośrednich. W latach 50. i 60., w warunkach niskiej i stabilnej inflacji, Fed stosował regulację cenową, wykorzystując jako cel pośredni krótkookresową stopę procentową. W 70., gdy inflacja była wysoka, a realna stopa procentowa niepewna, bank centralny przeszedł na bezpośrednie ustalanie wielkości podaży pieniądza. Po 1980 roku, gdy inflacja została obniżona, Fed powrócił do ustalania stopy procentowej (Hall i Taylor 2000, s. 383–386).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

a inflacją jest określony funkcją popytu na pieniądz. Rozpoznanie i oszacowanie tej funkcji stanowi więc warunek skuteczności polityki pieniężnej wykorzystującej cel pośredni.

Tabela 5.11. Cele pośrednie w polityce pieniężnej

Cel pośredni	Działanie banku centralnego (BC)
Baza monetarna (M0)	(BC) ustala wielkość bazy monetarnej M0, stopa procentowa dostosowuje się przy danym popycie banków i ludzi
Agregaty pieniężne M1, M2, M3	BC wpływa na wielkość M1, M2 lub M3 poprzez zmiany wielkości bazy monetarnej i/lub stopy procentowej
Wielkość kredytu krajowego	BC utrzymuje tempo wzrostu kredytu nominalnego w określonych granicach (przy pomocy stóp procentowych, stopy rezerwy obowiązkowej i limitów kredytowych)
Krótkookresowa stopa procentowa	BC ustala poziom krótkookresowej stopy procentowej, podaż M0 dostosowuje się przy danym popycie
Kurs walutowy	BC ustala i broni pożądanego poziomu kursu walutowego, ilość pieniądza krajowego staje się zmienną endogeniczną

Źródło: opracowanie własne.

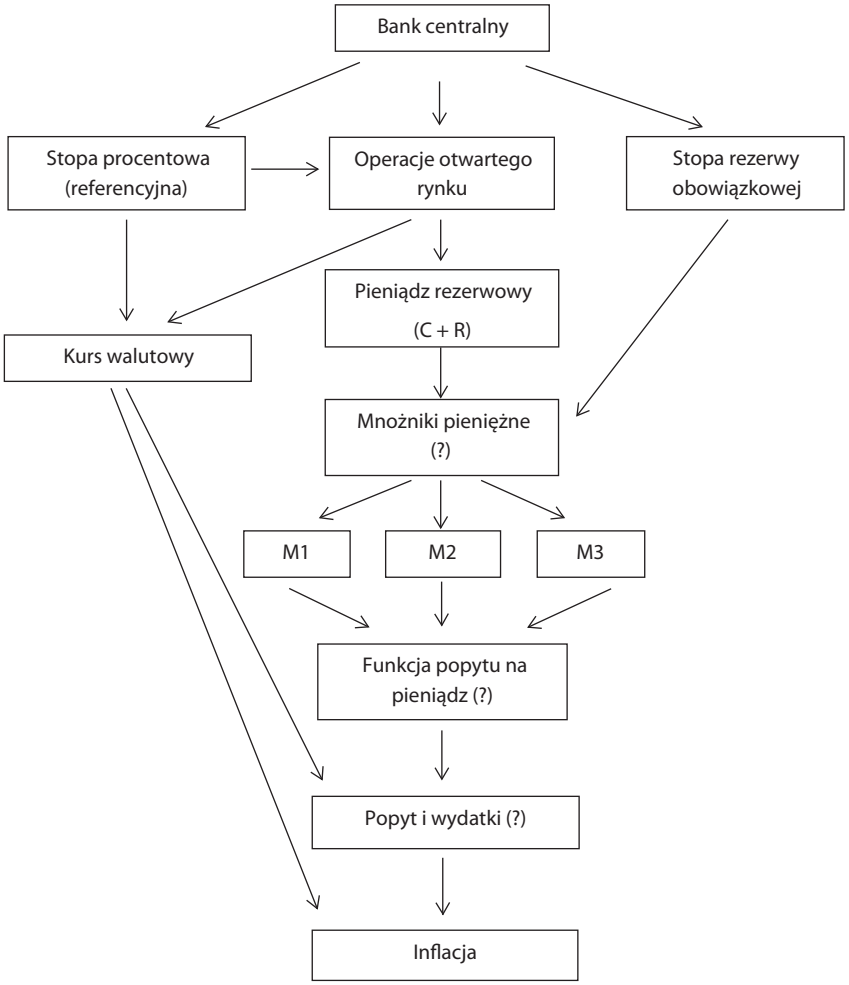
Rysunek 5.2 pokazuje, że utrzymanie stabilności cen nie jest łatwym zadaniem. Między decyzjami banku centralnego a zmianami inflacji pojawia się wiele ogniw pośrednich. Ich rozpoznanie jest warunkiem skuteczności transmisji impulsów polityki pieniężnej. Znaki zapytania umieszczone w diagramie wskazują na te ogniwa procesu transmisji, w przypadku których występuje szczególnie duża niepewność co do rzeczywistego charakteru relacji między poszczególnymi zmiennymi. Dotyczy to zwłaszcza wielkości mnożników pieniężnych oraz postaci funkcjonalnej i wartości parametrów funkcji popytu na pieniądz. Duża zmienność tych elementów utrudnia prawidłowe prowadzenie polityki pieniężnej.

Bank centralny sprawuje bezpośrednią kontrolę tylko nad ilością pieniądza rezerwowego (bazy monetarnej), który sam tworzy, udostępniając go bankom komercyjnym i regulując jego podaż za pomocą stopy procentowej, kwot operacji otwartego rynku i innych operacji na rynku pieniężnym. Nie decyduje on jednak o tym, jaka część pieniądza rezerwowego przyjmuje postać gotówki – zależy to od popytu na gotówkę, zgłaszanego przez uczestników rynku. Co więcej, bank centralny ma tylko pośrednią i niepełną kontrolę nad ilością pieniądza depozytowego (pieniądz banków komercyjnych), którą kształtuje przy pomocy zmian stóp procentowych i stopy rezerwy obowiązkowej. Jak pamiętamy, całkowita ilość pieniądza depozytowego, przy danej ilości pieniądza rezerwowego, zależy od (nieznanego

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w określonym momencie) mnożnika, który jest funkcją stopy rezerwy i udziału gotówki w obiegu pieniężnym (zob. punkt 5.2). Z kolei inflacja, przy danej podaży, ma związek z poziomem popytu i rzeczywistych wydatków na dobra i usługi, a te zależą od (nieznanej w określonym momencie) ilości pieniądza oraz (nieznanej wówczas) funkcji popytu na pieniądz.

Rysunek 5.2. Oddziaływanie banku centralnego na inflację



Źródło: opracowanie własne.

W latach 50. i 60. XX wieku wykorzystywanie podaży pieniądza jako celu pośredniego w polityce pieniężnej było bardzo popularne wśród banków centralnych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ilość pieniądza znajdującego się w obiegu stanowiła dobry wskaźnik wyprzedzający w przypadku inflacji. W następnych latach związek ten osłabł jednak, częściowo wskutek stopniowego znoszenia barier i liberalizacji transgranicznych przepływów kapitału, a częściowo także pod wpływem rozwoju nowych instrumentów finansowych, w tym instrumentów pochodnych. W efekcie w latach 90. wiele banków centralnych odeszło od reguły polityki pieniężnej polegającej na ścisłej kontroli podaży pieniądza.

Komplikacje związane ze stosowaniem celów pośrednich, w tym zwłaszcza rosnąca niepewność co do wielkości mnożników pieniężnych i kształtu funkcji popytu na pieniądz, sprawiają, że banki centralne poszukują alternatywnych sposobów prowadzenia polityki pieniężnej w sposób zapewniający skuteczniejszą realizację celu inflacyjnego. Metodą, która zyskała dużą popularność w okresie ostatnich dwudziestu lat, jest tzw. strategia (reguła) bezpośredniego celu inflacyjnego (*Direct Inflation Targeting* – DIT). Polega ona na dostosowaniu bieżącej polityki pieniężnej do różnicy między prognozowaną inflacją a ustalonym wcześniej celem inflacyjnym. W ramach tej strategii bank centralny rezygnuje więc z posługiwania się jakimkolwiek celem pośrednim i stara się realizować cel inflacyjny w sposób bezpośredni. Szczegółowe omówienie strategii DTI znajduje się w dalszej części rozdziału.

5.4. Instrumenty polityki pieniężnej

5.4.1. Mechanizm zmian podaży pieniądza przez bank centralny

Przypomnijmy, że bank centralny emituje jedynie pieniądz rezerwowy, z kolei banki komercyjne tworzą pieniądz depozytowy, który stanowi zdecydowaną większość podaży pieniądza ogółem. Aby mieć więc kontrolę nad całością podaży pieniądza, banki centralne musiały więc znaleźć sposób, aby kontrolować proces tworzenia depozytów. Tym sposobem jest tzw. system ułamkowej rezerwy (*fractional reserve banking*), który umożliwia bankom utrzymywanie tylko części posiadanych depozytów w postaci pieniądza rezerwowego (gotówka lub środki na rachunku w banku centralnym), tak aby zapewnić sobie zdolność do dokonywania bieżących wypłat. Relacja tej części depozytu, która pozostaje w banku w postaci pieniądza rezerwowego do całego depozytu, nazywana jest stopą rezerwy (obowiązkowej lub dobrowolnej – zob. punkt 5.2 powyżej). Gdy bank komercyjny od każdego utworzonego

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

depozytu musi zatem odprowadzić rezerwę w wysokości określonej stopą rezerwy, tempo tworzenia depozytów będzie zależeć od tego, ile pieniądza rezerwowego posiada bank komercyjny. W ten sposób, regulując wielkość pieniądza rezerwowego znajdującego się w posiadaniu banków komercyjnych, bank centralny – przy danej stopie rezerwy – wpływa na wielkość tworzonych przez banki depozytów i – w konsekwencji – na podaż pieniądza w gospodarce (Sławiński 2011).

Bank komercyjny tworzy depozyt w dwóch różnych sytuacjach. Pierwsza z nich ma miejsce, gdy osoba lub firma dokonuje wpłaty gotówkowej na swój rachunek w danym banku. Wówczas wzrostowi stanu środków na rachunku bankowym klienta towarzyszy spadek gotówki w obiegu o dokładnie tę samą kwotę i w efekcie ogólna podaż pieniądza nie ulega zmianie. Druga – znacznie bardziej interesująca – sytuacja ma miejsce wtedy, gdy bank udziela osobie lub firmie kredytu. Na rachunku kredytobiorcy zapisuje się wówczas kwotę udzielonego kredytu. W ten sposób powstaje depozyt, który zwiększa ogólną podaż pieniądza. To właśnie w tej drugiej sytuacji banki tworzą nowy pieniądz – depozytowy.

Gdy na rachunku klienta powstaje jednak depozyt, bank komercyjny musi od tego depozytu odprowadzić na rachunek w banku centralnym wymaganą rezerwę w wysokości określonej stopą rezerwy obowiązkowej. W tym celu bank komercyjny potrzebuje pewnej ilości pieniądza rezerwowego, czyli gotówki, przy czym może pozyskać go zasadniczo z dwóch źródeł: poprzez zaciągnięcie pożyczki (od innych banków lub od banku centralnego) lub sprzedaż części posiadanych aktywów (np. papierów wartościowych). Gdy bank komercyjny pożycza płynne środki od innego banku lub sprzedaje aktywa innym uczestnikom rynku, ilość pieniądza rezerwowego w gospodarce nie ulega zmianie. Lecz w sytuacji, gdy bank komercyjny pożycza płynne środki w banku centralnym albo sprzedaje mu aktywa w zamian za pieniądz rezerwowy, ilość pieniądza rezerwowego w gospodarce zwiększa się.

Bank centralny dostarcza więc rezerwy do banków poprzez nabywanie od nich papierów wartościowych i innych aktywów (np. walut obcych) lub udzielanie pożyczek. Dzięki temu, regulując skalę zarówno tych zakupów, jak i udzielanych pożyczek, bank centralny ma pełną kontrolę nad podażą pieniądza rezerwowego. Wielkość tych operacji wyznacza górny pułap tworzenia nowych depozytów, a zatem określa maksymalne tempo wzrostu podaży pieniądza.

Jak wspomnieliśmy, tworzenie nowych depozytów dokonuje się przede wszystkim w wyniku udzielania kredytów przez banki komercyjne. Przyznanie kredytu oznacza stworzenie depozytu, a stworzenie depozytu wymaga od banku odprowadzenia rezerwy. W praktyce oznacza to, że bank komercyjny musi posiadać pewien

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zasób pieniądza rezerwowego, aby udzielić kredytu i stworzyć depozyt. Proces tworzenia depozytów obejmuje więc dwa etapy: najpierw bank centralny dostarcza bankom komercyjnym rezerwy, a dopiero potem banki komercyjne powiększają kredyty i depozyty.

Popyt na rezerwy ze strony banków jest więc określony przez skalę ich akcji kredytowej. Bank centralny zapewnia podaż rezerw, lecz tylko na taką skalę, jaką uważa za niezbędną dla realizacji swych statutowych celów (stabilność cen). Gdy popyt na rezerwy jest w opinii banku centralnego zbyt duży, dostosowuje on ten popyt do określonej podaży pieniądza rezerwowego, podnosząc stopy procentowe i ograniczając wielkości zakupu aktywów od banków. Gdy popyt na rezerwy jest zaś w opinii banku centralnego zbyt mały, obniża on stopy procentowe i zwiększa zakupy aktywów.

Tabela 5.12 przedstawia zmiany w bilansie banku komercyjnego, które ilustrują funkcjonowanie mechanizmu regulowania podaży pieniądza. Przyjmijmy, że bank chce udzielić kredytu swojemu klientowi w wysokości 1000 PLN, lecz nie posiada wolnych zasobów pieniądza rezerwowego. Załóżmy też, że stopa rezerwy obowiązkowej wynosi 10%. Bank sprzedaje więc bankowi centralnemu obligacje ze swojego portfela aktywów za kwotę 100 PLN. W ten sposób w bilansie banku jeden rodzaj aktywów (papiery wartościowe) zamieniony zostaje na inny rodzaj aktywów (pieniądze na rachunku w banku centralnym). Ten pierwszy etap operacji ukazany jest w lewej części tabeli 5.12. Następnie bank udziela kredytu swojemu klientowi na kwotę 1000 PLN. Decyzja o przyznaniu kredytu oznacza, że bank uznaje rachunek klienta na kwotę 1000 PLN, tworząc depozyt. Jednocześnie bank przekazuje rezerwę obowiązkową w wysokości 10% od sumy nowego depozytu, czyli w kwocie 100 PLN, do banku centralnego. Ten drugi etap operacji ukazany został w prawej części tabeli 5.12.

Tabela 5.12. Mechanizm tworzenia pieniądza depozytowego przez bank komercyjny (dane w PLN)

Etap I		Etap II	
Aktywa	Pasywa	Aktywa	Pasywa
Obligacje: –100 Pieniądz rezerwowy na rachunku banku w banku centralnym: +100 Nowy kredyt: 0	Nowy depozyt: 0	Obligacje: –100 Rezerwa obowiązkowa: +100 Nowy kredyt: +1000	Nowy depozyt: +1000

Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Z powyższego przykładu płyną dwa ważne wnioski. Po pierwsze, to decyzja banku komercyjnego o udzieleniu kredytu powoduje powstanie depozytu. Kredyt jest zatem źródłem tworzenia (kreacji) pieniądza depozytowego, a nie odwrotnie. Ogólnie, w skali całej gospodarki, akcja kredytowa banków prowadzi do tworzenia depozytów, czyli określa tempo wzrostu podaży pieniądza szerokiego agregatu. Jeśli więc kredyt jest źródłem kreacji pieniądza depozytowego, bank centralny, chcąc wpływać na wielkość podaży pieniądza depozytowego, reguluje dostęp do kredytu poprzez zmiany krótkookresowej stopy procentowej. Określają one popyt na kredyty, z kolei popyt decyduje o podaży kredytów, a tworzone przez nie depozyty wyznaczają zapotrzebowanie na płynne rezerwy banków, czyli na pieniądz rezerwowy (za pośrednictwem mechanizmu rezerwy obowiązkowej). W ten sposób zmiany stopy procentowej określają popyt banków na pieniądz rezerwowy. Po drugie, udzielenie kredytu byłoby niemożliwe, gdyby bank nie posiadał środków płynnych w wysokości co najmniej wymaganej rezerwy obowiązkowej. Tak więc bez udziału pieniądza rezerwowego cała operacja udzielenia kredytu nie mogłaby dojść do skutku. Dlatego to właśnie pieniądz rezerwowy nazywany bywa czasem „pieniądzem wielkiej mocy" (*high-powered money*) – jego zmiany silnie wpływają zarówno na podaż całego pieniądza (M2, M3), jak i na krótkookresowe rynkowe stopy procentowe.

5.4.2. Operacje otwartego rynku

Tabela 5.12 przedstawia najbardziej typowy – choć jednocześnie bardzo uproszczony – mechanizm regulowania podaży pieniądza przez bank centralny. Polega on na nabywaniu (lub sprzedaży) papierów wartościowych od banków komercyjnych. Gdy bank centralny kupuje papiery wartościowe od banków komercyjnych, płaci im pieniądzem rezerwowym, uznając ich rachunki prowadzone w banku centralnym. W ten sposób zasób pieniądza rezerwowego znajdującego się w posiadaniu banków komercyjnych zwiększa się, co umożliwia im zwiększanie akcji kredytowej. W rezultacie całkowita podaż pieniądza w gospodarce wzrasta. Gdy bank centralny sprzedaje papiery wartościowe bankom komercyjnym, płacą one za te papiery pieniądzem rezerwowym, wykorzystując środki zgromadzone na własnych rachunkach w banku centralnym. Wskutek tego stan tych rachunków obniża się, co zmniejsza możliwości udzielania kredytów przez banki. W rezultacie ograniczenie akcji kredytowej wywołuje spadek podaży pieniądza w gospodarce. Transakcje zakupu i sprzedaży papierów wartościowych miedzy bankiem

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

centralnym a bankami komercyjnymi nazywają się operacjami otwartego rynku (*open market operations* – OOR). Stanowią one główny instrument polityki pieniężnej, przy pomocy którego bank centralny reguluje wielkość podaży pieniądza i poziom stóp procentowych.

OOR polegają na kupnie i sprzedaży papierów wartościowych przez bank centralny. Transakcje w ramach OOR dokonywane są wyłącznie z bankami, żadne inne podmioty nie uczestniczą w tym systemie. Z punktu widzenia kierunku przepływu środków pieniężnych transakcje te dzielą się na operacje zasilające, czyli dostarczające płynność z banku centralnego do systemu bankowego, i operacje absorbujące, czyli ściągające płynność z systemu bankowego z powrotem do banku centralnego. Z perspektywy warunków, na jakich dokonywane są poszczególne transakcje, operacje otwartego rynku dzielą się na warunkowe i bezwarunkowe. Wreszcie z punktu widzenia częstotliwości i systematyczności przeprowadzania operacji otwartego rynku można umownie podzielić je na podstawowe, dostrajające i strukturalne.

Najbardziej typowe i najczęściej stosowane są operacje warunkowe. Należy do nich porozumienie w sprawie odkupu (*repurchase agreement*), czyli tak zwana operacja *repo*, oraz porozumienie w sprawie odsprzedaży (*reverse repurchase agreement*), czyli tak zwana operacja *reverse repo*. W ramach operacji *repo* bank komercyjny zawiera z bankiem centralnym umowę sprzedaży bankowi centralnemu papierów wartościowych (skarbowych papierów wartościowych, własnych papierów wartościowych, ewentualnie papierów wartościowych emitowanych przez inne podmioty) pod warunkiem ich odkupu (*repurchase*) od banku centralnego w określonym terminie (najczęściej po siedmiu dniach). W ten sposób bank centralny przejściowo zwiększa rezerwy banków, a więc zasila je w płynność (operacja zasilająca). Z kolei w ramach operacji *reverse repo* bank komercyjny zawiera z bankiem centralnym umowę kupna od banku centralnego papierów wartościowych pod warunkiem ich odsprzedaży (*reverse repurchase*) przez bank komercyjny w określonym terminie (najczęściej po siedmiu dniach). W ten sposób bank centralny przejściowo zmniejsza rezerwy banków, czyli absorbuje płynność z rynku pieniężnego (operacja absorbująca).

W obu przypadkach transakcja jest warunkowa w tym sensie, że bank komercyjny zobowiązuje się odkupić papiery wartościowe od banku centralnego w ramach operacji *repo*, ewentualnie odsprzedać papiery wartościowe bankowi centralnemu w ramach operacji *reverse repo*. Ta warunkowość sprawia, że, choć z formalnego punktu widzenia operacje otwartego rynku mają charakter transakcji kupna i sprzedaży (i ten sposób są też księgowane w bilansach banków), pod względem

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ekonomicznym mamy w istocie do czynienia z umową pożyczki. W ramach transakcji *repo* bank centralny pożycza bankowi komercyjnemu płynne środki na dany okres pod zastaw papierów wartościowych, z kolei w ramach operacji *reverse repo* to bank komercyjny „pożycza" bankowi centralnemu środki płynne na wyznaczony okres pod zastaw papierów wartościowych.

Jeśli mamy więc do czynienia z *de facto* pożyczkami, środki będące przedmiotem tych pożyczek muszą być oprocentowane. O wysokości oprocentowania decyduje zasadniczo bank centralny, ustalając krótkookresową stopę procentową, na podstawie której dokonywane są operacje *repo* i *reverse repo*. To właśnie ta stopa procentowa banku centralnego jest współcześnie najważniejszym instrumentem polityki pieniężnej, określającym „cenę" za dostęp do płynności (do pieniądza rezerwowego). Tę podstawową stopę procentową omówimy szerzej w dalszej części rozdziału.

Operacje otwartego rynku mają ograniczony zakres podmiotowy (mogą w nich uczestniczyć tylko uprawnione banki), przedmiotowy (przedmiotem zastawu mogą być tylko ściśle określone kategorie papierów wartościowych) oraz czasowy (operacje dokonywane są na określone terminy zapadalności). Ogólne zasady funkcjonowania OOR są do siebie bardzo zbliżone w poszczególnych krajach, choć szczegółowe przepisy mogą się różnić. W Polsce uczestnikami OOR mogą być jedynie banki mające status tzw. banków-dealerów rynku pieniężnego. Aby uzyskać taki status, czyli prawo uczestnictwa w OOR, banki muszą posiadać odpowiedni potencjał finansowy i spełnić cały szereg wymogów formalnych, tj. być w dobrej sytuacji finansowej (na przykład nie mogą być objęte postępowaniem naprawczym), spełniać wszystkie wymogi nadzorcze co do wysokości kapitału, płynności oraz jakości i struktury aktywów, a także wymogi sprawozdawcze, muszą uczestniczyć w mechanizmie ustalania stawek WIBOR i WIBID – czyli w tzw. fixingu – oraz zawrzeć odpowiednią umowę z Narodowym Bankiem Polskim[37]. Jeśli chodzi

[37] Narodowy Bank Polski co roku (lub częściej w marę potrzeb) ogłasza listę banków mających status Dealerów Rynku Pieniężnego i uprawnionych do uczestniczenia w operacjach rynku otwartego. W 2016 roku następujące banki w Polsce posiadały status Dealerów Rynku Pieniężnego (kolejność według numeru rozliczeniowego banku):

1) Powszechna Kasa Oszczędności Bank Polski SA,
2) Bank Handlowy w Warszawie SA,
3) ING Bank Śląski SA,
4) Bank BPH SA,
5) Bank Zachodni WBK SA,
6) Bank Gospodarstwa Krajowego,
7) mBank SA,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

o zakres aktywów będących przedmiotem OOR, to katalog ten jest w Polsce dość wąsko określony, w przeciwieństwie do wielu innych państw. Narodowy Bank Polski akceptuje w ramach OOR skarbowe papiery wartościowe (SPW), bony pieniężne emitowane przez NBP (obecnie w formie niematerialnej), obligacje Banku Gospodarstwa Krajowego, niektóre obligacje komunalne (np. emitowane przez miasto stołeczne Warszawa) oraz niektóre bankowe listy zastawne (*collateralized debt obligations* – CDO). W innych krajach katalog papierów wartościowych przyjmowanych w zastaw przez banki centralne jest ogólnie znacznie szerszy i obejmuje, poza wyżej wymienionymi aktywami, także obligacje korporacyjne, regularne obligacje bankowe oraz niektóre „syntetyczne" papiery wartościowe, czyli papiery emitowane pod zastaw aktywów niepłynnych, takich jak kredyty hipoteczne, konsumpcyjne lub kredyty udzielone w ramach leasingu[38].

Z punktu widzenia okresu objętego transakcją, OOR dzielą się generalnie na operacje krótkoterminowe i operacje długoterminowe. Te pierwsze, realizowane najczęściej i w sposób systematyczny, obejmują z reguły okres siedmiu dni. Operacje 7-dniowe są powszechnie stosowane m.in. w strefie euro i w Wielkiej Brytanii. W USA najczęściej dokonuje się operacji na jeszcze krótsze okresy – trwające dwa, trzy lub cztery dni. W Polsce NBP także przeprowadza głównie operacje 7-dniowe (w latach 2003–2005 były to operacja 14-dniowe, a przed 2003 rokiem operacje 28-dniowe). Z kolei operacje długoterminowe obejmują wszystkie operacje o dłuższych terminach zapadalności, najczęściej od trzech do sześciu miesięcy. W okresie kryzysu finansowego po 2008 roku wiele banków centralnych wydłużyło terminy OOR, dążąc do zapewnienia dostatecznej płynności w sektorze bankowym i ustabilizowania warunków funkcjonowania poszczególnych banków[39].

Najważniejsze znaczenie w regulowaniu poziomu płynności w gospodarce mają operacje krótkookresowe. Pozwalają one na bieżąco regulować poziom płynnych rezerw w systemie bankowym. Przeprowadzane są z reguły raz w tygodniu

8) Bank Millennium SA,
9) Bank Polska Kasa Opieki SA,
10) Raiffeisen Bank Polska SA,
11) Societé Generale SA Oddział w Polsce,
12) Deutsche Bank Polska SA,
13) Bank BGŻ BNP Paribas SA.

[38] Proces emisji papierów syntetycznych pod zastaw niepłynnych aktywów, takich jak kredyty, nazywa się sekurytyzacją.

[39] Na przykład, Europejski Bank Centralny przeprowadził w latach 2011–2012 dwie operacje *repo* z terminami zapadalności trzy lata (tzw. *long-term refinancing operations* – LTRO) na łączną kwotę ponad 1000 mld EUR.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(w Polsce w piątki). Krótki okres zapadalności tych operacji jest wygodny dla obu stron transakcji. Banki komercyjne decydują co tydzień o tym, czy rolować operacje zasilające lub absorbujące, przedłużając okres korzystania z dodatkowej płynności (lub inwestowania w papiery), czy raczej zrezygnować z dodatkowej płynności lub przeznaczyć ją na inne cele. Z kolei bank centralny ma do dyspozycji bardzo elastyczny instrument regulowania poziomu płynności. Reagując na dane płynące z gospodarki, bank centralny może co tydzień przedłużać na kolejne tygodnie zawarte wcześniej operacje zasilające i absorbujące, lecz może też je ograniczać lub zwiększać. Banki centralne mogą wykorzystywać również tzw. operacje dostrajające (*fine-tuning operations*) – stosowane poza normalnymi terminami OOR, w reakcji na nagłe, jednorazowe zmiany poziomu płynności w bankach. Ich celem jest ograniczenia wpływu tych zmian na poziom krótkoterminowych stawek rynkowych, w tym zwłaszcza ograniczanie ich zmienności. Z kolei operacje z dłuższymi terminami zapadalności mają charakter strukturalny, a ich celem jest uzupełnianie lub absorbowanie płynności w systemie bankowym w reakcji na zmiany sytuacji płynnościowej, dokonujące się w okresie kilku miesięcy lub nawet kilku lat, a wynikające na przykład z kryzysu finansowego lub będące skutkiem dużych przepływów kapitału.

Oprócz operacji warunkowych banki centralne stosują w ramach OOR także bezwarunkowe operacje kupna i sprzedaży papierów wartościowych – są to tzw. operacje *outright*. Ich celem jest trwałe, strukturalne dostosowanie płynności do potrzeb gospodarki. Gdy bank centralny kupuje (sprzedaje) bezwarunkowo papiery wartościowe od banków komercyjnych, zwiększa (zmniejsza) ich płynne rezerwy w sposób trwały[40] i powiększa (obniża) także sumę aktywów w swoim bilansie.

Operacje otwartego rynku są zawsze poprzedzone próbą oszacowania stanu nadpłynności lub niedoboru płynności na rynku pieniężnym – na tej podstawie bank centralny określa następnie wielkość oferty płynnych środków w ramach sukcesywnych OOR. Jest to ważny element bieżących analiz w banku centralnym, ponieważ stan płynności w sektorze bankowym ulega często dużym zmianom. Do najważniejszych czynników zewnętrznych kształtujących płynność banków należą następujące zjawiska:

a) wymiana przez rząd walut obcych na złote, np. w ramach wykorzystania środków europejskich (czasem wymiana walut przez inne podmioty, np. banki);

[40] To znaczy do momentu zapadalności papierów wartościowych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

b) zmiany poziomu wydatków z rachunków instytucji publicznych utrzymywanych w banku centralnym (wydane środki trafiają do banków i zwiększają wielkość posiadanych przez nie płynnych rezerw);
c) zmiany skali wykupu zapadających papierów wartościowych sprzedanych wcześniej bankom.

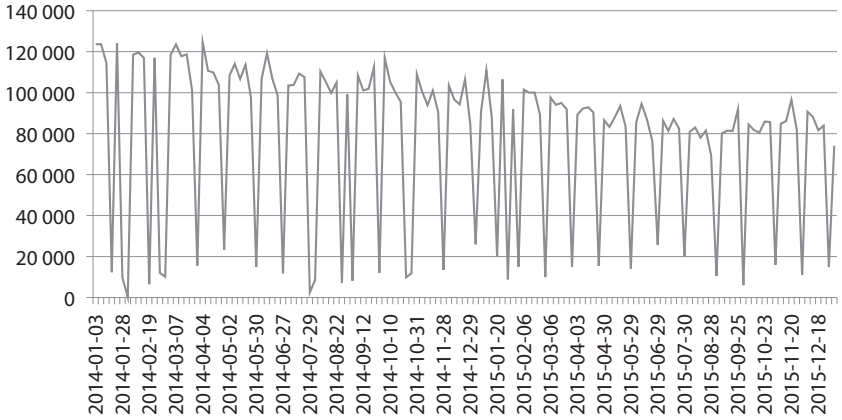
Czynniki te pozostają poza kontrolą banku centralnego, a wycena ich oddziaływania na płynność wymaga dość skomplikowanych analiz.

Przy strukturalnym nadmiarze płynności (nadpłynność) bank centralny dokonuje systematycznych operacji *reverse repo* oraz dodatkowo sprzedaje swoje bony pieniężne bankom w zamian za płynne środki, czyli pożycza od nich pieniądze. W ten sposób ogranicza skalę nadpłynności w sektorze bankowym. Przy braku nadpłynności lub przy strukturalnym niedoborze płynności bank centralny dokonuje operacji *repo*. W Polsce od wielu lat mamy do czynienia ze strukturalną nadpłynnością sektora bankowego. Wykres 5.9 przedstawia wyniki przetargów na 7-dniowe bony pieniężne NBP w ramach OOR w latach 2014–2015. Jak widać, w 2014 roku na cotygodniowych sesjach w ramach podstawowych operacji *reverse repo* banki komercyjne systematycznie kupowały od NBP bony pieniężne na kwotę 100–120 mld PLN. W następnym roku skala operacji nieco zmalała do około 80–100 mld PLN. Oprócz tego NBP sprzedawał dodatkowo bony w ramach operacji dostrajających poza terminami regularnych operacji, jednorazowo w kwotach od 5 do 20 mld PLN. Oznacza to, że, zgodnie z szacunkami NBP, skala nadpłynności w polskim systemie bankowym systematycznie przekraczała w 2014 roku 100 mld PLN, a w 2015 roku – 80 mld PLN.

Operacje otwartego rynku realizowane są w większości krajów na zasadzie przetargu. Bank centralny zbiera najpierw oferty od banków co do skali ich potrzeb w zakresie płynności (lub skali wolnych środków pozostających w ich dyspozycji) oraz co do ceny (stopy procentowej), po jakiej banki są gotowe nabywać (sprzedawać) płynność. Następnie na tej podstawie określa skalę danej operacji otwartego rynku i dokonuje transakcji z bankami oferującymi najwyższe stopy procentowe, a w każdym razie nie niższe niż oficjalna stopa procentowa banku centralnego. W celu zapewnienia równomiernego dostępu do płynności bank centralny może zdecydować się na ograniczenie skali zasilania w płynność największych banków, aby umożliwić również mniejszym bankom dostęp do płynności. Dlatego też w trakcie dokonywania operacji bank centralny często przydziela bankom środki (lub oferuje papiery) proporcjonalne do ich wielkości i zgłoszonego zapotrzebowania.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 5.9. Wyniki przetargów na 7-dniowe bony pieniężne NBP w ramach OOR w okresie 1.01.2014–31.12.2015 (dane w mln PLN)



Źródło: opracowanie własne na podstawie danych NBP.

5.4.3. Stopa procentowa: podstawowe pojęcia

Stopa procentowa jest ceną za wynajem (użyczenie) kapitału. Jeśli przedsiębiorca chce pożyczyć kapitał, idzie do banku, aby wziąć kredyt. Ceną za skorzystanie z niego jest stopa procentowa, którą przedsiębiorca musi płacić bankowi. Uiszczanie tej swoistej opłaty jest uzasadnione tym, że wypożyczony kapitał mógłby przynosić jego właścicielowi dochód gdzie indziej, na przykład jako lokata terminowa w banku lub pożyczka udzielona innemu przedsiębiorcy. Ten utracony przez właściciela dochód nazywamy kosztem alternatywnym (*opportunity cost*). Stopa procentowa powinna być więc co najmniej równa kosztowi alternatywnemu.

Realna stopa procentowa to stopa procentowa nominalna pomniejszona o stopę oczekiwanej inflacji. W praktyce, w związku z trudnościami w oszacowaniu oczekiwanej inflacji, odejmujemy od stopy nominalnej aktualną stopę inflacji. Uważa się, że są dwa ważne powody, dla których realna rynkowa stopa procentowa jest z reguły dodatnia. Po pierwsze, ludzie są z natury niecierpliwi i jeśli mają wybór, to decydują się na konsumpcję dziś, a nie na przykład za rok, a skoro już mają wstrzymać się z konsumpcją, chcą w zamian odpowiedniej rekompensaty. Oznacza to, że w subiektywnym odczuciu poszczególnych osób dana kwota pieniędzy ma większą wartość dziś niż za rok. Tę charakterystyczną cechę zachowań ludzkich nazywamy – za wybitnym ekonomistą amerykańskim Irvingiem Fisherem – czystą

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

preferencją czasową (*pure time preference* – Fisher 1930). Zatem kapitał – jako niewydana część dochodu – powinien przynosić realne wynagrodzenie jako rekompensata za odroczoną konsumpcję. Po drugie, kapitał ten może zawsze przynosić dochód, gdy zostanie zainwestowany gdzie indziej (koszt alternatywny).

W finansach rynkowa stopa procentowa jest sumą stopy wolnej od ryzyka (*risk-free*) oraz premii za ryzyko niewypłacalności dłużnika – ryzyko kredytowe. Stopa wolna od ryzyka pozostaje na ogół równa stopie czystej preferencji czasowej – jej przybliżenie w warunkach realnej gospodarki stanowi krótkookresowa stopa procentowa, ustalana przez bank centralny. Im dłuższy staje się okres kredytowania, tym wyższa będzie premia za ryzyko kredytowe. Długookresowa stopa procentowa jest natomiast równa średniej oczekiwanej krótkookresowej stopie procentowej w przyszłości (np. stopa 5-letnia jest równa przewidywanej stopie krótkiej za pięć lat).

Szwedzki ekonomista Knut Wicksell w 1898 roku wprowadził do analizy makroekonomicznej pojęcie naturalnej stopy procentowej (NSP), definiując ją jako realną stopę procentową zapewniającą równowagę między podażą a popytem na rynku produktów. Innymi słowy, NSP to taka realna stopa, która jest neutralna względem cen, a zatem nie wywołuje ani inflacji, ani deflacji. Gdy NSP jest zbyt niska, popyt wzrasta i rośnie inflacja, gdy NSP jest zbyt wysoka, popyt maleje i spada inflacja[41]. Zależność między NSP a inflacją możemy więc zapisać następująco:

$$\Delta p = \alpha(NSP - r), \quad \alpha > 0, \tag{5.17}$$

gdzie p oznacza poziom cen, a r realną rynkową stopę procentową. Formuła pokazuje, że między stopą inflacji a odchyleniem r od NSP występuje odwrotna zależność. Tak zdefiniowana NSP jest oczywiście zmienną nieobserwowalną i ma czysto empiryczny charakter. Jej szacunki mogą podlegać dużym wahaniom w zależności od zmian poziomu cen p. Odchylenie stopy rynkowej r od NSP, nazywane luką stopy procentowej, dobrze odzwierciedla stopień restrykcyjności polityki

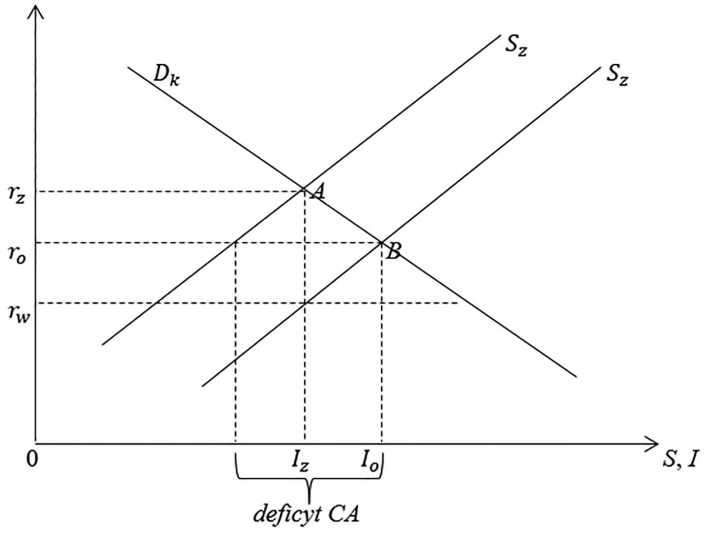
[41] Zgodnie z definicją Wicksella, „there is a certain rate of interest on loans which is neutral in respect to commodity prices, and tends neither to raise nor to lower them" (Wicksell 1936, s. 102). Współcześnie definiuje się NSP także jako realną stopę procentową zapewniającą PKB na poziomie potencjalnym („the natural rate is defined to be the real fed funds rate consistent with real GDP equaling its potential level (potential GDP) in the absence of transitory shocks to demand. Potential GDP, in turn, is defined to be the level of output consistent with stable price inflation, absent transitory shocks to supply. Thus, the natural rate of interest is the real fed funds rate consistent with stable inflation absent shocks to demand and supply" – zob. US Federal Reserve).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pieniężnej – im wyższe staje się *r* w relacji do NSP, tym bardziej restrykcyjna okazuje się polityka pieniężna.

Od czego zależy jednak poziom NSP? Odpowiedź na to pytanie można uzyskać, dokonując dekompozycji rynkowej stopy równowagi. Pierwszym składnikiem określającym wysokość stopy rynkowej jest średnia realna stopa zwrotu z kapitału w danej gospodarce. Zależy ona od czynników fundamentalnych, takich jak relacja zasobów kapitału do zasobów pracy oraz produktywność zasobów. Drugi składnik stanowi premia z tytułu ograniczenia swobody transgranicznego przepływu kapitału między daną gospodarką a zagranicą. Im silniejsze ograniczenia, tym wyższa będzie NSP w danej gospodarce. Trzecim składnikiem jest suma premii za rozmaite kategorie ryzyka – kursowe, płynności i niewypłacalności. Nominalna rynkowa stopa procentowa pozostaje natomiast równa stopie realnej, powiększonej o stopę oczekiwanej inflacji (tzw. równanie Fishera)[42].

Rysunek 5.3. Naturalna stopa procentowa w gospodarce zamkniętej i otwartej



Źródło: opracowanie własne.

W gospodarce otwartej NSP jest niższa, ponieważ napływ kapitału z zagranicy zwiększa podaż dostępnych w kraju oszczędności i obniża cenę rynkową za

[42] Więcej na temat naturalnej stopy procentowej zob. Brzoza-Brzezina (2011).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ich wykorzystanie. Zarazem NSP będzie z reguły wyższa niż stopa procentowa na rynku międzynarodowym, ponieważ: 1) napływ kapitału idący w parze ze wzrostem zadłużenia powoduje wzrost ryzyka niewypłacalności, 2) istnieje ryzyko kursowe (w przypadku odrębnych walut), oraz 3) wśród inwestorów zagranicznych występuje zjawisko preferencji wobec inwestycji w krajowe papiery wartościowe (tzw. *home bias*).

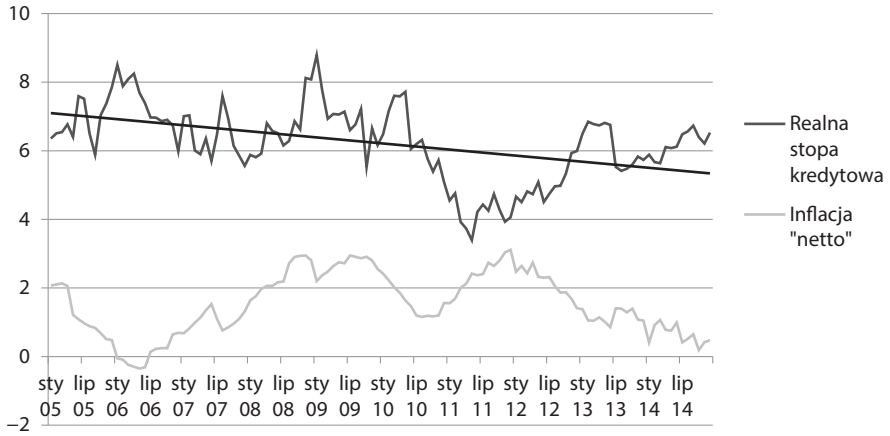
Rysunek 5.3 przedstawia naturalną stopę procentową w gospodarce zamkniętej i otwartej. W gospodarce zamkniętej naturalna stopa procentowa, czyli realna stopa procentowa równowagi r, ustali się na poziomie r_z, równoważącym podaż krajowych oszczędności S_z z krajowym popytem inwestycyjnym D_I. W ten sposób inwestycje zrównają się z oszczędnościami (S = I), czyli rynek produktów znajdzie się w równowadze. Otwarcie gospodarki na przepływy kapitału, gdy realna stopa na rynku światowym wynosi $r_w < r_z$, spowoduje napływ oszczędności zagranicznych. Krzywa podaży dostępnych oszczędności przesunie się do położenia S_o, a punkt równowagi z A do B. Realna stopa procentowa równowagi – czyli NSP – spadnie w tej sytuacji do r_o, powodując wzrost inwestycji z I_z do I_o. Jednocześnie pojawi się deficyt na rachunku bieżącym (CA), finansowany napływem kapitału zagranicznego. Nowa naturalna stopa procentowa r_o będzie niższa niż r_z, ale z reguły wyższa niż realna stopa rynku światowego r_w ze względu na krajowe preferencje inwestorów zagranicznych (*home bias*).

Czy opisana na rysunku 5.3 hipoteza o obniżaniu się krajowej NSP na skutek otwarcia gospodarki na napływ kapitału jest znajduje potwierdzenie w rzeczywistości? Próba empirycznej weryfikacji tej hipotezy w przypadku Polski została przedstawiona na wykresie 5.10. Ponieważ NSP jest bezpośrednio nieobserwowalna, za jej przybliżenie przyjęto realną rynkową stopę procentową obliczoną jako średnia nominalna stopa oprocentowania złotowych kredytów bankowych pomniejszona o wskaźnik CPI.

Jeśli wartość realnej rynkowej stopy procentowej może być uznana za przybliżenie NSP, to przebieg zmian linii trendu sugeruje, że w Polsce występuje wyraźna spadkowa tendencja NSP, zgodna z logiką postępującej integracji finansowej (rosnąca skala napływu kapitału). Z równania trendu wynika, że w latach 2005–2014 NSP obniżała się średnio w tempie około 18 punktów bazowych na rok, przy względnie stabilnej inflacji. Występuje również wyraźna ujemna korelacja pomiędzy NSP a inflacją „netto" (współczynnik korelacji wynosi –0,38), co potwierdza formułę Wicksella opisaną równaniem (5.17): odchylenia r od trendu są ujemnie skorelowane ze zmianami inflacji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 5.10. Realna rynkowa stopa procentowa oraz inflacja po wyłączeniu cen żywności i energii (tzw. inflacja „netto") w Polsce w latach 2005–2014 (dane w%)[a]



a) Realna stopa rynkowa obliczona jako średnia nominalna stopa oprocentowania stanu kredytów złotowych pomniejszona o wskaźnik.

Źródło: opracowanie własne na podstawie danych NBP.

5.4.4. Stopa procentowa jako instrument polityki pieniężnej

Jak wspomniano wcześniej, bank centralny reguluje wielkość obiegu pieniężnego w gospodarce przy pomocy różnych instrumentów, oddziałujących na cenę oraz ilość pieniądza rezerwowego. Do narzędzi tych należą:

a) krótkookresowa stopa procentowa: jest to cena, po jakiej bank centralny zasila w płynność (pieniądz rezerwowy) banki komercyjne, lub po jakiej absorbuje nadmierną płynność banków komercyjnych (operacje *repo, reverse repo i outright*), i w ten sposób zwiększa lub zmniejsza zdolność banków do udzielania kredytów;
b) stopy procentowe dla pozostałych operacji kredytowo-depozytowych;
c) stopa rezerwy obowiązkowej: określa relację rezerwy obowiązkowej do depozytów i stanowi instrument regulowania wielkość emisji kredytu przy danym poziomie rezerw;
d) nastawienie w polityce pieniężnej: neutralne, restrykcyjne (większe prawdopodobieństwo zaostrzenia), ekspansywne (większe prawdopodobieństwo poluzowania);

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

e) limity kredytowe nałożone na poszczególne banki (instrument administracyjny, stosowany wyjątkowo).

Największe znaczenie spośród wymienionych instrumentów polityki pieniężnej mają krótkookresowe stopy procentowe. Przypomnijmy, że stopa procentowa jest ceną za wynajem (użyczenie) kapitału – jeśli ktoś chce wziąć kredyt, musi płacić stopę procentową, ponieważ wypożyczony kapitał mógłby przynosić jego właścicielowi dochód gdzie indziej (*opportunity cost*). Podobnie dzieje się w przypadku pożyczek udzielanych bankom przez bank centralny (lub odwrotnie). Gdy bank komercyjny chce pożyczyć płynne środki z banku centralnego, uiszcza opłatę w postaci krótkookresowej stopy procentowej, którą ustala bank centralny. Na przykład, jeśli bank chce pożyczyć w ramach operacji *repo* 1 milion PLN na 7 dni, a stopa procentowa w przypadku tych operacji wynosi 3% w skali roku, musi uiścić kwotę $1\,000\,000 \times (7 \times 0{,}03/360) = 583{,}33$ PLN. Im wyższa jest stopa procentowa, tym wyższe stają się koszty pożyczek i tym niższy będzie popyt na płynność ze strony banków. Dlatego bank centralny, ustalając stopę procentową w przypadku operacji *repo* na określonym poziomie, reguluje też popyt banków komercyjnych na płynność. Stopa procentowa stosowana w operacjach *repo* i *reverse repo* jest najważniejszą stopą procentową polityki pieniężnej (tzw. *policy rate*). Stopa ta jest różnie nazywana w różnych krajach – na przykład w Polsce jest to tzw. stopa referencyjna, w strefie euro – główna stopa refinansowa (*main refinancing rate – refi*), a w Wielkiej Brytanii – tzw. stopa bankowa (*Bank Rate*). Nazywa się ją także oficjalną stopą procentową, ponieważ jest wyznaczana przez instytucję państwową.

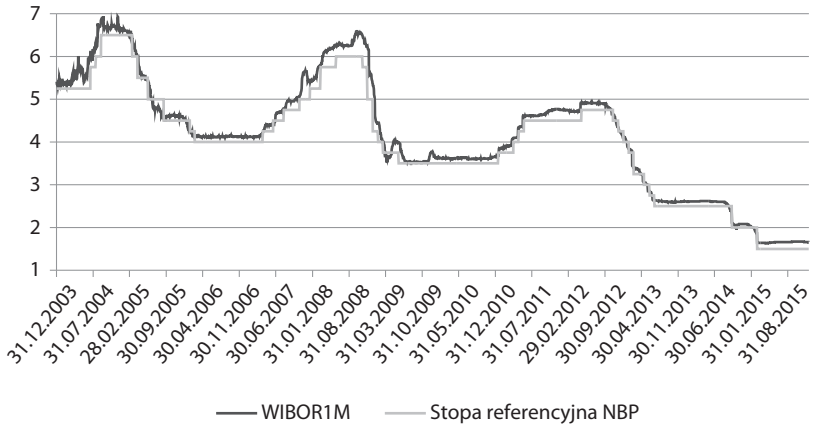
Główna stopa procentowa banku centralnego jest nie tylko ceną, jaką banki komercyjne płacą bankowi centralnemu za płynność, lecz pozostaje także ważnym punktem odniesienia dla kształtowania się krótkookresowych stóp rynku pieniężnego. Rynek pieniężny, zwany też rynkiem międzybankowym, jest miejscem, w którym banki komercyjne udzielają sobie wzajemnie pożyczek. Poszczególne banki znajdują się w różnej sytuacji płynnościowej – w danym momencie niektóre z nich mają nadwyżkę, a inne niedobór płynności. Wynika to z niedopasowania terminów zapadalności aktywów i pasywów bankowych, pojawiających się możliwości inwestowania, zmian w rynkowej cenie aktywów, określonej strategii biznesowej banku oraz innych czynników. Krótkookresowe rynkowe stopy procentowe są z reguły zbliżone do głównej stopy procentowej banku centralnego ze względu na mechanizm arbitrażu cenowego. Gdyby stopy rynkowe były wyraźnie wyższe od stopy banku centralnego, banki zaopatrywałyby się w płynność w banku centralnym w ramach operacji *repo*, gdyby zaś stopy rynkowe okazały się wyraźnie niższe,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

banki pożyczałyby wolne środki bankowi centralnemu w ramach operacji *reverse repo*. W obu przypadkach rynkowa cena płynności będzie kształtować się w pewnym przedziale wartości w pobliżu głównej stopy procentowej banku centralnego

Wykres 5.11 pokazuje, że jednomiesięczna stopa rynku pieniężnego w Polsce (WIBOR1) w okresie ostatnich dwunastu lat kształtowała się mniej więcej na poziomie stopy referencyjnej NBP, przy czym średnia różnica między tymi stopami wynosiła 16 punktów bazowych. Wykres potwierdza, że polski bank centralny, zgodnie z przyjętym celem operacyjnym, z powodzeniem utrzymywał stopy rynku pieniężnego na poziomie lub nieco powyżej stopy referencyjnej, co było istotnym warunkiem utrzymania kontroli nad podażą pieniądza w warunkach silnej strukturalnej nadpłynności w polskim sektorze bankowym. Jedynie w kilku przypadkach stopa WIBOR1M spadła poniżej stopy referencyjnej, i to tylko wówczas, gdy rynek oczekiwał na obniżenie stóp oficjalnych.

Wykres 5.11. WIBOR1M i stopa referencyjna NBP w okresie 31.12.2003–31.12.2015 (dane w %)



Źródło: opracowanie własne na podstawie danych NBP.

Aby uniknąć nadmiernych wahań rynkowych stóp procentowych wokół poziomu, na którym znajduje się główna (oficjalna) stopa procentowa, bank centralny ustala jeszcze dwie ważne stopy procentowe, wyznaczające limit górny i limit dolny dla wahań stóp rynkowych. Pierwsza z nich dotyczy oprocentowania krótkookresowych pożyczek zaciąganych przez banki komercyjne w banku centralnym w celu doraźnego uzupełnienia płynności. Pożyczki takie udzielane są z reguły na 1 dzień (z możliwością rolowania pożyczki na kolejne dni) bankom, które z jakiegoś powodu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

znalazły się przejściowo w sytuacji niedoboru płynności i nie są w stanie uzupełnić tych braków z innych źródeł. Oprocentowanie takich pożyczek jest odpowiednio wyższe od oprocentowania regularnych operacji *repo*. W Polsce tego rodzaju jednodniowy kredyt nazywa się kredytem lombardowym, ponieważ jest on udzielany pod zastaw papierów wartościowych, a stopa procentowa od tych kredytów nosi miano stopy kredytu lombardowego lub – krócej – stopy lombardowej. Druga ze stóp limitujących wahania stóp rynkowych dotyczy oprocentowania krótkookresowych depozytów, które banki komercyjne mogą złożyć w banku centralnym, gdy mają nadwyżkę płynności i nie są w stanie ulokować tej nadwyżki na rynku pieniężnym. Lokata taka składana jest na jeden dzień (*overnight*), a jej oprocentowanie pozostaje odpowiednio niższe niż oprocentowanie regularnych operacji *reverse repo*. W Polsce stopę oprocentowania tych jednodniowych depozytów nazywa się stopą depozytową. Stopa kredytu lombardowego i stopa depozytowa wyznaczają więc górny i dolny limit w przypadku krótkookresowych wahań rynkowych stóp procentowych na rynku międzybankowym.

Podobne rozwiązania stabilizujące rynek pieniężny istnieją również w innych krajach. Na przykład w strefie euro Europejski Bank Centralny także wyznacza, oprócz głównej stopy refinansowej, także dwie krańcowe stopy procentowe – krańcową stopę kredytową (*marginal lending rate*) i krańcową stopę depozytową (*marginal deposit rate*).

Czasami banki centralne stosują dodatkowe stopy procentowe w przypadku pewnych określonych operacji depozytowo-kredytowych, dotyczących niektórych operacji finansowych czy wybranych sektorów gospodarki, takich jak rolnictwo czy eksport. W Polsce Narodowy Bank Polski wyznacza jeszcze tzw. stopę redyskonta weksli, która jest wykorzystywana w rozliczaniu kredytów udzielanych rolnikom na sfinansowanie produkcji rolnej. Oprócz tego NBP określa także tzw. stopę kredytu refinansowego, która odnosi się do kredytów udzielonych w przeszłości na finansowanie tzw. inwestycji centralnych, realizowanych przez rząd[43]. W odróżnieniu od rynkowych stóp procentowych wszystkie wymienione stopy procentowe są stopami oficjalnymi, ponieważ ustala je instytucja publiczna, jaką jest bank centralny.

[43] Jest to instrument pochodzący jeszcze z czasów poprzedniego systemu. NBP zaprzestał finansowania nowych inwestycji centralnych w połowie lat 90., ale w jego bilansie figurują nadal niektóre inwestycje rozpoczęte jeszcze w latach 80. i dotychczas niezakończone.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wśród stóp procentowych ustalanych przez bank centralny znajduje się również stopą oprocentowania rezerw utrzymywanych przez banki komercyjne na rachunkach w banku centralnym. Rezerwy te są oprocentowane po to, aby zmniejszyć koszt alternatywny ich utrzymywania. Uważa się bowiem, że banki mogłyby inwestować te środki w bardziej dochodowe przedsięwzięcia. Oprocentowanie rezerw, zwłaszcza obowiązkowych, ma na celu częściowe zrekompensowanie bankom tych utraconych dochodów. W Polsce rachunki rezerwowe banków są oprocentowane od 2004 roku, a stopa oprocentowania wynosi obecnie 0,9 stopy referencyjnej.

Tabela 5.13 zawiera syntetyczne zestawienie instrumentów stosowanych w celu dostarczania i absorbowania płynności w gospodarce na przykładzie strefy euro. Jak widać, Eurosystem (system banków centralnych państw należących do strefy euro wraz z Europejskim Bankiem Centralnym) korzysta z szerokiej gamy narzędzi, które mogą być stosowane na okres od jednego dnia do kilku lat, w sposób ciągły lub okresowo, a także wykorzystujących różne rodzaje zabezpieczeń. Zapewnia to dużą elastyczność w polityce pieniężnej.

Tabela 5.13. Instrumenty dostarczania i absorbowania płynności w Eurosystemie

Operacje polityki pieniężnej	Typ transakcji		Okres zapadalności (*maturity*)	Częstotliwość
	dostarczanie płynności	absorbowanie płynności		
Operacje otwartego rynku				
Główne operacje refinansowe (MRO)	transakcje *repo*		tydzień	tygodniowo
Długoterminowe transakcje finansowe (LTRO)	transakcje *repo*		3 miesiące – 3 lata	miesięcznie, nieregularnie
Operacje dostrajające	transakcje *repo*, swapy walutowe	transakcje *reverse repo*, depozyty terminowe, swapy walutowe	zróżnicowany	nieregularnie
Operacje strukturalne	transakcje *repo*, zakupy *outright*	sprzedaż certyfikatów EBC, sprzedaż *outright*	zróżnicowane	zróżnicowane
Instrumenty stałe (*standing facilities*)				
Marginal lending facility (MLF)	transakcje *repo*		1 dzień (*overnight*)	na życzenie
Marginal deposit facility (MDF)		depozyty	1 dzień (*overnight*)	na życzenie

Źródło: Europejski Bank Centralny.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.4.5. Pozostałe instrumenty

Ważne miejsce wśród instrumentów stosowanych przez bank centralny zajmuje stopa rezerwy obowiązkowej. Jak już wiadomo, określa ona relację rezerwy obowiązkowej do depozytów i stanowi instrument regulowania wielkości emisji kredytu przy danym poziomie rezerw. Utrzymywanie rezerw na określonym poziomie umożliwia bankom niezakłócone zaspokajanie potrzeb płynnościowych, co zapobiega nadmiernym wahaniom krótkoterminowych rynkowych stóp procentowych. W przeciwieństwie do oficjalnych stóp procentowych, które wpływają na poziom stóp rynkowych i w ten sposób kształtują popyt na kredyty, stopa rezerwy obowiązkowej oddziałuje na podaż kredytów oferowanych przez banki. Podniesienie stopy rezerwy zmusza banki do ograniczenia tempa wzrostu akcji kredytowej, ponieważ większa część depozytów musi być odprowadzona na rachunek rezerwowy w banku centralnym – a zatem zostaje wycofana z obiegu pieniężnego. I odwrotnie, obniżenie stopy rezerwy pozwala zwiększyć akcję kredytową, ponieważ uwalnia część rezerw. Przypomnijmy, że zależność pomiędzy stopą rezerwy obowiązkowej i podażą pieniądza opisuje wzór na mnożnik pieniężny, którego wielkość zależy odwrotnie proporcjonalnie od wielkości stopy rezerwy (zob. punkt 5.2).

Ogólnie rzecz biorąc, uważa się, że stopa procentowa regulująca popyt na kredyty jest skuteczniejszym instrumentem kształtowania wielkości emisji kredytu niż stopa rezerwy obowiązkowej. Zmiany oficjalnych stóp procentowych przenoszą się stosunkowo szybko na zmiany rynkowych stóp, co bezpośrednio wpływa na koszt kredytu i na popyt. Z kolei wpływ zmian stopy rezerw obowiązkowych na wielkość podaży kredytu jest słabszy i przebiega z dużym opóźnieniem. Dzieje się tak dlatego, że kredyt stanowi tylko część sumy aktywów w bilansach banków komercyjnych, co umożliwia bankom utrzymanie akcji kredytowej kosztem obniżenia udziału pozostałych aktywów nawet w obliczu wzrostu stopy rezerwy obowiązkowej. W związku z tym głównym instrumentem zmian podaży pieniądza pozostają stopy procentowe, a stopy rezerwy obowiązkowej zmieniane są rzadko i z reguły w odpowiedzi na strukturalne zmiany w sektorze bankowym.

Rezerwy obowiązkowe stanowią z reguły niewielki ułamek depozytów. Stopa rezerwy obowiązkowej wynosi w Polsce 3,5% stanu depozytów, w strefie euro 2%, a w Stanach Zjednoczonych waha się od 0% w przypadku depozytów o wartości poniżej 15,5 mld USD aż do 10% odnośnie depozytów przekraczających 115,1 mln USD (stan na koniec 2013 roku). W krajach o mniej stabilnych systemach finansowych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i wysokiej inflacji stopy rezerw obowiązkowych są wyższe – na przykład w Polsce w pierwszych latach po transformacji systemowej stopa rezerwy obowiązkowej przekraczała 20%.

Niektóre banki centralne wykorzystują w polityce pieniężnej komunikaty określające prawdopodobny kierunek polityki pieniężnej w najbliższej przyszłości. Ogłaszając tzw. nastawienie polityki (*policy stance*), informują uczestników rynku o swoich intencjach co do przyszłych decyzji z zakresu polityki pieniężnej. Ogłaszając zmianę nastawienia,, bank centralny wpływa na oczekiwania uczestników rynku, kieruje się potrzebą prowadzenia polityki pieniężnej zorientowanej na przyszłość, uwzględniającej opóźnienia, jakie występują między decyzją dotyczącą stóp procentowych i jej najsilniejszym wpływem na inflację. Gdy bank centralny ogłasza nastawienie restrykcyjne (*restrictive policy stance, restrictive bias*), sygnalizuje tym samym, że w krótkim okresie istnieje większe prawdopodobieństwo zaostrzenia polityki pieniężnej niż jej poluzowania. Z kolei odwrotnie, gdy bank ogłasza nastawienie ekspansywne (*expansionary policy stance, expansionary bias*), istnieje większe prawdopodobieństwo poluzowania polityki finansowej niż jej zaostrzenia. Przy nastawieniu neutralnym bank centralny daje do zrozumienia, że zarówno zaostrzenie, jak i luzowanie polityki pieniężnej jest w krótkim okresie jednakowo (mało) prawdopodobne[44].

Obecnie banki centralne rzadziej stosują instrument nastawienia, przede wszystkim dlatego, że ma on charakter silnie krótkookresowy. Ogłoszone nastawienie obowiązuje w zasadzie tylko do następnego posiedzenia organu decyzyjnego banku centralnego – z reguły jest to horyzont od dwóch tygodni do jednego miesiąca, a to oznacza, że ogłaszanie nastawienia jedynie w niewielkim stopniu zmniejsza niepewność na rynkach finansowych[45]. Dlatego też w ostatnich latach, w obliczu potrzeby stabilizowania oczekiwań dłuższym horyzoncie, banki centralne największych krajów zaczęły wykorzystywać inny sposób komunikowania rynkom swoich zamiarów – tzw. *forward guidance*. Omówimy go w dalszej części rozdziału.

Banki centralne większości krajów mogą również stosować pewne instrumenty regulacji administracyjnej. Należą do nich m.in. limity kredytowe nakładane na poszczególne banki, mające na celu hamowanie akcji kredytowej w warunkach

[44] Instrument nastawienia może też być zdefiniowany w sposób bardziej precyzyjny. Na przykład nastawienie restrykcyjne może oznaczać, że jakiekolwiek poluzowanie polityki pieniężnej jest wykluczone w krótkim okresie, a nastawienie ekspansywne z kolei oznacza, że jakiekolwiek zaostrzenie jest wykluczone.

[45] Na przykład Narodowy Bank Polski stosował instrument nastawienia tylko w latach 1999–2005.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

głębokiego naruszenia równowagi makroekonomicznej. Stanowią one narzędzie zakazu administracyjnego i co do zasady powinny być stosowane tylko w wyjątkowych okolicznościach, gdy rynkowe (parametryczne) instrumenty polityki pieniężnej są z jakiegoś powodu nieskuteczne. Ma to miejsce, na przykład, w warunkach bardzo wysokiej inflacji, kiedy zachodzi potrzeba radykalnego ograniczenia tempa wzrostu podaży pieniądza. W Polsce stosowano limity kredytowe w ramach systemu gospodarki planowej oraz w pierwszych latach transformacji, kiedy instytucje rynkowe i mechanizm polityki pieniężnej nie były dostatecznie rozwinięte[46]. Również niektóre kraje Zachodniej Europy stosowały limity kredytowe w okresie bezpośrednio następującym po II wojnie światowej.

Ramka 5.4. Naliczanie i rozliczanie rezerwy obowiązkowej w Polsce

Rezerwą obowiązkową jest wyrażona w złotych część środków pieniężnych zgromadzonych na rachunkach bankowych i przyjętych przez banki oraz podlegających zwrotowi, z wyjątkiem środków przyjętych od innego banku krajowego, a także pozyskanych z zagranicy na co najmniej 2 lata[47]. Rezerwa obowiązkowa utrzymywana jest przez banki na rachunkach w Narodowym Banku Polskim. Wysokość stopy rezerwy obowiązkowej ustala Rada Polityki Pieniężnej. Od 31 października 2003 roku stopa rezerwy obowiązkowej wynosi 3,5% w przypadku wszystkich rodzajów depozytów (w okresie od połowy 2009 do końca 2010 stopa rezerwy była przejściowo obniżona do 3,0% w związku z kryzysem finansowym). Od 30 września 2003 roku wszystkie banki pomniejszają naliczoną rezerwę obowiązkową o równowartość 500 tysięcy EUR. Środki rezerwy obowiązkowej od 1 maja 2004 roku są oprocentowane.

Rezerwa obowiązkowa jest utrzymywana od ostatniego dnia miesiąca następującego po miesiącu, za który nalicza się rezerwę, do dnia poprzedzającego kolejny okres utrzymywania rezerwy. Jeżeli pierwszy dzień okresu utrzymywania rezerwy, zwanego dalej „okresem rezerwowym", jest dniem wolnym od pracy, to utrzymywanie rezerwy rozpoczyna się pierwszego dnia roboczego po tym dniu. Banki składają w Narodowym Banku Polskim – nie później niż do 25. dnia miesiąca

[46] Ustawa o NBP z 1997 roku przewiduje możliwość zastosowania ograniczeń wielkości środków pieniężnych przekazywanych przez banki do dyspozycji kredytobiorców i pożyczkobiorców – czyli limitów kredytowych (art. 46, ust. 1). NBP nie korzystał dotychczas z tej możliwości.

[47] Uchwała nr 15/2004 Zarządu NBP, DzU NBP 2004, nr 3.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

następującego po miesiącu, za który nalicza się rezerwę obowiązkową – deklarację w sprawie wysokości rezerwy obowiązkowej banku.

Banki komercyjne (z wyjątkiem banków spółdzielczych) utrzymują rezerwę obowiązkową na swoich rachunkach bieżących prowadzonych w NBP. Nie muszą one jednak utrzymywać rezerwy o wymaganej wysokości każdego dnia – bank centralny oczekuje jedynie, aby średnia dzienna stanu rezerwy w miesięcznym okresie jej utrzymywania była równa rezerwie naliczonej (czyli wymaganej). Daje to bankom możliwość elastycznego wykorzystywania płynnych środków w trakcie okresu rozliczeniowego na inne cele (na przykład na krótkookresowe inwestycje w dochodowe aktywa). Aby zapewnić średni dzienny poziom utrzymywanej rezerwy zgodny z wyliczonym limitem, banki często pod koniec okresu rozliczeniowego dokonują transakcji uzupełniających stan płynności (lub odwrotnie), co tłumaczy duże fluktuacje krótkookresowych stóp procentowych pod koniec każdego miesiąca (wykres 5.10).

Jeżeli rezerwa obowiązkowa utrzymywana przez bank w danym okresie rezerwowym jest rzeczywiście niższa od rezerwy wymaganej, to od kwoty stanowiącej różnicę pomiędzy kwotą wymaganej rezerwy obowiązkowej a kwotą rezerwy obowiązkowej faktycznie utrzymaną na rachunku bank uiszcza na rzecz NBP odsetki karne. Jako faktycznie utrzymaną kwotę rezerwy w danym okresie rezerwowym rozumie się średnią arytmetyczną stanu środków na rachunku bieżącym lub rachunku rezerwy obowiązkowej ze wszystkich dni okresu rezerwowego.

Odsetki karne naliczane są według następującego wzoru:

$$O_k = \frac{K * n * p}{100 * 365}, \tag{5.18}$$

gdzie:
O_k – kwota odsetek z tytułu nieutrzymania kwoty wymaganej rezerwy obowiązkowej w okresie rezerwowym (odsetki „karne”),
K – różnica między kwotą wymaganej rezerwy obowiązkowej w okresie rezerwowym a kwotą rezerwy faktycznie utrzymaną przez bank w tym okresie,
n – liczba dni kalendarzowych w okresie rezerwowym,
p – wysokość oprocentowania, które stosuje się przy naliczaniu odsetek z tytułu nieutrzymania kwoty wymaganej rezerwy obowiązkowej w okresie rezerwowym.

Odsetki karne są naliczane według stopy procentowej, stanowiącej dwukrotność stopy oprocentowania kredytu lombardowego, obowiązującej w okresie rezerwowym, za który należą się odsetki.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Środki rezerwy obowiązkowej, utrzymywane przez banki na rachunkach bieżących lub rachunkach rezerwy obowiązkowej są oprocentowane w okresie rezerwowym w wysokości ustalonej przez Radę Polityki Pieniężnej. Oprocentowaniu podlegają środki tylko do wysokości wymaganej rezerwy obowiązkowej, nie wyższej jednakże niż rezerwa obowiązkowa faktycznie utrzymana przez bank w okresie rezerwowym.

Odsetki należne bankom z tytułu oprocentowania środków rezerwy obowiązkowej w danym okresie rezerwowym oblicza się według następującego wzoru:

$$O_r = \frac{R * n * i}{100 * 365}, \tag{5.19}$$

gdzie:
O_r – należne odsetki od kwoty rezerwy obowiązkowej w okresie rezerwowym,
R – średnia dzienna kwota rezerwy obowiązkowej w okresie rezerwowym,
n – liczba dni kalendarzowych w okresie rezerwowym,
i – wysokość oprocentowania stosowana przy naliczaniu odsetek od środków rezerwy obowiązkowej w okresie rezerwowym.

Ramka 5.5. Operacje kredytowo-depozytowe NBP

W sytuacji, gdy NBP prowadzi podstawowe operacje otwartego rynku z 7-dniowym terminem zapadalności, może dochodzić do znacznych wahań najkrótszych, zwłaszcza jednodniowych, stóp rynku międzybankowego. Łagodzeniu tych wahań służą operacje kredytowo-depozytowe, prowadzone z bankami komercyjnymi z ich inicjatywy. Operacje te obejmują udzielanie kredytu lombardowego oraz przyjmowanie lokaty terminowej od banków (w formie depozytu na koniec dnia). Operacje kredytowo-depozytowe NBP wpływają na wysokość stóp procentowych na rynku pieniężnym i mają na celu utrzymanie tych stóp w przedziale, którego górną granicę stanowi oprocentowanie kredytu lombardowego, a dolną – oprocentowanie depozytu w NBP.

NBP udziela bankom kredytu lombardowego pod zastaw skarbowych papierów wartościowych. Umożliwia on im pokrywanie krótkookresowych niedoborów płynności i udzielany jest na następujących zasadach:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

- zastawem są skarbowe papiery wartościowe, a wysokość kredytu nie może przekroczyć 80% ich wartości nominalnej,
- termin spłaty kredytu przypada w następnym dniu operacyjnym po dniu jego udzielenia,
- warunkiem przyznania kredytu jest uprzednia spłata wcześniej zaciągniętego zadłużenia.

NBP oferuje też bankom możliwość składania krótkookresowego (jednodniowego) depozytu w banku centralnym. Lokaty przyjmowane są do końca dnia operacyjnego, a zwrot kwoty depozytu wraz z należnymi odsetkami następuje w kolejnym dniu operacyjnym. Lokaty są oprocentowane według stopy depozytowej ustalanej przez Radę Polityki Pieniężnej. Pozwalają one bankom komercyjnym na zagospodarowanie nadwyżek płynnych środków, co przeciwdziała spadkowi krótkookresowych procentowych stóp na rynku międzybankowym poniżej stopy depozytowej.

NBP oferuje też tzw. kredyt techniczny, umożliwiający bankom komercyjnym zaciągnięcie nieprocentowanego kredytu w ciągu dnia operacyjnego. Ułatwia on dokonywanie rozliczeń bieżących, przy ograniczeniu możliwości powstawania napięć o charakterze płynnościowym w sektorze bankowym. Podobnie jak miało to miejsce w przypadku pozostałych operacji zasilających, także kredyt techniczny udzielany jest pod zabezpieczenie w postaci skarbowych papierów wartościowych.

5.4.6. Wskaźnik restrykcyjności polityki pieniężnej

Często zadajemy sobie pytanie o to, czy w danym momencie polityka pieniężna jest restrykcyjna czy ekspansywna? Najprościej można odpowiedzieć, że polityka pieniężna pozostaje restrykcyjna, gdy dostęp do pieniądza okazuje się utrudniony, czyli gdy realna „cena" pieniądza krajowego jest wysoka. I odwrotnie, polityka pieniężna jest ekspansywna wówczas, gdy realna „cena" pieniądza krajowego staje się niska. Lecz co stanowi cenę krajowego pieniądza? Większość z nas zapewne odpowie, że jest nią krajowa stopa procentowa. Ale taka odpowiedź jest niepełna.

Gdy zwróci się uwagę na skonsolidowany bilans krajowego systemu bankowego (zob. punkt 5.1), to natychmiast daje się zauważyć, że podaż krajowego pieniądza (widoczna po stronie pasywów) może pochodzić z dwóch źródeł (widocznych po stronie aktywów) – jednym z nich jest kredyt krajowy, a drugim aktywa

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zagraniczne. Gdy bank udziela kredytu, tworzy depozyt, a kiedy skupuje waluty obce wypłaca gotówkę lub uznaje rachunek sprzedawcy. Stopa procentowa jest więc rzeczywiście „ceną" pieniądza krajowego, lecz tylko takiego, który został pozyskany z udzielonego kredytu. Z kolei „cenę" uzyskania pieniądza krajowego z wymiany walut stanowi kurs walutowy. Dlatego, jeśli chce się określić stopień restrykcyjności polityki pieniężnej, należy wziąć pod uwagę oba wskaźniki – krajową stopę procentową i kurs walutowy.

Na tej podstawie możemy wyznaczyć syntetyczny wskaźnik restrykcyjności polityki pieniężnej (*Monetary Condition Index* – MCI) jako średnią ważoną dwóch wskaźników – wskaźnika realnej stopy procentowej (r) i wskaźnika realnego kursu walutowego (e). Otrzymujemy więc (zob. np. Eriksson i in. 1998, Batini i Turnbull 2000, Kot 2003):

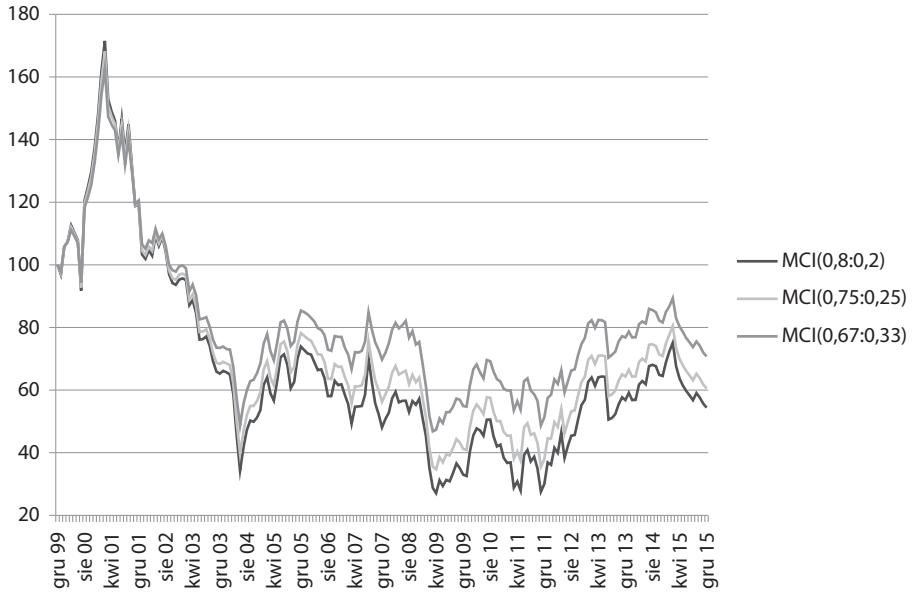
$$\mathrm{MCI} = \alpha(\textit{indeks } r) + \beta(\textit{indeks } e), \text{ przy czym } \alpha + \beta = 1. \tag{5.20}$$

Wskaźnik realnej stopy procentowej to wskaźnik stopy nominalnej podzielony przez wskaźnik cen (na przykład CPI), z kolei wskaźnik realnego kursu walutowego to wskaźnik zmian nominalnego kursu podzielony przez wskaźnik cen. Waga przy wskaźniku realnej stopy procentowej będzie z reguły kilkakrotnie większa niż waga określająca wskaźnik realnego kursu waluty ze względu na większy udział kredytu w tworzeniu pieniądza krajowego[48]. Na przykład z tabeli 5.5 wynika, że proporcje, w jakich po stronie aktywów skonsolidowanego bilansu krajowego systemu bankowego występują aktywa krajowe i aktywa zagraniczne, wynoszą odpowiednio około 70% i 30% całkowitej sumy aktywów. Interpretując wartości wskaźnika MCI, trzeba pamiętać, że tak obliczony wskaźnik jest miarą relatywną – ilustruje zmiany restrykcyjności polityki pieniężnej w porównaniu z okresem bazowym, a nie w odniesieniu do jakiegoś poziomu „optymalnego". Należy też pamiętać, że wskaźnik MCI tylko częściowo odzwierciedla *zamierzoną* restrykcyjność polityki pieniężnej, ponieważ o ile krótkookresowa stopa procentowa jest z reguły kontrolowana przez bank centralny, o tyle kurs walutowy jest najczęściej zmienną w całości zależną od rynku.

[48] W Polsce na koniec 2013 roku suma kredytów udzielonych przez sektor bankowy jest niemal dwa i pół razy większa niż suma aktywów zagranicznych znajdujących się w posiadaniu sektora bankowego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 5.12. Wskaźnik MCI dla Polski w latach 2000–2015 (grudzień 1999 roku = 100)[a]



a) Wskaźnik realnej stopy procentowej wyliczony jako wskaźnik stopy referencyjnej NBP deflowany wskaźnikiem CPI; wskaźnik kursu realnego wyliczony jako wskaźnik zmian kursu „koszyka" (kurs euro i kurs dolara amerykańskiego w proporcji 50:50), deflowany wskaźnikiem CPI.

Źródło: opracowanie własne na podstawie danych NBP.

Na wykresie 5.12 przedstawiono zmiany wskaźnika MCI w przypadku Polski w latach 2000–2015, wyliczonego w trzech wariantach: dla $\alpha = 0{,}8$ i $\beta = 0{,}2$ dla $\alpha = 0{,}75$ i $\beta = 0{,}25$ i dla $\alpha = 0{,}67$ i $\beta = 0{,}33$[49]. W latach 2000–2004 występowały duże wahania MCI spowodowane zmianami stóp procentowych, wahaniami kursu i zmianami stopy inflacji. Gwałtowny wzrost restrykcyjności polityki pieniężnej w okresie od lipca 2000 roku do lutego 2001 roku był spowodowany głównie silnym podniesieniem stóp procentowych w odpowiedzi na wzrost inflacji. Z kolei systematyczny spadek restrykcyjności w okresie od marca 2001 roku do kwietnia 2004 roku był głównie wynikiem obniżania stóp procentowych w odpowiedzi na spadek inflacji. W kolejnych latach wskaźnik MCI był bardziej stabilny, przy czym coraz większą rolę zaczęła odgrywać aprecjacja kursu walutowego. Jedynie w drugiej połowie 2008 roku gwałtowna deprecjacja kursu złotego spowodowała silny

[49] W strefie euro wskaźnik MCI jest liczony przy relacji wag 6:1 na korzyść wskaźnika stóp procentowych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

spadek restrykcyjności. Można też zauważyć, że od końca 2011 roku do początku 2015 roku następowało systematyczne zaostrzenie polityki pieniężnej mimo bardzo słabej koniunktury i stosunkowo niskich stóp procentowych. Wyjątkiem była pierwsza połowa 2013 roku, kiedy dokonano serii obniżek oficjalnej stopy procentowej (z 4,25% w grudniu 2012 roku do 2,50 w lipcu 2013 roku). Od połowy 2013 roku do początku 2015 roku wskaźnik MCI ponownie zaczął rosnąć, co było spowodowane malejącą inflacją przy niezmienionych stopach procentowych. W rezultacie restrykcyjność polityki pieniężnej na przełomie 2014 i 2015 roku powróciła do poziomu, jaki odnotowano w okresie silnego wzrostu gospodarczego w latach 2005–2008. Od marca 2015 roku wartość wskaźnika MCI spadała, co miało związek z obniżeniem stopy referencyjnej NBP w tym miesiącu i ze stopniowym osłabieniem kursu złotego.

Wykres 5.12 pozwala także zauważyć, że praktycznie w całym okresie po 2002 roku kurs walutowy zwiększał restrykcyjność polityki pieniężnej – w przypadku wyższych wag wskaźnika realnego kursu wartości wskaźnika MCI są odpowiednio wyższe. Może to częściowo wynikać z zaniżonego poziomu kursu „koszyka" w okresie bazowym (grudzień 1999 roku).

5.5. Transmisja impulsów polityki pieniężnej do gospodarki

5.5.1. Mechanizm transmisji monetarnej

Jak już wiadomo, bank centralny nie decyduje wprawdzie bezpośrednio o wysokości i zmianach cen w gospodarce, lecz wpływa na inne zmienne ekonomiczne, które z kolei oddziałują na poziom i zmiany cen. Impulsy polityki pieniężnej – takie jak zmiany oficjalnych stóp procentowych, wielkość i częstotliwość operacji *repo* i *reverse repo*, zmiany stopy rezerwy obowiązkowej – wpływają na wielkość podaży pieniądza w gospodarce, a w konsekwencji także na efektywny popyt, wydatki i – ostatecznie – na ceny (czyli na inflację). Jak pokazano na rysunku 5.2, w procesie tym występuje wiele ogniw, połączonych złożonymi zależnościami przyczynowo-skutkowymi, których rzeczywista postać często nie jest do końca rozpoznana. Całość tego procesu nazywamy mechanizmem transmisji monetarnej. Jego głównymi elementami składowymi są rynek pieniężny (międzybankowy), gdzie impulsy polityki pieniężnej oddziałują na zachowania banków komercyjnych, oraz poszczególne kanały transmisji, za pośrednictwem których zachowania banków

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przenoszą się i oddziałują na zachowania uczestników rynku w pozostałych częściach gospodarki. Ich sposób postępowania wpływa zaś ostatecznie na inflację[50].

5.5.2. Rynek pieniężny

Rynek pieniężny, inaczej mówiąc rynek międzybankowy, to rynek na którym banki udzielają sobie wzajemnie pożyczek krótkoterminowych w celu uzupełnienia niedoborów płynności lub zagospodarowania nadwyżek płynności. Niedobór płynności może pojawić się w związku z koniecznością budowania rezerwy obowiązkowej pod koniec okresu rozliczeniowego (zob. ramka 5.6), w związku z wzrostem akcji kredytowej, lub w związku z koniecznością uzupełnienia płynności na bieżące wypłaty dokonywane przez klientów. Nadwyżka płynności może pojawić się na skutek wcześniejszej spłaty kredytu lub uwolnienia części rezerw. Zarówno niedobór, jak i nadwyżka wynikają z niedopasowania terminów zapadalności aktywów i pasywów oraz z tego, że trudno jest przewidzieć wszystkie wydarzenia mające wpływ na stan płynności w banku. Wobec tego potrzebny jest mechanizm pozwalający bankom zarządzać płynnością w taki sposób, aby uniknąć znaczących strat. Są nim właśnie krótkookresowe pożyczki międzybankowe na rynku pieniężnym.

Transakcje pożyczkowe zawierane są między bankami po określonych rynkowych stopach procentowych, bilansujących podaż płynności z popytem na płynność (Sławiński 2011). Stawki referencyjne w przypadku tych stóp w żargonie finansowym nazywane są stopami IBOR (*interbank offered rate* – międzybankowa stopa ofertowa), a ustala się je (czyli kwotuje) na podstawie obserwowanych faktycznych stóp rynkowych. Najbardziej typowa procedura wyznaczania stawek IBOR polega na tym, że wyodrębniona grupa banków kwotuje stawki referencyjne na rynku pieniężnym w jednakowym czasie na okres pożyczki od jednego dnia to jednego roku. Ustalenie stóp procentowych na dany dzień – tzw. *fixing* – odbywa się z reguły o określonej godzinie (na rynkach europejskich o 11.00 czasu GMT) i wtedy wybrana grupa banków podaje stawki, po jakich jest skłonna zaoferować środki na poszczególne okresy. Z punktu widzenia funkcjonowania rynku kredytowego największą rolę odgrywają stawki *overnight* (ON), stawki jednomiesięczne (1M) i trzymiesięczne (3M), gdyż na te okresy najczęściej banki udzielają sobie

[50] Więcej na temat mechanizmu transmisji pieniężnej w Polsce – zob. Demchuk i in. (2012) oraz Kapuściński i in. (2014).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pożyczek. Stawki w przypadku kredytów denominowanych w walutach obcych kwotowane są na rynku londyńskim (LIBOR) i europejskim (EURIBOR).

W normalnych warunkach rynkowych stawki IBOR podlegają dwóm głównym tendencjom. Po pierwsze, im dłuższy okres, tym stawka z reguły jest wyższa, odzwierciedlając tym samym rosnące ryzyko związane z udzielaniem pożyczki na dłuższy okres. Po drugie, stawki IBOR powinny, chociażby w części, uwzględniać oczekiwania co do zmian stóp na rynku międzybankowym. Przy silnym oczekiwaniu na obniżki stóp procentowych krzywa złożona z kwotowań na rosnące okresy powinna być opadająca.

Międzybankowa stawka referencyjna kwotowana na polskim rynku pieniężnym nazywa się WIBOR (*Warsaw Interbank Offered Rate*). Wybrane banki komercyjne (mające status dealerów rynku pieniężnego) kwotują każdego dnia o godzinie 11.00 dwie stawki: stawkę, po której przyjmą depozyty (*Warsaw Interbank Interest on Deposits* – WIBID), i stawkę, po której są skłonne udzielać pożyczek (WIBOR). Z kwotowań tych banków eliminuje się kwotowania najwyższe i najniższe, a z pozostałych wyciągana jest średnia i w ten sposób ustalane są stawki WIBOR i WIBID, które następnie są ogłaszane oficjalnie przez Narodowy Bank Polski. Różnica między tymi stawkami – tzw. *spread* – nie może być zbyt duża i wynosi w przypadku transakcji powyżej 1M maksimum 20 punktów bazowych (0,2 punktu procentowego). Ponadto stawka WIBOR w Polsce jest stawką transakcyjną, co oznacza, że po zakwotowaniu przez przynajmniej jakiś czas banki powinny zawierać transakcje po zaproponowanych przez siebie stawkach. Stopy WIBOR i WIBID kwotuje się dla operacji na jeden dzień (*overnight* – O/N), trzy dni (*tomorrow-next* – T/N), siedem dni, czternaście dni, 1 miesiąc (1M), 3 miesiące (3M), 6 miesięcy (6M), 9 miesięcy (9M) i rok (360 dni).

Stopy WIBOR pełnią istotną rolę w mechanizmie transmisji monetarnej, ponieważ są punktem odniesienia w przypadku ustalania stóp oprocentowania kredytów udzielanych przez banki podmiotom gospodarczym. Ogromna większość kredytów w Polsce pozostaje oprocentowana zmienną stopą procentową, która jest kalkulowana przez banki na poziomie wybranej bieżącej stopy WIBOR (głównie WIBOR 1M lub WIBOR 3M), powiększonej o marżę banku (*spread*). Wysokość marży jest ustalona z klientem w umowie kredytowej na stałym poziomie, z kolei stawka WIBOR zależy od bieżących notowań. Jeśli umowa kredytowa przewiduje spłatę kredytu w ratach miesięcznych, na przykład 5. dnia każdego miesiąca, bank stosuje do wyliczenia wysokości raty kredytu stawkę WIBOR z tego dnia. W ten sposób zmiany stóp procentowych rynku pieniężnego (czyli WIBOR-y) przenoszą

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

się na zmiany kosztu obsługi kredytu, wywierając tym samym wpływ na popyt na kredyty ze strony przedsiębiorstw i gospodarstw domowych. Zarazem stopy WIBOR zależą – ze względu na oddziaływanie mechanizmu arbitrażu cenowego – od stóp oficjalnych NBP. W ten sposób bank centralny, zmieniając poziom oficjalnych stóp, wpływa na poziom rynkowych stóp procentowych i koszt kredytu w całej gospodarce.

Stawki WIBOR nie nadają się jednak do rozliczania transakcji pożyczkowych o bardzo krótkim okresie zapadalności, na przykład jednodniowych. Początkowo stosowano w tym celu stawkę WIBOR O/N, lecz nie odzwierciedlała ona prawidłowo warunków rynku transakcji O/N, ponieważ wszystkie stawki referencyjne WIBOR prezentują stan rynku o godzinie 12.00. O ile w przypadku dłuższych terminów jest to wystarczające, o tyle odnośnie do transakcji O/N taka stawka nie oddaje w pełni tego, co dzieje się na rynku pieniężnym w ciągu całego dnia. Stawki O/N z reguły cechują się dużą zmiennością w ciągu dnia i jedno kwotowanie z godziny 12.00 nie wystarcza, aby prawidłowo ocenić zachowanie rynku w innych porach dnia. Była to szczególnie kłopotliwa niedogodność w przypadku transakcji o najkrótszych terminach, rozliczanych pod koniec dnia. Dlatego Narodowy Bank Polski, a także wiele innych banków centralnych, kwotuje także stawki rynku pieniężnego na koniec dnia operacyjnego (notowania są dokonywane między godziną 16.00 a 17.00). W Polsce ta krótkookresowa stawka *overnight* nazywa się POLONIA (*Polish Overnight Interest Average*), ze względu na analogię z podobną stawką notowaną w strefie euro (EONIA – *Euro Overnight Interest Average*). Sposób jej wyznaczania i notowania omówiony został w ramce 5.6.

Ramka 5.6. Stawka POLONIA

Stawka procentowa POLONIA stosowana jest na polskim rynku międzybankowym w transakcjach pieniężnych w odniesieniu do depozytów O/N (w strefie euro stosuje się stopę EONIA). Stawka ta jest więc obliczana jako średnia ważona ze stawek oraz kwot transakcji przekazanych przez uczestników rynku do godziny 16.30 każdego dnia roboczego. Każdy uczestnik rynku zobowiązany jest do przekazania NBP wykazu transakcji zawartych i przesłanych do rozliczenia w danym dniu do godziny 16.00 na termin O/N. Skalkulowana w ten sposób stawka zostaje opublikowana w serwisie informacyjnym NBP o godzinie 17.00. Obecnie stawka POLONIA opisuje średnio około 60% zawartych transakcji międzybankowych. Do

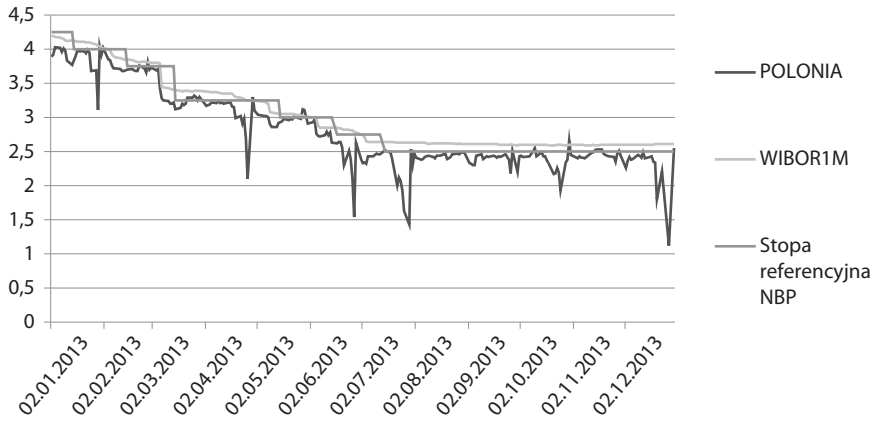
Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

indeksu liczą się codzienne transakcje O/N (ich dzienny poziom w 2013 roku wynosił średnio 3,5 mld PLN). Stawki POLONIA znajdują główne zastosowanie przy rozliczaniu kontraktów OIS, lecz jednak coraz częściej stawka ta jest wykorzystywana w bankach także jako *benchmark* służący do wyznaczenia kosztu finansowania.

Od 2008 roku celem operacyjnym NBP jest takie kształtowanie stawki rynkowej POLONIA, aby była ona jak najbardziej zbliżona poziomem do stopy referencyjnej. W praktyce polega to na wykorzystaniu operacji otwartego rynku do oddziaływania na skalę płynności w sektorze bankowym w sposób umożliwiający utrzymywanie krótkoterminowych rynkowych stóp procentowych na poziomie zbliżonym do stopy referencyjnej, co ma sprzyjać realizacji założonego celu inflacyjnego. W szczególności bank centralny reguluje rozmiary operacji otwartego rynku w celu odpowiedniego kształtowania stawki POLONIA. Wcześniej NBP opierał cele operacyjne na stawkach dłuższych. Do końca 2002 roku celem operacyjnym była stawka WIBOR 1M, w latach 2003–2004 – stawka WIBOR 2W, a w latach 2005–2007 – stawka WIBOR 1W.

Wykres 5.13. pokazuje, że choć stawka POLONIA w 2013 roku generalnie kształtowała się w okolicach poziomu stopy referencyjnej, to okresowo występowały wyraźne spadki poniżej stopy referencyjnej. Odchylenia te miały z reguły miejsce w końcowych dniach każdego miesiąca i wynikały z rozwiązywania kontraktów OIS w związku z miesięcznym cyklem rozliczania rezerwy obowiązkowej.

Wykres 5.13. Stawka POLONIA, stawka WIBOR1M i stopa referencyjna NBP w okresie 01.01.2013–31.12.2013 (dane w %)



Źródło: dane NBP.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.5.3. Instrumenty dłużne

Instrumenty dłużne to papiery wartościowe stwierdzające istnienie zobowiązania finansowego dłużnika wobec wierzyciela, a więc zarazem istnienie wierzytelności wierzyciela wobec dłużnika. Emisja i sprzedaż instrumentów dłużnych jest formą pozyskiwania kapitału pożyczkowego przez uczestników rynku, którzy potrzebują środków finansowych na określone cele. Zaciągnięte w ten sposób pożyczki są oprocentowane i muszą być spłacane w określonych terminach. Różnią się zatem zasadniczo od instrumentów udziałowych, które stanowią tytuł własności i przynoszą dochód w postaci udziału w zysku (dywidenda), nie podlegają natomiast zwrotowi i nie są oprocentowane. Instrumenty dłużne mają charakter płynny, czyli mogą być przedmiotem obrotu na rynkach finansowych, w przeciwieństwie do pożyczek i kredytów, które takiego charakteru nie mają.

Instrumenty dłużne różnią się głównie poziomem oprocentowania i terminami zapadalności. Najczęściej przyjmują postać obligacji (*bonds*) i bonów oraz weksli (*bills*). Bon jest krótkoterminowym papierem wartościowym na okaziciela o terminie zapadalności (*maturity*) z reguły od jednego tygodnia do 12 miesięcy. Sprzedawany jest z dyskontem w stosunku do ceny nominalnej, czyli poniżej jej poziomu. Wykup bonu w zapadającym terminie następuje po cenie nominalnej, a dochód inwestora równy jest różnicy pomiędzy ceną nominalną a ceną zakupu, czyli pozostaje równy kwocie dyskonta. Z kolei obligacja jest średnio- lub długoterminowym papierem wartościowym na okaziciela o terminie zapadalności od roku do 20–30 lat (a niekiedy nawet do 50 lat). Oprocentowanie obligacji w większości przypadków jest stałe, określone w postaci tzw. kuponu (*coupon*), i wypłacane na tej podstawie inwestorowi raz lub dwa razy w roku. Stosunek kuponu do wartości nominalnej stanowi nominalne oprocentowanie obligacji[51]. Lecz emitowane są też obligacje o zmiennym oprocentowaniu, a także tzw. obligacje zerokuponowe, czyli obligacje sprzedawane z dyskontem i zapewniające wypłatę skumulowanych odsetek w momencie wykupu.

Gdy obligacje o stałym oprocentowaniu – czyli o stałym kuponie – są przedmiotem obrotu na rynku wtórnym, ich cena rynkowa zmienia się w zależności od zmian podaży i popytu. Ponieważ wielkość kuponu, czyli kwota odsetek, nie ulega jednak zmianie, stosunek wartości kuponu do bieżącej ceny rynkowej obligacji waha

[51] Są to tzw. instrumenty stałokuponowe (*fixed income instruments*).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

się w kierunku przeciwnym do ceny rynkowej. Relacja ta nazywa się rentownością – lub dochodowością – obligacji (*bond yield*). Spadek ceny rynkowej obligacji wywołuje zatem wzrost rentowności, a wzrost ceny obligacji powoduje spadek rentowności. Warto tę zależność dobrze zapamiętać, ponieważ wbrew intuicji wzrost rentowności jest zjawiskiem niekorzystnym dla inwestora, bo oznacza spadek cen posiadanych przez niego papierów wartościowych. Wyjaśnienie relacji zachodzących między ceną a rentownością obligacji przedstawione zostało w ramce 5.7, a relacji między dyskontem a stopą rentowności w ramce 5.8.

Ramka 5.7. Ceny i rentowności obligacji

Załóżmy, że rząd emituje długoterminowe obligacje o wartości nominalnej (emisyjnej) 100 PLN za sztukę, terminie zapadalności 10 lat i kuponie 5 PLN, co oznacza, że nominalne oprocentowanie obligacji w momencie emisji – czyli rentowność – wynosi 5%. Obligacje te są następnie przedmiotem kupna i sprzedaży na rynku wtórnym. Załóżmy, że ich cena po pewnym czasie spada do 80 PLN. Ponieważ kupon wynosi nadal 5 PLN rocznie, oznacza to, że rentowność wzrasta z 5/100 = 5% do 5/80 = 6,25%. Wzrost rentowności jest spowodowany spadkiem ceny. Z kolei odwrotnie, gdy cena obligacji wzrośnie, na przykład, do 120 PLN, rentowność spadnie do 5/120 = 4,17%.

Na tej podstawie, znając nominalne oprocentowanie w momencie emisji i rentowności bieżące notowane na rynku, możemy obliczyć aktualną cenę rynkową danej obligacji. Na przykład, rentowności obligacji dziesięcioletnich rządu Grecji, które w latach 2006–2007 były emitowane z oprocentowaniem nominalnym na poziomie średnio około 4,5% rocznie, w szczytowym momencie kryzysu zadłużenia w lutym 2012 roku wzrosły do 29%. Oznacza to, że ich cena rynkowa spadła w czasie kryzysu do zaledwie 15,5% wartości nominalnej (4,5/29 = 0,155). Świadczyło to tym, że rynki finansowe były przekonane, że rząd grecki stał się praktycznie niewypłacalny

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 5.8. Dyskonto i stopa rentowności bonów skarbowych

W większości krajów krótkoterminowe bony skarbowe (*Treasury Bills*) sprzedawane są z dyskontem, czyli poniżej ceny nominalnej, a wykupywane po cenie nominalnej. Wynagrodzeniem inwestora z tytułu udzielenia pożyczki emitentowi bonu jest więc kwota dyskonta, stanowiąca różnicę między ceną wykupu bonu (wartością nominalną) a ceną jego nabycia. Dyskonto jest zatem równe różnicy między wartością nominalną i ceną nabycia. Aby ocenić opłacalność takiej transakcji z punktu widzenia inwestora, trzeba obliczyć stopę dyskonta i porównać ja z aktualną rynkową stopą procentową. Jeśli *kwota* dyskonta D z definicji wynosi:

$$D = P_E - P_z, \tag{5.21}$$

to *stopa* dyskonta d jest równa (w skali rocznej):

$$d = \frac{P_E - P_Z}{P_E} \cdot \frac{360}{n}, \tag{5.22}$$

gdzie: P_E – nominalna wartość bonu równa cenie wykupu, P_Z – cena nabycia bonu (z dyskontem), n – liczba dni do wykupu. Z kolei stopa rentowności r wynosi:

$$r = \frac{P_E - P_Z}{P_Z} \cdot \frac{360}{n}. \tag{5.23}$$

Różnica między stopą dyskontową a stopą rentowności polega na tym, że stopa dyskontowa jest stopą aktualizacji przyszłego dochodu w czasie n na dzień dzisiejszy, podczas gdy stopa rentowności jest stopą zwrotu z dziś dokonanej inwestycji w określonym czasie n. Załóżmy, że rząd sprzedaje 13-tygodniowy bon skarbowy o wartości nominalnej 1000 PLN za 985 PLN, czyli z dyskontem $D = 15$ PLN. Stopa dyskonta będzie równa zatem:

$$d = \frac{15}{1000} \cdot \frac{360}{91} = 0{,}0593,$$

czyli wyniesie 5,93% w skali rocznej. Z kolei stopa rentowności stanie się równa:

$$r = \frac{15}{985} \cdot \frac{360}{91} = 0{,}0602,$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

czyli wyniesie 6,02% w skali rocznej. Oczywiście, między stopą dyskonta a stopą rentowności występuje ściśle określona zależność:

$$d = \frac{r}{1+r}, \tag{5.24}$$

oraz

$$r = \frac{d}{1-d}. \tag{5.25}$$

Ogólnie rzecz biorąc, stopa rentowności jest zawsze nieco wyższa niż stopa dyskonta, a obie stopy zmieniają się w tym samym kierunku.

5.5.4. Krzywa rentowności

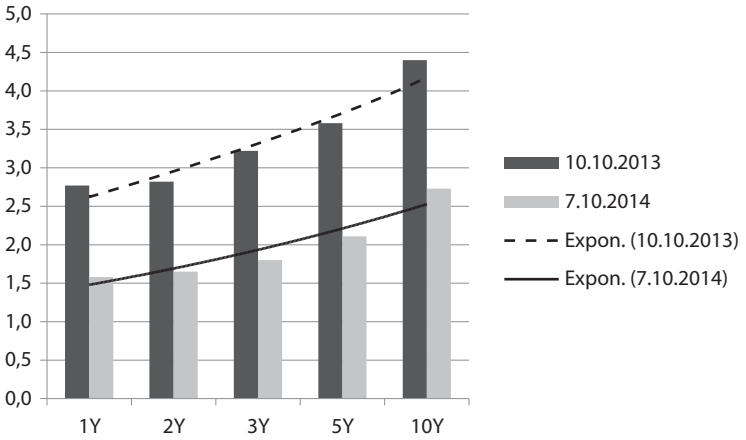
Krzywa rentowności – zwana też krzywą dochodowości (*yield curve*) – przedstawia rynkowe stopy rentowności rynkowych instrumentów dłużnych (papierów wartościowych, lokat) od najkrótszych do najdłuższych okresów zapadalności. W normalnych warunkach (przy niskiej i stabilnej inflacji) krzywa rentowności położona jest nisko (stopy rentowności są umiarkowanie niskie w przypadku wszystkich okresów zapadalności), a jej naturalnym kształtem pozostaje krzywa lekko wznosząca się. Z kolei w okresach zwalczania wysokiej inflacji krzywa rentowności jest położona wysoko (stopy rentowności są wysokie, zwłaszcza w przypadku krótszych terminów) i ma kształt wyraźnie opadający, ponieważ uczestnicy rynku spodziewają się obniżenia inflacji i w ślad za tym spadku stóp procentowych. Krzywa rentowności może mieć kształt opadający także wówczas, gdy rynki oczekują pogorszenia sytuacji gospodarczej – dzięki tej właściwości krzywa rentowności może być dobrym wskaźnikiem wyprzedzającym koniunktury (*lead indicator*).

W Polsce w okresie ostatnich dwudziestu lat krzywa rentowności systematycznie obniżała się na całej długości w miarę postępów w obniżaniu inflacji – do 2003 roku była wyraźnie opadająca, odzwierciedlając rynkowe oczekiwania spadającej inflacji. Wykres 5.14 prezentuje przykładowe krzywe rentowności z dwóch przykładowych okresów. Jak widać, krzywa rentowności była w latach 2013–2014 lekko rosnąca i zbliżona kształtem oraz poziomem do krzywych w innych rozwiniętych gospodarkach rynkowych. Zarazem widać, że w latach 2013–2014 miał miejsce w Polsce silny spadek średnich rentowności na całej długości krzywej, czyli w przypadku wszystkich okresów zapadalności. Było to spowodowane przede

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wszystkim bardzo niskimi oficjalnymi stopami procentowymi w większości krajów rozwiniętych, w tym w Polsce.

Wykres 5.14. Krzywa rentowności w Polsce: rentowności obligacji skarbowych stałokuponowych o okresach zapadalności od 1 roku do 10 lat (dane w %, wartości rzeczywiste i trend wykładniczy)



Źródło: opracowanie własne na podstawie danych portalu bossa.pl

Bank centralny na ogół stara się, aby krzywa rentowności miała kształt naturalny, tj. lekko rosnący, i aby była położona względnie nisko. Ponieważ bank centralny ustala jednak tylko krótkookresowe stopy procentowe, ma bezpośredni wpływ tylko na „krótki koniec" krzywej, z kolei „długi koniec" zależy od oczekiwań rynku co do przyszłego poziomu stóp procentowych. Oczekiwania te z kolei zależą w dużym stopniu od prognoz przyszłego stanu koniunktury, w tym od oczekiwań inflacyjnych. Ponieważ bankowi centralnemu zależy, aby długi koniec krzywej rentowności nie był położony ani zbyt wysoko (bo sygnalizuje to wysokie oczekiwania inflacyjne), ani zbyt nisko (bo świadczy to o pesymistycznych nastrojach uczestników rynku), będzie się starał utrzymywać oczekiwania co do stóp długookresowych na umiarkowanym poziomie. Związek między stopami krótkimi i stopami długimi pozwala bankom centralnym odpowiednio kształtować oczekiwania uczestników rynku.

Co sprawia jednak, że długoterminowe stopy rynkowe kształtują się na wyższym lub niższym poziomie w stosunku do wartości bieżących? Istnieje kilka koncepcji wyjaśniających zachowanie długoterminowych stóp procentowych. Zgodnie z *teorią oczekiwań*, stopa długookresowa podlega bieżącej krótkookresowej stopie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

procentowej i jej oczekiwanemu poziomowi w przyszłości. Na przykład wysokość oprocentowania dwunastomiesięcznych (12M) dłużnych papierów wartościowych $i_{0,12}$ zależy od aktualnej wysokości oprocentowania sześciomiesięcznych (6M) papierów $i_{0,6}$ i oczekiwanej wysokości stopy sześciomiesięcznej za 6 miesięcy od dziś $i_{6,12}$ (Sałata 2011, s. 240–241):

$$1+i_{0,12}=\left(1+0,5i_{0,6}\right)\left(1+0,5i_{6,12}\right). \tag{5.26}$$

Gdy lewa strona równania (5.25) jest mniejsza niż prawa strona, rentowności papierów 12M są relatywnie niskie, a więc ich ceny pozostają relatywnie wysokie względem oczekiwanych stóp za 6 miesięcy. Inwestorzy będą się zatem spodziewać, że rentowności 12M wzrosną, dostosowując się do oczekiwanych stóp 6M za 6 miesięcy (ponieważ stopy $i_{6,12}$ są wyższe), i będą sprzedawać papiery 12M. Spowoduje to spadek ich cen, wzrost rentowności i wyrównanie się lewej strony oraz prawej strony w równaniu (5.26). Gdy z kolei lewa strona równania (5.26) jest większa niż prawa, mechanizm dostosowawczy działa w przeciwną stronę.

Zgodnie z *teorią preferencji płynności* banki komercyjne preferują papiery krótkoterminowe ze względów płynnościowych. Ponieważ ceny papierów krótkich są mniej wrażliwe na zmiany stóp procentowych, ich ceny podlegają mniejszym wahaniom niż ceny papierów długich, i dzięki temu pozostają one wygodniejszym i mniej kosztownym instrumentem bieżącego uzupełniania płynności niż papiery długie. Dlatego jeśli banki mają inwestować w papiery długie (obligacje), to domagają się stosownej premii za ryzyko zmiany cen, co powoduje wzrost rentowności papierów długich – stąd klasyczny, lekko wznoszący się kształt krzywej rentowności.

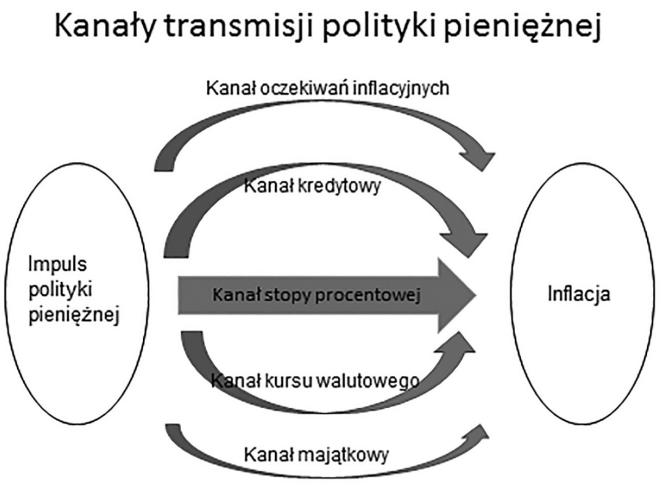
Ponieważ wysokość długoterminowych stóp procentowych zależy od oczekiwanego w przyszłości poziomu krótkoterminowych stóp procentowych, dobrą wskazówkę co do przewidywanego w przyszłości poziomu stóp procentowych stanowią terminowe stopy procentowe notowane na rynku i tzw. kontrakty FRA (*forward rate agreement*). Rynek kontraktów FRA oferuje zabezpieczenie przed odchylaniem się przyszłej wysokości stóp krótkich od ich oczekiwanego poziomu w dniu zawarcia kontraktu. Na przykład terminowa stopa procentowa w kontrakcie FRA 3x6, odzwierciedlająca oczekiwany poziom stopy 3M za 3 miesiące, jest określona przez różnicę pomiędzy aktualną (bieżącą) stopą 3M i aktualną stopą 6M. Oprocentowanie kontraktów FRA pozostaje dobrym wskaźnikiem oczekiwań rynków finansowych co do decyzji banków centralnych dotyczących stóp procentowych: jeśli terminowe stopy rosną, rynki oczekują podwyżek stóp procentowych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.5.5. Kanały transmisji

Oddziaływanie polityki pieniężnej na gospodarkę i ceny przebiega za pośrednictwem kilku kanałów. Różnią się one między sobą kierunkiem działania, szybkością transmisji impulsów pieniężnych, siłą ich wpływu i skutecznością oddziaływania. Najczęściej wymienia się kanał stopy procentowej, kanał kredytowy, kanał kursu walutowego, kanał oczekiwań inflacyjnych i kanał majątkowy (zob. rysunek 5.4).

Rysunek 5.4. Kanały transmisji impulsów polityki pieniężnej



Źródło: opracowanie własne.

Kanał stopy procentowej łączy zmiany oficjalnej stopy procentowej ze zmianami oprocentowania kredytu bankowego i ze zmianami oprocentowania depozytów. Zmiany oficjalnych stóp procentowych dokonywane przez bank centralny oddziałują nie tylko na nominalne rynkowe stopy procentowe, lecz także – przy dość dużej sztywności cen i płac nominalnych w krótkim okresie – na realne stopy procentowe. Realne stopy (zarówno bieżące, jak i oczekiwane w przyszłości) oddziałują z kolei na decyzje uczestników rynku o inwestowaniu oraz konsumowaniu lub oszczędzaniu.

Załóżmy, że bank centralny podnosi stopy procentowe, obawiając się wzrostu inflacji. Ten impuls ze strony polityki pieniężnej wywołuje wzrost rynkowych stóp procentowych na rynku pieniężnym (dzięki mechanizmowi arbitrażu), a wzrost

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

stawek rynkowych (WIBOR) przekłada się na wzrost oprocentowania kredytów i depozytów w bankach, ponieważ banki starają się utrzymać swoje marże (*spready*) na niezmienionym poziomie. Wzrost ceny kredytu skłania uczestników rynku do ograniczenia popytu na kredyty, co wpływa na zmniejszenie wydatków (lub powoduje spadek tempa ich wzrostu), a wzrost oprocentowania depozytów zwiększa skłonność do oszczędzania. W rezultacie wydatki maleją i słabnie presja inflacyjna (przy danej podaży towarów). Odwrotny skutek wywiera obniżenie stóp procentowych – kredyt tanieje, co sprzyja wzrostowi wydatków, a spadek oprocentowania depozytów obniża skłonność do oszczędzania. Skuteczność funkcjonowania kanału stopy procentowej staje się tym większa, im większy jest udział kredytów bankowych w finansowaniu gospodarki, oraz tym mniejsza, im większy jest udział kredytów walutowych w całości akcji kredytowej.

Skuteczność kanału stopy procentowej zależy także od tego, które stopy rynkowe są najważniejsze dla uczestników rynku i jak zmiana stopy procentowej banku centralnego wpływa na wahania tych stóp. Na przykład stopy oprocentowania kredytów hipotecznych mogą być zmienne i indeksowane do wybranej stopy rynku pieniężnego (na przykład do WIBOR 3M). Wówczas impuls w postaci podwyżki stopy banku centralnego wpływa zarówno na wzrost bieżących kosztów obsługi kredytu, jak i na spadek zainteresowania zaciąganiem nowych kredytów. Oba efekty działają na rzecz spadku inflacji. Tak funkcjonuje rynek kredytów hipotecznych między innymi w Stanach Zjednoczonych i w Polsce. Alternatywnie stopy oprocentowania kredytów mogą być stałe, tak jak ma to miejsce na przykład w Niemczech. Wówczas impuls w postaci podwyżki stopy banku centralnego wpływa tylko na spadek zainteresowania zaciąganiem nowych kredytów, z kolei bieżące koszty obsługi kredytu pozostają bez zmiany. Wpływ na inflację jest więc słabszy niż w pierwszym przypadku.

Kanał kredytowy oddziałuje na wielkość akcji kredytowej banków poprzez wpływ impulsu polityki pieniężnej na koszt pozyskania pieniądza rezerwowego przez sektor bankowy. Załóżmy, że bank centralny podnosi stopy procentowe w celu obniżenia inflacji. Powoduje to wzrost kosztu pozyskiwania płynności w ramach operacji otwartego rynku, co zmniejsza skalę popytu banków na pożyczki z banku centralnego. Mniejsze zasoby pieniądza rezerwowego w bankach powodują ograniczenie skali kredytów (Bernanke i Gertler 1995). Bank centralny może też bezpośrednio ograniczyć skalę operacji *repo*, co wywoła spadek płynnych rezerw w posiadaniu banków. W takiej sytuacji, jeśli banki nie są w stanie skompensować spadku rezerw (na przykład poprzez sprzedaż innych aktywów), muszą ograniczyć skalę akcji kredytowej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Zmniejszenie podaży kredytów wpływa na zmniejszenie popytu i wydatków w gospodarce, co zmniejsza presję inflacyjną. Efekt ten będzie dodatkowo wzmocniony wzrostem rynkowych stóp procentowych wywołanym ograniczeniem płynności w sektorze bankowym. W pewnych warunkach kanał kredytowy może jednak nie funkcjonować prawidłowo. Dzieje się tak w warunkach istnienia dużej nadpłynności w sektorze bankowym. Wówczas ograniczenie skali operacji *repo* może nie wywołać pożądanego spadku podaży kredytów, ponieważ banki posiadają obfite zasoby płynności. Podobnie może być w sytuacji, gdy banki zdecydują się na sprzedaż innych aktywów (na przykład obligacji rządowych lub korporacyjnych). Skuteczność kanału kredytowego staje się tym niższa, im wyższy jest udział środków własnych w finansowaniu działalności przedsiębiorstw.

Kanał kursu walutowego polega na przenoszeniu wpływu impulsu polityki pieniężnej na kurs walutowy, którego zmiana wywołuje następnie odpowiednie zmiany w łącznym popycie i wydatkach. Załóżmy, że bank centralny podnosi oficjalne stopy procentowe. Impuls ten oddziałuje dwutorowo. Po pierwsze, wzrost stóp procentowych wywołuje spadek zagregowanego popytu, w tym popytu na import. Przy niezmienionym eksporcie, spadek importu przekłada się na poprawę stanu salda bilansu obrotów bieżących i w efekcie wywołuje aprecjację waluty krajowej. Umocnienie kursu waluty krajowej powoduje z kolei spadek cen towarów importowanych i będących przedmiotem eksportu, co ogranicza presję inflacyjną. W tym przypadku polityka pieniężna oddziałuje na kurs walutowy za pośrednictwem zmian na rachunku bieżącym bilansu płatniczego, będących skutkiem zmian zagregowanego popytu. Po drugie, wzrost stóp procentowych zwiększa dochodowość krajowych instrumentów dłużnych (w tym lokat bankowych), co pociąga za sobą napływ kapitału z zagranicy. Spowodowany tym wzrost podaży walut obcych wywołuje aprecjację kursu waluty krajowej, spadek cen towarów importowanych i będących przedmiotem eksportu, a w efekcie spadek presji inflacyjnej. W tym przypadku polityka pieniężna oddziałuje na kurs walutowy za pośrednictwem zmian na rachunku kapitałowym bilansu płatniczego, będących skutkiem napływu kapitału. W obu przypadkach warunkiem skuteczności funkcjonowania kanału kursowego jest system zmiennego kursu waluty.

Kanał oczekiwań przenosi impuls polityki pieniężnej na przewidywaną inflację. Uczestnicy rynku – banki, przedsiębiorstwa, gospodarstwa domowe – starają się przewidywać przyszłe warunki gospodarowania i zmiany w otoczeniu ekonomicznym. W ramach tego procesu starają się także przewidzieć przyszłą inflację i do tych oczekiwań dostosowują swoje bieżące oraz planowane decyzje. Przewidywania

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

te formułowane są na podstawie dostępnych informacji, w tym bieżących odczytów inflacji. Szczególnie ważny punkt odniesienia stanowią sygnały i decyzje pochodzące z banku centralnego. Wzrost oczekiwań inflacyjnych uruchamia zachowania „obronne" w postaci spadku popytu na pieniądz krajowy i wzrostu popytu na inne aktywa, takie jak nieruchomości, waluty obce, złoto, a także dobra trwałego użytku. Wzrost oczekiwań powoduje także zmianę zachowań konkurencyjnych uczestników rynku – dostawcy w większym stopniu skłonni są podwyższać ceny, a nabywcy w większym stopniu gotowi są te podwyżki akceptować. W efekcie wzrost cen przyspiesza, co może zmusić do zaostrzenia polityki pieniężnej. Dlatego też banki centralne starają się utrzymywać oczekiwania inflacyjne na niskim poziomie poprzez odpowiednie komunikowanie się z rynkami i budowanie wiarygodności w polityce pieniężnej. Kanał oczekiwań ma mniejsze znaczenie przy niskiej i stabilnej inflacji, lecz w warunkach wysokiej lub rosnącej inflacji oczekiwania ulegają utrwaleniu, co zagraża stabilności cen.

Nalży przy tym podkreślić, że związki między polityką banku centralnego i zmianami oczekiwań oraz między oczekiwaniami a rzeczywistą inflacją są złożone i nie w pełni rozpoznane. W dalszej części rozdziału przyjrzymy się bliżej mechanizmowi oczekiwań inflacyjnych.

Kanał majątkowy (*wealth channel*), zwany też kanałem cen aktywów (*asset-price channel*), przenosi skutki impulsu polityki pieniężnej na inflację poprzez wpływ na wartość aktywów w posiadaniu przedsiębiorstw i gospodarstw domowych. Gdy bank centralny podnosi stopy procentowe, zwiększa tym samym opłacalność lokat bankowych i rentowność instrumentów dłużnych. Oznacza to, że – zgodnie z opisaną wyżej zależnością między ceną a rentownością obligacji (ramka 5.7) – cena instrumentów dłużnych spada, co zmniejsza bieżącą wartość majątku posiadacza obligacji. Lecz wzrost stóp procentowych oznacza także, że relatywna dochodowość pozostałych aktywów, takich jak akcje czy udziały, maleje, a to pociąga za sobą ich wyprzedaż i – w konsekwencji – spadek cen. Z kolei spadek cen akcji i udziałów zmniejsza wartość rynkową przedsiębiorstwa w relacji do kosztu zastąpienia aparatu wytwórczego (relacja ta nosi nazwę współczynnika q Tobina). Sprawia to, że przedsiębiorstwa pozyskują mniej kapitału na rynku, co ogranicza inwestycje. Podobnie jest z nieruchomościami, których wartość spada, gdy rosną stopy procentowe. Co więcej, wzrost stóp procentowych podnosi koszty obsługi kredytów hipotecznych, a to zmusza osoby spłacające kredyty do ograniczenia innych wydatków. Wszystko to hamuje wzrost zagregowanego popytu i działa w kierunku łagodzenia presji inflacyjnej. Kanał majątkowy jest szczególnie skuteczny

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w gospodarkach, w których majątek gospodarstw domowych składa się w znacznej mierze z akcji i nieruchomości, a lokaty bankowe odgrywają mniejszą rolę. Taka struktura majątkowa gospodarstw domowych występuje na przykład w USA i w Wielkiej Brytanii. Z kolei w krajach Europy kontynentalnej, w tym w Polsce, udział akcji czy też nieruchomości finansowanych kredytem w całości aktywów pozostających w posiadaniu ludności jest niższy i w efekcie kanał majątkowy ma mniejsze znaczenie w mechanizmie transmisji.

Charakterystyka poszczególnych kanałów transmisji wskazuje, że ich „drożność", skuteczność i szybkość przenoszenia impulsów pieniężnych zależy od wielu czynników, na które bank centralny w krótkim okresie nie ma wpływu lub posiada jedynie wpływ ograniczony. Dotyczy to takich zjawisk jak stan nadpłynności w sektorze bankowym, udział kredytów walutowych w całości kredytu, zakres transgranicznych przepływów kapitału czy udział instrumentów udziałowych w oszczędnościach ludności.

Poszczególne kanały różnią się także pod względem czasu transmisji impulsów pieniężnych. Najszybszą transmisję zapewnia kanał kursu walutowego, zwłaszcza w jego części dotyczącej transgranicznego przepływu kapitału. Wzrost stóp procentowych wywołuje niemal natychmiastową reakcję kursu, a ta już w ciągu kilku lub kilkunastu tygodni pociąga za sobą zmiany cen części towarów importowanych, przede wszystkim tych, które nabywane są po cenach bieżących. Skalę wpływu zmian kursu walutowego na ceny krajowe mierzy tzw. efekt przenoszenia (efekt *pass-through*), który informuje o tym, o ile procent wzrosną (spadną) ceny krajowe, gdy kurs waluty obcej spadnie (wzrośnie) o 1% (Salvatore 2004, s. 562–566). Efekt przenoszenia staje się tym większy, im większy jest udział handlu zagranicznego w gospodarce oraz im mniejszy okazuje się poziom zapas towarów importowanych w kraju[52].

Rozpoznanie skali opóźnień występujących między impulsem polityki pieniężnej a efektem w postaci zmian w stopie inflacji ma kluczowe znaczenie dla zapewnienia skuteczności polityki pieniężnej. Ponieważ kanał kredytowy i kanał stopy procentowej wywierają wpływ na inflację poprzez wpływ na sektor realny (popyt i wydatki), zmiany w inflacji wystąpią dopiero po zmianach w sektorze realnym. Ponadto, ponieważ decyzje dotyczące wydatków nie zapadają natychmiast,

[52] Szacunki wielkości efektu *pass-trough* w odniesieniu do cen konsumpcyjnych w Polsce wskazują, że w ostatnich latach ulega on obniżeniu. Według wcześniejszych analiz, w latach 1998–2010 efekt ten wynosił około 0,2%, z kolei najnowsze badania dotyczące lat 2000–2013 wskazują, że nie przekracza on 0,1% (Kapuściński i in. 2014).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pojawiają się opóźnienia związane z dostosowaniem bieżących wydatków do nowych parametrów. Badania wskazują, że najsilniejszy efekt w zakresie wpływu zmian stopy procentowej na sektor realny w rozwiniętej gospodarce rynkowej występuje z opóźnieniem średnio 4–6 kwartałów, podczas gdy w przypadku inflacji opóźnienie to wynosi średnio 6–8 kwartałów. W Polsce efekty te zachodzą nieco szybciej, z opóźnieniem odpowiednio 3–5 kwartałów i 4–6 kwartałów, głównie ze względu na wciąż mniej płynny i słabiej rozwinięty sektor finansowy[53].

Oznacza to, że jeśli bank centralny podejmuje decyzję dziś, musi liczyć się z tym, że najsilniejszy efekt wystąpi dopiero za około rok–półtora. Ta opóźniona reakcja inflacji na decyzje banku centralnego sprawia, że polityka pieniężna musi odnosić się do przyszłego, przewidywanego stanu gospodarki – inaczej mówiąc, powinna być „skierowana ku przyszłości" (*forward looking*).

5.5.6. Oczekiwania inflacyjne

Skala negatywnego wpływu inflacji na wzrost gospodarczy zależy w dużym stopniu od tego, czy jest ona przewidywana, czy nie. Gdy ludzie nie oczekują inflacji (lub jej przyspieszenia), nie zmieniają swoich zachowań rynkowych. Występuje wówczas zjawisko iluzji pieniężnej (*money illusion*), które polega na tym, że uczestnicy rynku utożsamiają zmiany nominalne ze zmianami realnymi. Wówczas ujemny wpływ inflacji na wzrost gospodarczy jest niewielki lub żaden. Gdy pojawiają się jednak oczekiwania inflacyjne (*inflationary expectations*), ludzie zaczynają uwzględniać utratę realnej wartości pieniądza w swoich kalkulacjach. Sprzedawcy stają się bardziej skłonni podnosić ceny towarów, nabywcy stają się zaś gotowi je akceptować, konsumenci i producenci kupują „na zapas", pracownicy domagają się podwyżek płac, a pracodawcy skłonni są je przyznawać – wszyscy pozostają bowiem przekonani, że ceny będą rosły.

Oczekiwania inflacyjne mają wieloraki wpływ na zachowania podmiotów gospodarczych, kształtując funkcję popytu na dobra i usługi, funkcję popytu na aktywa finansowe i funkcję popytu na pieniądz. Wzrost oczekiwań inflacyjnych oznacza, że ludzie spodziewają się przyspieszenia inflacji w przyszłości. To przekonanie

[53] Symulacje przeprowadzone w Narodowym Banku Polskim pokazują, że maksymalny spadek inflacji po wzroście krótkoterminowej stopy procentowej o 1 punkt procentowy wynosi –0,7 punktu procentowego i następuje między piątym a szóstym kwartałem po impulsie. Efekt ten w ponad połowie można przypisać funkcjonowaniu kanału stopy procentowej. W pozostałej części wpływ ten wynika głównie z działania kanału kursowego i kredytowego (NBP 2014, s. 40–41).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ma wpływ na ich sposób postępowania – powoduje spadek popytu na pieniądz i wzrost popytu na towary i usługi oraz pozostałe aktywa (rzeczowe i finansowe), co z kolei pociąga za sobą wzrost rzeczywistej inflacji. W ten sposób uruchamiany jest mechanizm samospełniających się obaw. Ponadto, wzrost oczekiwań inflacyjnych wywołuje wzrost nominalnych stóp procentowych zarówno w wyniku uwzględnienia oczekiwanej inflacji w przyszłości, jak i na skutek wzrostu premii za ryzyko utraty wartości kapitału (ponieważ wzrost inflacji oznacza także wzrost zmienności cen). Wzrost oczekiwań inflacyjnych powoduje także zmianę zachowań uczestników rynku i osłabienie mechanizmu konkurencji. Przedsiębiorcy łatwiej akceptują podwyżki cen dostawców i podwyżki płac, ponieważ wiedzą, że są w stanie jednocześnie podnosić ceny własnych wyrobów. W ten sposób mechanizm konkurencji rynkowej ulega wypaczeniu i pojawia się niebezpieczeństwo rozkręcenia spirali cen i płac.

Dlaczego oczekiwania inflacyjne są ważne w przypadku polityki pieniężnej? Przede wszystkim dlatego, że wywierają one wpływ na zachowania uczestników rynku, a więc na procesy gospodarcze, a w tym także na samą inflację. Ze względu na wbudowany mechanizm samospełniających się obaw, oczekiwania wskazują również, jakiej inflacji można spodziewać się w przyszłości, jeśli bank centralny nie podejmie działań na rzecz jej ograniczenia. Dlatego oczekiwania są także dobrą miarą wiarygodności i reputacji banku centralnego – im bardziej wiarygodny jest bank centralny, tym niższe będą oczekiwania inflacyjne.

Z tych powodów oczekiwania inflacyjne od wielu lat pozostają w centrum uwagi banków centralnych. Zbierają one informacje na temat oczekiwań i systematycznie dokonują ich pomiaru, głównie w formie ankiet i wywiadów prowadzonych wśród głównych uczestników rynku (gospodarstwa domowe, banki, przedsiębiorstwa). Zebrane w ten sposób materiały są następnie analizowane w celu rozpoznania mechanizmów rządzących tymi zjawiskami.

Pojęcie oczekiwań inflacyjnych pojawiło się już kilkadziesiąt lat temu w pracach Keynesa (1933), Fishera (1930) i Pigou (1949), lecz po raz pierwszy w kontekście polityki pieniężnej zwrócili na nie uwagę Phelps (1967) i Friedman (1970, 1975). Czy możemy powiedzieć coś na temat sposobu, w jaki tworzą się oczekiwania? Ekonomiści wskazują najczęściej na trzy teoretycznie możliwe mechanizmy formowania się oczekiwań wśród uczestników rynku.

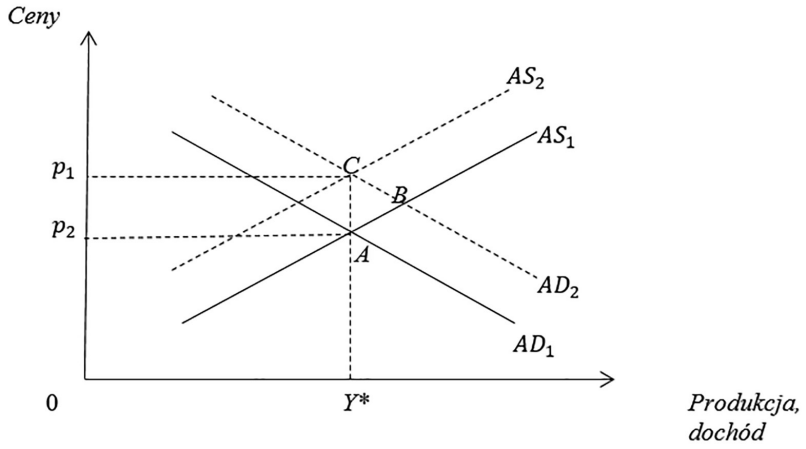
Oczekiwania statyczne kształtowane są przez dane odnoszące się do inflacji z poprzedniego okresu ($t-1$) i nazywane są także oczekiwaniami inercyjnymi. Przeciętne gospodarstwo domowe formuje swoje oczekiwania właśnie najczęściej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w sposób statyczny, to znaczy jest skłonne uważać, że inflacja w przyszłości ukształtuje się na poziomie zbliżonym do obserwowanego w chwili obecnej. Taki sposób kształtowania się oczekiwań wynika z tego, że gospodarstwa domowe nie analizują systematycznie informacji na temat zmian koniunktury gospodarczej i nie posiadają głębszej wiedzy na temat mechanizmu transmisji monetarnej i opóźnień występujących w polityce pieniężnej. Sytuacja ta, paradoksalnie, ułatwia wprawdzie wykorzystanie polityki pieniężnej dla pobudzania wzrostu gospodarczego, ale utrudnia walkę z inflacją. Bank centralny nie może ograniczyć się w takiej sytuacji do samych zapowiedzi obniżenia inflacji, ale musi rzeczywiście ją obniżyć, by trwale zakotwiczyć oczekiwania ludności na niższym poziomie.

Oczekiwania adaptacyjne (autoregresyjne) dostosowują się tylko częściowo do zachowań inflacji zaobserwowanych w przeszłości, a częściowo są kształtowane przez inne czynniki. Przedsiębiorstwa kształtują swoje oczekiwania najczęściej w sposób adaptacyjny, tzn. biorą stale pod uwagę najnowsze odczyty inflacji, choć uwzględniają też inne informacje. Taki sposób kształtowania oczekiwań wynika stąd, że w przedsiębiorstwach częściej mamy do czynienia z istnieniem wyspecjalizowanych służb analitycznych, odpowiedzialnych za prognozowanie sytuacji rynkowej i gospodarczej, w tym inflacji. W takich warunkach skuteczność polityki pieniężnej jest ograniczona – ma ona tylko krótkookresowy wpływ na zmienne realne, co utrudnia walkę z inflacją.

Rysunek 5.5. Oczekiwania inflacyjne w modelu AS-AD



Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ilustruje to rysunek 5.5, gdzie przedstawiono mechanizm tworzenia oczekiwań inflacyjnych w ramach prostego modelu AS-AD. Załóżmy, że gospodarka jest w punkcie równowagi *A*, któremu odpowiada potencjalny PKB na poziomie Y^*. Gdy bank centralny obniża stopę procentową, krzywa popytu przesuwa się z AD_1 do AD_2, a równowaga z punktu *A* do *B*. Lecz w punkcie *B* produkcja znajduje się powyżej poziomu potencjalnego Y^* i w rezultacie rośnie inflacja. W ślad za nią pojawiają się oczekiwania inflacyjne, które, zgodnie z mechanizmem krzywej Philipsa, powodują wycofanie części siły roboczej z rynku pracy i przesunięcie krzywej podaży z AS_1 do AS_2. Produkcja powraca do poziomu potencjalnego Y^*, lecz równowaga przesuwa się do punktu *C*, któremu odpowiada wyższy poziom cen p_2. Zmiana polityki pieniężnej miała więc tylko krótkookresowy wpływ na wzrost produkcji. Po pojawieniu się oczekiwań produkcja wraca do poprzedniego poziomu, a trwałym skutkiem ekspansji pieniężnej jest wzrost cen.

Trzecim sposobem formowania oczekiwań jest mechanizm **oczekiwań racjonalnych (antycypacyjnych)**. Hipoteza oczekiwań racjonalnych (Muth 1961) zakłada, że uczestnicy rynku wykorzystują całą dostępną informację rynkową dla prognozowania przyszłej inflacji, i są w stanie w pełni przewidzieć skutki działania banku centralnego. W takich warunkach nie występuje więc nawet krótkookresowy wpływ pobudzający polityki pieniężnej na zmienne realne. Nawiązując do rysunku 5.5, równowaga w gospodarce nie przesuwa się najpierw z punktu *A* do punktu *B,* a dopiero po pewnym czasie do punktu *C*, lecz bezpośrednio z punktu *A* do punktu *C*. W ramach hipotezy oczekiwań racjonalnych polityka pieniężna nie ma więc wpływu na zmienne realne i w skrajnym przypadku jest całkowicie nieskuteczna w pobudzaniu produkcji oraz zatrudnienia.

Hipoteza oczekiwań racjonalnych, choć intelektualnie atrakcyjna, opiera się jednak na bardzo restrykcyjnych założeniach[54]. W praktyce trudno uznać, że uczestnicy rynku mają dostęp do pełnej informacji o nim i znają zamiary banku centralnego oraz że wykorzystują tę wiedzę w swoich działaniach. Badania empiryczne wskazują, że ogólnie mamy raczej do czynienia jedynie z częściową racjonalnością (Kahneman i Tversky 2000; Łyziak 2003, 2012; Tomczyk 2008; Stanisławska 2010).

[54] Stopień racjonalności oczekiwań inflacyjnych ocenia się najczęściej z punktu widzenia realizacji dwóch warunków – warunku nieobciążoności (*unbiasedness*) i warunku ortogonalności błędów oczekiwań względem dostępnej informacji makroekonomicznej, zwanego też warunkiem makroekonomicznej efektywności (*macroeconomic efficiency*). Pierwszy z nich jest spełniony wówczas, gdy oczekiwania nie są obciążone błędem systematycznym, tzn. średnia oczekiwana stopa inflacji jest równa inflacji faktycznej. Z kolei drugi warunek oznacza, że przy formułowaniu oczekiwań w sposób efektywny brane są pod uwagę wszelkie dostępne informacje mające wpływ na przyszłą inflację (Łyziak 2012, s. 16).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wśród tych uczestników rynku, którzy najbardziej wnikliwie analizują dostępne informacje, znajdują się analitycy bankowi i finansowi. Mechanizm formowania oczekiwań inflacyjnych w tej grupie uczestników rynku jest stosunkowo najbliższy modelowi oczekiwań racjonalnych.

Obecność i charakter oczekiwań mają istotne znaczenie w przypadku kosztów polityki obniżania inflacji. Makroekonomiczne koszty polityki dezinflacyjnej poprzez podnoszenie stóp procentowych obejmują przede wszystkim koszt (przejściowego) spowolnienia lub spadku produkcji i spadku zatrudnienia, wywołanego podniesieniem stóp. Ten efekt polityki dezinflacji mierzy tzw. współczynnik wyrzeczenia (*sacrifice ratio*), definiowany jako wielkość procentowego spadku PKB na jeden punkt procentowy spadku inflacji. Skala tego kosztu zależy od mechanizmu oczekiwań oraz ich poziomu i trwałości. Gdy oczekiwania są adaptacyjne lub zwłaszcza statyczne, koszty dezinflacji okazują się wysokie, bo uczestnicy rynku kierują się w swych zachowaniach głównie obserwacjami z przeszłości, a nie tym, co bank centralny zapowiada i co zamierza zrobić. Deklaracje i zapowiedzi banku centralnego wskazujące na to, że poskromi inflację, nie są więc brane pod uwagę przez uczestników rynku. W takiej sytuacji bank musi znacznie podnieść stopy procentowe i utrzymywać je na wysokim poziomie przez dostatecznie długi czas, aby wywołany spowolnieniem gospodarczym spadek inflacji przełożył się ostatecznie na zmianę oczekiwań.

Gdy oczekiwania są racjonalne, koszt dezinflacji może być niski, bo już sama zapowiedź obniżenia inflacji wywołuje spadek oczekiwań i w efekcie spadek samej inflacji. W takich warunkach bank centralny nie musi drastycznie podnosić stóp procentowych. Warunkiem jest jednak wysoka wiarygodność banku centralnego w oczach uczestników rynku. Gdy bank pozostaje instytucją dostatecznie wiarygodną, może nawet w ogóle nie zaostrzać polityki pieniężnej, ponieważ uczestnicy rynku dają wiarę jego zapowiedziom i, obniżając swoje oczekiwania inflacyjne, dostosowują odpowiednio swoje zachowania. Współczynnik wyrzeczenia jest w tym przypadku niski[55].

Obecność racjonalnych oczekiwań inflacyjnych może niekiedy prowadzić do zjawiska, które nazywamy niespójnością czasową (dynamiczną) polityki pieniężnej. Jest to hipoteza, zgodnie z którą jakiekolwiek starania banku centralnego zmierzające do tego, aby przyspieszyć wzrost i zmniejszyć bezrobocie poniżej poziomu

[55] Szacunki dokonane przez NBP w przypadku Polski wskazują, że współczynnik wyrzeczenia odnośnie do oczekiwań statycznych wynosi 0,8% PKB na 1 punkt procentowy spadku inflacji, a odnośnie do oczekiwań w pełni racjonalnych tylko 0,05% PKB (Łyziak 2011, s. 114).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

długookresowej równowagi, są z góry skazane na niepowodzenie (Kydland i Prescott 1977). Jest to potencjalnie ważne ograniczenie skuteczności polityki pieniężnej.

5.6. Reguły polityki pieniężnej

5.6.1. Dyskrecjonalność polityki pieniężnej i zjawisko niespójności czasowej

Polityka pieniężna może być prowadzona w sposób dyskrecjonalny lub być oparta na określonych regułach[56]. W pierwszym przypadku bank centralny nie informuje uczestników rynku, w jaki sposób będzie realizował swoje cele, a w szczególności nie zobowiązuje się do stosowania jakichkolwiek stałych reguł w polityce pieniężnej. W drugim przypadku bank centralny ujawnia sposób, w jaki będzie podejmował decyzje, a przede wszystkim publicznie ogłasza reguły i zasady, zgodnie z którymi będzie kształtową politykę, reagując na informacje napływające z rynku.

Polityka dyskrecjonalna może wydawać się atrakcyjna z punktu widzenia władzy gospodarczej. W jej ramach władza nie jest skrępowana żadnymi – choćby dobrowolnymi – ograniczeniami wynikającymi z ogłoszonych wcześniej reguł i zobowiązań i dzięki temu dysponuje pełną swobodą działania. W szczególności może reagować na nieprzewidziane wydarzenia w sposób dowolny, niestandardowy, uwzględniając także doraźne kalkulacje polityczne. Immanentne wady takiego sposobu uprawiania polityki są jednak na tyle istotne, że koszty, jakie w polityce pieniężnej pociąga za sobą dyskrecjonalność, okazują się znacznie większe od ewentualnych korzyści.

Gdy polityka pieniężna jest prowadzona dyskrecjonalnie, uczestnicy rynków finansowych nie wiedzą, czym kieruje się bank centralny w podejmowaniu decyzji i w związku z tym nie są w stanie przewidzieć reakcji banku centralnego na napływające z gospodarki sygnały i informacje. W takich warunkach planowanie działalności gospodarczej, zwłaszcza inwestycji, jest utrudnione ze względu na podwyższoną niepewność co do działań banku centralnego. Większa niepewność pociąga za sobą wzrost premii za ryzyko w rynkowych stopach procentowych, co z kolei podnosi koszty działania firm i w warunkach awersji do ryzyka ogranicza aktywność gospodarczą. Ponadto dyskrecjonalność w polityce rodzi podejrzenia co do prawdziwych

[56] Na temat różnic między polityką dyskrecjonalną a polityką opartą na regułach w ujęciu generalnym zob. następny punkt.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

intencji banku centralnego. Nawet jeśli oficjalnym celem polityki pieniężnej ma być utrzymanie stabilności cen, uczestnicy rynku mają prawo obawiać się, że bank centralny może ulegać presji polityków i oportunistycznie wykorzystywać politykę pieniężną do realizacji innych celów, na przykład do pobudzania wzrostu i ograniczania bezrobocia kosztem zwiększania inflacji. Ponieważ wiadomo jednak, że przy pomocy polityki pieniężnej nie można w sposób trwały zwiększać produkcji i zatrudnienia, to taka „oportunistyczna" polityka może jedynie doprowadzić do wyższej inflacji. Dyskrecjonalność polityki zwiększa więc niepewność, w tym także ryzyko inflacji w przyszłości, co przekłada się na wzrost oczekiwań inflacyjnych (więcej na ten temat zob. m.in. Wojtyna 1999, Jakóbik 2008).

Polityka dyskrecjonalna wywołuje problem niespójności czasowej (lub dynamicznej – *time inconsistency, dynamic inconsistency*). Zjawisko to polega na tym, że zastosowanie określonej polityki, która w danych warunkach rynkowych i przy określonych zachowaniach uczestników rynku jest optymalna, pociąga za sobą zmianę tych zachowań, a w konsekwencji także potrzebę zmiany samej polityki. Inaczej mówiąc, optymalna polityka zastosowana w czasie t przestaje być optymalna w czasie $t + 1$ wskutek zmian zachowań uczestników rynku wywołanych samym zastosowaniem tej polityki. W ten sposób polityka optymalna przestaje być polityka spójną w czasie (*time consistent*), co oznacza, że optymalność *ex ante* przestaje być optymalnością *ex post* (Lucas 1976, Kydland i Prescot 1977, Barro i Gordon 1983, Marczewski 2008).

Jednym z najdalej idących ataków na dyskrecjonalną politykę gospodarczą jest tzw. krytyka Lucasa. Amerykański ekonomista Robert Lucas (1976) zakwestionował użyteczność modeli makroekonomicznych szacowanych na podstawie metod ekonometrycznych w celu oceny skutków zmian w polityce gospodarczej[57]. Ponieważ modele te są estymowane na podstawie danych historycznych, ich struktura i wartości parametrów odzwierciedlają określoną ścieżkę polityki gospodarczej zastosowanej w przeszłości. W szczególności, ponieważ systematycznie podejmowane (powtarzalne) decyzje z zakresu polityki gospodarczej wpływają na utrwalenie się określonych oczekiwań wśród uczestników rynku, modele te prezentują oczekiwane reakcje gospodarki pod warunkiem kontynuowania dotychczasowej polityki. Gdy dochodzi jednak do zmiany tej polityki, uczestnicy rynku dostosowują się do niej, co wywołuje zmianę parametrów modeli. W rezultacie nie można polegać na dostępnych modelach makroekonomicznych w celu oceny skutków

[57] Lucas (1976), s. 19–46.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zmian w polityce (zob. ramka 5.9). Lucas podważył w ten sposób sens modelowania makroekonomicznego – jednego z najbardziej obiecujących i zarazem najszybciej rozwijających się analitycznych sposobów oceny polityki gospodarczej[58].

Ramka 5.9. Krytyka Lucasa raz jeszcze

Krytyka Lucasa (*Lucas critique*) dotyczy głównie dużych, wielorównaniowych modeli makroekonomicznych o ogólnej postaci:

$$Y_t = f\left(X_t, Y_{t-1}, \varepsilon_t\right), \tag{5.27}$$

gdzie Y_t jest wektorem zmiennych ekonomicznych w czasie t, X_t to wektor zmiennych polityki gospodarczej, a ε_t oznacza wektor wartości składnika losowego. Oszacowanie modelu wymaga przekształcenia równania (5.27) w następującą postać:

$$Y_t = F\left(X_t, Y_{t-1}, \alpha\right) + u_t, \tag{5.28}$$

gdzie składowe wektora parametrów α są szacowane na podstawie wartości historycznych zmiennych $X_1 \ldots. X_T$ oraz $Y_1 \ldots. Y_T$. Gdy model ma być wykorzystany do oceny skutków nowej polityki, wprowadza się do niego nowe wartości wektora X_t. Taka procedura jest jednak uzasadniona tylko wówczas, gdy funkcja F i parametry α są stałe w czasie i nie zmieniają się wraz ze zmianami X_t. Tymczasem zachowanie uczestników rynku (wektor Y_t) zależy od ich oczekiwań co do przyszłych wartości zmiennych X_t, i jeśli wartości te ulegają zmianie, zmieniają się także ich oczekiwania. W szczególności decyzje uczestników rynku dotyczące konsumpcji czy inwestycji zależą od oczekiwań co do przyszłego stanu gospodarki. Na przykład, gdy następuje dewaluacja waluty krajowej, zmieniają się także oczekiwania co do przyszłej inflacji, a to z kolei wpływa na podejmowanie decyzji o inwestycjach. Ponieważ każda zmiana w dotychczasowej polityce wywołuje zmianę oczekiwań, oznacza to, że jednocześnie pociąga za sobą zmiany w samej strukturze modelu. Uniemożliwia to w praktyce wykorzystanie dotychczasowego modelu do oceny skutków zastosowania nowej polityki. Traktowana dosłownie, krytyka Lucasa prowadzi do bardzo pesymistycznego wniosku – wobec niemożności oceny skutków

[58] Więcej na ten temat zob. Taylor, s. 20–22; Benassy-Querée (2011), s. 80–82. Zob. także rozdział 1.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

własnych decyzji władza gospodarcza powinna powstrzymać się od prowadzenia jakiejkolwiek polityki.

Sens krytyki Lucasa można zilustrować często przywoływanym przykładem. Jak powszechnie wiadomo, rząd amerykański od wielu lat przechowuje swoje rezerwy złota w skarbcu w Fort Knox. Skarbiec jest bardzo silnie strzeżony i nigdy nie zdarzyło się, aby ktoś podjął próbę obrabowania skarbca. Na tej podstawie można byłoby wysnuć wniosek, że rezerwy złota są generalnie bezpieczne i dlatego można zmniejszyć stopień ochrony skarbca. Byłby to oczywiście wniosek bardzo ryzykowny, ponieważ potencjalni rabusie nie podejmują prób kradzieży właśnie dlatego, że Fort Knox jest tak dobrze chroniony, w związku z czym dokonanie udanej kradzieży jest bardzo mało prawdopodobne. Lecz zmiana polityki ochronnej i obniżenie poziomu ochrony, na przykład poprzez znaczną redukcję liczby strażników, skłoniłaby potencjalnych rabusiów do ponownej analizy szans powodzenia i, być może, podjęcia próby obrabowania skarbca. Innymi słowy, to, że nie dochodziło dotychczas do kradzieży nie oznacza, że nie dojdzie do nich wówczas, gdy polityka ochrony ulegnie zmianie[59].

Głęboko pesymistyczny wydźwięk krytyki Lucasa prowokował wielu ekonomistów do podejmowania polemiki i prób podważenia jego wniosków. Z jednej strony podjęto badania mające na celu uzupełnienie modeli makroekonomicznych zależnościami mikroekonomicznymi, objaśniającymi rzeczywiste zachowania uczestników rynku oraz uwzględniając długookresowe reguły ich postępowania, takie jak wygładzanie konsumpcji w ujęciu wielookresowym czy wpływ przejściowych zmian w podatkach na popyt. Z drugiej strony rozwinięto modele uwzględniające autoregresję zmiennych (modele VAR – *vector autoregressive*), które pozwalają na bardziej precyzyjne szacowanie zależności między zmiennymi (Sims 1980).

[59] Innym przykładem jest nadmiernie głęboka dewaluacja złotego dokonana w Polsce na początku transformacji (w styczniu 1990 roku). Reformatorzy polskiej gospodarki, na podstawie wieloletnich obserwacji „czarnego" rynku walutowego, byli przekonani o istnieniu stałego, dużego popytu na waluty wymienialne. Dlatego, wprowadzając wymienialność złotego i dysponując tylko ograniczonymi rezerwami walut wymienialnych, postanowili ustalić kurs dolara USA na odpowiednio wysokim poziomie, tak aby zapewnił on równowagę na rynku walutowym. Tymczasem okazało się, że popyt na waluty wymienialne drastycznie spadł w 1990 roku, właśnie dlatego, że wprowadzenie wiarygodnego programu reform i bardzo wysokie oprocentowanie depozytów złotowych skłoniło ludzi do zamiany dolarów na złote w portfelach. Ustalony kurs dolara okazał się zdecydowanie za wysoki, co przyczyniło się do pogłębienia przejściowej recesji i przyspieszyło inflację. Pozwala to zauważyć, że ocena skuteczności nowej polityki na podstawie historycznych obserwacji nie daje wiarygodnych wyników, ponieważ nie uwzględnia wpływu zmiany reżimu polityki na oczekiwania i zachowania uczestników rynku (więcej na ten temat w rozdziale 6).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 5.10. Problem niespójności czasowej: ujęcie z perspektywy funkcji popytu na pieniądz

Problem niespójności czasowej polityki pieniężnej można także opisać, wykorzystując standardowe równanie popytu na pieniądz (zob. równania 2.10 i 2.11b, rozdział 2). Załóżmy, że bank centralny zapowiada utrzymanie niskiej inflacji poprzez utrzymanie stałego tempa podaży pieniądza. Rynkowe stopy procentowe spadają i popyt na pieniądz rośnie zgodnie z daną funkcją popytu na pieniądz realny:

$$L/P = hY - kr, \tag{5.29}$$

gdzie r oznacza realną stopę procentową.

Wzrost popytu na pieniądz skłania jednak oportunistyczny bank centralny do zwiększenia podaży pieniądza w tempie szybszym niż rośnie popyt – w nadziei, że przyniesie to spadek bezrobocia i nie wywoła inflacji. Lecz wzrost podaży pieniądza większy niż wzrost popytu przynosi po pewnym czasie przyspieszenie inflacji, a następnie wzrost oczekiwań inflacyjnych. To wywołuje zarówno wzrost stóp procentowych r, jak i zmianę parametrów funkcji popytu na pieniądz, ponieważ spada współczynnik h i rośnie współczynnik k. W rezultacie popyt na pieniądz spada, nie tylko dlatego, że rośnie stopa procentowa r, lecz także dlatego, że zmieniają się parametry prawej strony funkcji (5.29). Oznacza to wzrost wydatków i w efekcie inflacja zaczyna rosnąć coraz szybciej. Polityka dyskrecjonalna staje się nieskuteczna w zwalczaniu inflacji.

Przykładem niespójności czasowej w ramach polityki pieniężnej jest właśnie nadmiernie „krótkowzroczna" polityka banku centralnego. Załóżmy, że bank zapowiada utrzymanie niskiej inflacji. Jeśli jest wiarygodny, oczekiwania inflacyjne spadną, a w ślad za nimi spadną także długookresowe stopy procentowe, co sprzyja wzrostowi gospodarczemu. Lecz wobec braku oczekiwań inflacyjnych opłacalne może okazać się zwiększenie podaży pieniądza w celu dalszego przyspieszenia wzrostu gospodarczego, ponad tempo potencjalne. Wzrost podaży pieniądza umożliwia początkowo wzrost zatrudnienia, spadek bezrobocia i wzrost produkcji. Wzrost zatrudnienia pociąga za sobą wzrost wynagrodzeń zarówno nominalnych, jak i realnych, ponieważ ceny początkowo pozostają stabilne. Bank centralny

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wykorzystuje więc zależność opisaną ujemnie nachyloną krótkookresową krzywą Philipsa, aby uzyskać dodatkowe korzyści w postaci wzrostu produkcji i zatrudnienia. Przy danym położeniu krzywej Philipsa zwiększanie podaży pieniądza staje się polityką optymalną, ponieważ powoduje spadek bezrobocia i wzrost produkcji, przy początkowo niewielkim wzroście inflacji.

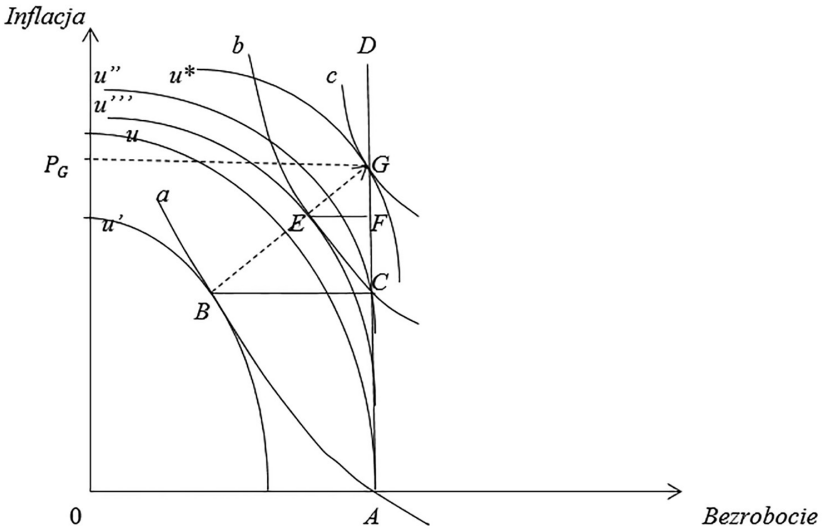
Przekonanie uczestników rynku, że nastąpił trwały wzrost *realnych* wynagrodzeń, jest jednak krótkotrwałe. Po pewnym czasie uświadamiają oni sobie, że, utożsamiając nominalny wzrost wynagrodzeń ze wzrostem realnym, padli ofiarą iluzji pieniężnej. Wzrost podaży pieniądza ponad poziom odpowiadający potencjalnemu wzrostowi PKB musi bowiem prędzej czy później doprowadzić do wzrostu inflacji. Z kolei wzrost inflacji spowoduje – po pewnym czasie – zanik iluzji pieniężnej, pojawienie się (lub wzrost) oczekiwań inflacyjnych, wzrost żądań płacowych i wzrost długookresowych stóp procentowych. W rezultacie produkcja spada, a bezrobocie ponownie rośnie, powracając do poprzedniego stanu. Krótkookresowa krzywa Philipsa przesuwa się w prawo i w górę. Polityka zwiększania podaży pieniądza przestaje być optymalna, ponieważ jej skutkiem jest wzrost inflacji i pojawienie się oczekiwań inflacyjnych, które powodują zanik korzyści wynikających z przejściowego wzrostu produkcji i zatrudnienia.

Takie „oszukiwanie" rynku jest możliwe tylko w warunkach polityki dyskrecjonalnej. Bank centralny swobodnie zwiększa podaż pieniądza, ponieważ nie czuje się skrępowany żadnymi regułami, i wykorzystuje doraźne możliwości chwilowej poprawy sytuacji gospodarczej. Lecz na dłuższą metę jest to polityka nieskuteczna i szkodliwa. Pojawienie się i wzrost oczekiwań inflacyjnych sprawia, że kolejne dawki podaży pieniądza nie dają już żadnych efektów realnych, a ich jedynym skutkiem jest wzrost inflacji. W efekcie w dłuższym okresie krzywa Philipsa zmienia swój kształt na pionowy.

Na rysunku 5.6 przedstawiono model sformułowany przez Barro i Gordona (1983), ilustrujący graficznie mechanizm funkcjonowania zjawiska niespójności czasowej. Na osi poziomej odłożono stopę bezrobocia, a na osi pionowej stopę inflacji. Krzywe *u* są krzywymi preferencji (użyteczności) władzy gospodarczej (banku centralnego). Punkty znajdujące się na danej krzywej *u* wyznaczają takie kombinacje stopy bezrobocia i stopy inflacji, które mają jednakową użyteczność z punktu widzenia władzy gospodarczej. Im niżej położona krzywa preferencji (im bliżej początku układu współrzędnych), tym większa użyteczność kombinacji bezrobocia oraz inflacji znajdujących się na danej krzywej. Krzywe *a*, *b* to krótkookresowe krzywe Philipsa, a pionowa prosta *AD* wyznacza poziom bezrobocia naturalnego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 5.6. Niespójność czasowa w polityce pieniężnej: model Barro-Gordona



Źródło: na podstawie Barro i Gordon (1983) oraz Kydland i Prescott (1976).

Załóżmy, że gospodarka znajduje się w punkcie *A*, gdzie inflacja jest równa zeru, a bezrobocie utrzymuje się na poziomie naturalnym (zob. ramka 5.11). Załóżmy również, że władza gospodarcza jest „krótkowzroczna" (na przykład jej horyzont decyzji jest ograniczony przez termin następnych wyborów) i że obiecuje utrzymanie inflacji na poziomie zero. Taka kombinacja inflacji i bezrobocia zapewnia władzy gospodarczej użyteczność na poziomie *u* (krzywa preferencji *u* przechodzi przez punkt *A*). Znajdując się jednak w punkcie *A* władza gospodarcza zyskuje oczywisty bodziec do zwiększenia podaży pieniądza, tak aby obniżyć bezrobocie kosztem wzrostu inflacji, przesuwając się wzdłuż krótkookresowej krzywej Phillipsa *a* aż do punktu *B*. Punkt *B* cechuje się znacznie korzystniejszą kombinacja inflacji i bezrobocia, ponieważ leży na niżej położonej krzywej preferencji *u'*. Początkowo wzrost wynagrodzeń wywołany wzrostem podaży pieniądza spowoduje wzrost podaży pracy, wzrost zatrudnienia i spadek bezrobocia. Gospodarka znajdzie się w punkcie *B*. Lecz wzrost inflacji ponad ogłoszony cel inflacyjny na poziomie zero uświadamia pracownikom, że za wzrostem wynagrodzeń nominalnych nie idzie wzrost wynagrodzeń realnych. Pojawiają się oczekiwania inflacyjne. W tej sytuacji część pracowników rezygnuje z pracy, i bezrobocie powraca do poziomu naturalnego. W rezultacie, w następnym okresie krzywa Phillipsa przesuwa się z *a* do *b*,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

a gospodarka przemieszcza się do punktu *C*. Lecz, ponieważ punkt *C* znajduje się na wyżej położonej krzywej preferencji *u"* niż punkt *A*, jego użyteczność jest niższa, bo przy tym samym poziomie bezrobocia występuje teraz inflacja. Załóżmy dalej, że niezrażona tym niepowodzeniem krótkowzroczna władza ponownie obiecuje stabilizację inflacji – tym razem na poziomie odpowiadającym punktowi *C*, lecz zarazem ma silną motywację, aby ponownie obniżyć bezrobocie kosztem zwiększenia inflacji, i w ten sposób ponownie „oszukać" pracowników. Zwiększając podaż pieniądza, władza gospodarcza obniża więc bezrobocie kosztem wzrostu inflacji, przesuwając się wzdłuż nowej krzywej Philipsa *b* z punktu *C* do punktu *E*, ponieważ punkt *E* leży na korzystniejszej krzywej preferencji *u'"*. Pracownicy szybko się jednak zorientują, że wynagrodzenia realne nie rosną, ponieważ wzrosła inflacja, i bezrobocie ponownie powróci do poziomu naturalnego w punkcie *F*. Władza gospodarcza będzie w stanie powtarzać ten manewr tak długo, aż gospodarka dojdzie do punktu *G* znajdującego się na styku krzywej preferencji u^* i krzywej Philipsa *c*. Bezrobocie powraca do poziomu naturalnego, tak jak w punkcie wyjścia (*A*), z kolei inflacja okaże się bardzo wysoka (P_G).

Ramka 5.11. Naturalna stopa bezrobocia i krzywa Philipsa

Naturalna stopa bezrobocia (NSB) mierzy bezrobocie, które występuje nawet wtedy, gdy rynek pracy pozostaje w równowadze. Oznacza to, że nawet jeśli podaż pracy równa się popytowi na pracę, istnieje pewna grupa pracowników, którzy nie są gotowi podjąć pracy przy danym poziomie wynagrodzeń realnych i wolą pozostać bezrobotni. Mamy wówczas do czynienia ze zjawiskiem niepełnego oczyszczania rynku pracy. Ekonomiści wskazują, że przyczyn takiej sytuacji może być wiele. Należą do nich tzw. bezrobocie frykcyjne, które pojawia się, gdy pracownik zmienia miejsce zatrudnienia i przez pewien czas pozostaje bez pracy (Pissarides 1990, Blanchard i Diamond 1989), zakres stosowania tzw. płac proefektywnościowych (Akerlof i Yellen 1986, Shapiro i Stiglitz 1984), skala korzyści płacowych z tytułu przynależności do związków zawodowych (Summers 1988), a także instytucjonalne rozwiązania istniejące na rynku pracy (polityka płacy minimalnej, zasiłki dla bezrobotnych – zob. Hall i Taylor 2000). NSB to zatem taka stopa bezrobocia, w przypadku której zatrudnieni są wszyscy pracownicy, którzy chcą pracować przy danym poziomie realnych wynagrodzeń. Gdy bezrobocie znajduje się powyżej NSB, pojawia się presja deflacyjna (recesja), a kiedy jest poniżej tego

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

progu, występuje presja inflacyjna. Oznacza to, że NSB odpowiada potencjalnemu tempu wzrostu PKB.

Koncepcja naturalnej stopy bezrobocia pojawiła się w kontekście krytyki keynesowskiej polityki stabilizacyjnej, nakierowanej na wzrost zatrudnienia. Friedman wskazał, że zależność Phillipsa opisująca związek między stopą inflacji a stopą bezrobocia ma charakter krótkookresowy i że w dłuższym czasie nie jest możliwe sprowadzenie bezrobocia poniżej pewnej granicy, którą nazwał naturalną stopą bezrobocia. Istnienie NSB uniemożliwia więc realizację postulatu pełnego zatrudnienia – gospodarka w dłuższym okresie dąży do poziomu NSB.

Rysunek 5.6 pozwala zauważyć, że jakakolwiek próba obniżenia stopy bezrobocia poniżej poziomu naturalnego wywołuje wzrost inflacji. Dlatego też naturalną stopę bezrobocia nazywa się alternatywnie stopą bezrobocia nie przyspieszającą inflacji lub stopą NAIRU (*Non-Accelerating Inflation Rate of Unemployment*). Jest to stopa długookresowej równowagi na rynku pracy, występująca przy założeniu istnienia oczekiwań inflacyjnych, które sprawiają, że płace i ceny rosną w tym samym tempie.

NSB stanowi ważną wartość referencyjną dla polityka gospodarczego, ponieważ, porównując NSB z rzeczywistym poziomem bezrobocia, może on ustalić, ile wynosi bezrobocie wymuszone i czy działania na rzecz przyspieszenia wzrostu przyniosą zatem trwały spadek bezrobocia. Badania empiryczne wskazują, że NSB przyjmuje odmienne wartości odnośnie do różnych krajów. Co więcej, NSB nie jest wielkością stałą w czasie i może ulegać zmianom pod wpływem oddziaływania innowacji technologicznych i zmian w gustach uczestników rynku (teoria realnego cyklu koniunkturalnego). Na przykład w latach 1970–1990 naturalna stopa bezrobocia w gospodarce amerykańskiej była szacowana na poziomie około 5–6% (Hall i Taylor 2000, s. 143).

Próby trwałego obniżenia bezrobocia przy pomocy dyskrecjonalnej polityki pieniężnej okazują się więc nieskuteczne. Sprawiają, że w każdym następnym okresie polityka zastosowana uprzednio staje się nieoptymalna. Na skutek wzrostu oczekiwań inflacyjnych krótkookresowa krzywa Philipsa z okresu na okres przesuwa się do góry. W rezultacie w długim okresie zmienia ona swój kształt z opadającego na pionowy i pokrywa się z linią *ACFD* (zob. rysunek 5.6 i ramka 5.11).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 5.12. Niespójność czasowa polityki według Kydlanda i Prescotta (1976)*

W swoim słynnym eseju z 1976 roku Kydland i Prescott pokazali, że w warunkach występowania racjonalnych oczekiwań polityka optymalna w danym okresie jest polityką niespójną w czasie. Przyjmijmy za nimi, że $\pi = (\pi_1, \pi_2, \ldots, \pi_T)$ jest ciągiem decyzji polityki gospodarczej w kolejnych okresach od 1 do T, a $x = (x_1, x_2, \ldots, x_T)$ stanowi ciąg zachowań uczestników rynku. Zachowania te w danym okresie t zależą od wcześniejszych zachowań oraz od wszystkich decyzji podejmowanych w zakresie polityki – nie tylko wcześniejszych, lecz także przyszłych, ponieważ uczestnicy rynku starają się w sposób racjonalny przewidzieć przyszłe decyzje władzy gospodarczej i dostosowują do nich swoje zachowania. W ten sposób otrzymujemy więc:

$$x_t = X_t\left(x_1, \ldots, x_{t-1}, \pi_1, \ldots, \pi_T\right), \quad \text{gdzie } t = 1, 2, \ldots, T. \tag{5.30}$$

Istnieje także funkcja użyteczności społecznej polityki o postaci ogólnej:

$$U\left(x_1, \ldots, x_T, \pi_1, \ldots, \pi_T\right). \tag{5.31}$$

Wówczas polityką optymalną (*optimal policy*) będzie taki ciąg decyzji π, dla którego funkcja U osiąga maksimum przy ograniczeniu równania (5.29). Z kolei polityką spójną (*consistent policy*) stanie się taki ciąg decyzji π, dla którego funkcja U osiąga maksimum, lecz dla danych dotyczących przeszłych zachowań $(x_1, \ldots, x_{t-1})$, oraz jeśli przyszłe decyzje π_s dla okresów $s > t$ są podejmowane w taki sam sposób jak dotychczas.

Załóżmy, że dane są tylko dwa okresy, $t = 1, 2$. Dla okresu $T = 2$, polityka π_2 jest dobrana tak, aby maksymalizować funkcję U:

$$U\left(x_1, x_2, \pi_1, \pi_2\right) \tag{5.32}$$

przy ograniczeniach:

$$x_1 = X_1\left(\pi_1, \pi_2\right)$$

$$x_2 = X_2\left(x_1, \pi_1, \pi_2\right). \tag{5.33}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Warunkiem pierwszego rzędu w przypadku maksimum funkcji (5.32) jest, aby pochodna funkcji U była równa zeru:

$$\frac{\partial U}{\partial x_2}\frac{\partial X_2}{\partial \pi_2}+\frac{\partial U}{\partial \pi_2}=0. \tag{5.34}$$

Warunek równania (5.34) odpowiada polityce optymalnej dla danych *niezależnych* wartości π_1 oraz x_1. Jeśli pojawiają się jednak oczekiwania dotyczące *przyszłych* decyzji polityki, wartość x_1 staje się także zależna od π_2. Uwzględnienie tej zależności zmienia postać warunku pierwszego rzędu. Wówczas dla danego π_1 funkcja $U(\cdot)$ osiąga maksimum, gdy:

$$\frac{\partial U}{\partial x_2}\frac{\partial X_2}{\partial \pi_2}+\frac{\partial U}{\partial \pi_2}+\frac{\partial X_1}{\partial \pi_2}\left(\frac{\partial U}{\partial x_1}+\frac{\partial U}{\partial x_2}\frac{\partial X_2}{\partial x_1}\right)=0. \tag{5.35}$$

Równania (5.33) stanowią warunek niezbędny, aby polityka była spójna w czasie (*time-consistent*). Lecz jednocześnie równania (5.34) i (5.35) będą spełnione (czyli pochodne będą równe zeru) tylko wtedy, gdy:

a) albo $\frac{\partial X_1}{\partial \pi_2}=0$, czyli decyzje w okresie $t=2$ (π_2) nie mają wpływu na zachowania w okresie $t=1$ (x_1), co oznacza, że nie istnieją racjonalne oczekiwania;
b) albo wyrażenie w nawiasie w równaniu (5.35) jest równe zeru, czyli wpływ zachowań w okresie $t=1$ (x_1) nie ma wpływu na funkcję U, zarówno bezpośrednio ($\frac{\partial U}{\partial x_1}=0$), jak i za pośrednictwem zachowań w okresie $t=2$ ($\frac{\partial U}{\partial x_2}\frac{\partial X_2}{\partial x_1}=0$).

Obie te sytuacje są bardzo mało prawdopodobne, ponieważ w rzeczywistości mamy do czynienia z oczekiwaniami, które przynajmniej częściowo są racjonalne, a zarazem funkcja U zależy od zachowań uczestników rynku. Tak więc polityka optymalna w danym momencie jest z reguły niespójna w czasie. Ilustracją graficzną niespójności czasowej jest rysunek 5.6, gdzie punktem polityki optymalnej jest punkt 0 (funkcja użyteczności społecznej osiąga maksimum w przypadku

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zerowych wartości inflacji i bezrobocia[60]), a punktem polityki spójnej w czasie jest punkt *G*, gdzie władza gospodarcza nie czuje pokusy „oszukania" uczestników rynku (krzywa użyteczności społecznej jest styczna do krzywej Phillipsa).

5.6.2. Rozwiązanie problemu niespójności czasowej w obrębie jednego kraju. Niezależny bank centralny

W poprzednim punkcie stwierdziliśmy, że władza gospodarcza, której preferencje obejmują obniżanie inflacji i obniżanie bezrobocia, może znaleźć motywację do tego, aby, zwiększając inflację ponad ustalony poziom, obniżać stopę bezrobocia poniżej poziomu naturalnego. Wiemy już, że korzyści wynikające z takiego manewru mają co najwyżej charakter krótkookresowy i z reguły nie wystąpią wówczas, gdy uczestnicy rynku będą w stanie przewidywać postępowanie władzy gospodarczej. Warto więc zadać sobie pytanie, czy w ogóle możliwa jest sytuacja, w której władza gospodarcza nie będzie ulegać pokusie „oszukania" uczestników rynku w celu osiągnięcia krótkookresowej korzyści w postaci przejścia na korzystniejszą krzywą preferencji?

Teoretycznie można wyobrazić sobie dwie takie sytuacje. Po pierwsze, krótkookresowe korzyści nie będą kusiły władzy gospodarczej wówczas, gdy nie będzie ona przywiązywać w ogóle wagi do bezrobocia. Wtedy krzywe preferencji będą poziome i władza gospodarcza pozostanie w punkcie *A*. Lecz w praktyce jest mało prawdopodobne, aby władza gospodarcza nie była zainteresowana zatrudnieniem i bezrobociem – w systemie politycznym demokracji parlamentarnej jest to nierealne. Losy polityczne władzy gospodarczej zależą bowiem w państwie demokratycznym od głosów wyborców, a ci uzależniają swoje wybory od sytuacji na rynku pracy. Powszechna świadomość tego, że każdej władzy gospodarczej zależy na obniżaniu bezrobocia, sprawia, że polityka zerowej inflacji lub jakiejkolwiek stałej inflacji poniżej punktu *G*, będzie niewiarygodna w warunkach istnienia racjonalnych podmiotów i krótkowzrocznej władzy.

Druga sytuacja może mieć miejsce wówczas, gdy wzrost inflacji wiąże się z dużym kosztem politycznym (na przykład utrata reputacji). Jeśli ten koszt jest

[60] To pewne uproszczenie. Wspomnieliśmy wyżej, że zerowa inflacja nie jest pożądanym celem; podobnie można też przyjąć, że pożądanym społecznie poziomem bezrobocia jest naturalna stopa bezrobocia. Można więc zmienną π traktować jako odchylenie inflacji od celu inflacyjnego. Nie ma to wpływu na końcowy wniosek.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

większy niż odległość pomiędzy krzywymi preferencji *u* i *u'*, władza gospodarcza może rzeczywiście realizować politykę niskiej inflacji. Doświadczenie pokazuje jednak, że są to rzadkie przypadki. Jeśli koszt ten jest więc mniejszy, uczestnicy rynku nie uwierzą w obietnice składane przez władzę i zwiększą swoje oczekiwania inflacyjne.

Z dotychczasowej dyskusji wynika zatem, że istnienie problemu niespójności czasowej uzasadnia przekazanie polityki pieniężnej w gestię odrębnej instytucji, niezależnej od władzy politycznej, która rzeczywiście nie przywiązuje wagi do bezrobocia (zob. ramka 5.13). Wówczas polityka utrzymania inflacji na niskim poziomie staje się trwale możliwa. W krajach rozwiniętych taką instytucją niezależną od polityków jest bank centralny, wyposażony w mandat, który nakłada na niego obowiązek utrzymania stabilności cen (Wojtyna 1998).

Amerykański ekonomista Kenneth Rogoff zaproponował istotną korektę powyższego wniosku. Pokazał on, że bank centralny odpowiedzialny za politykę pieniężną nie powinien całkowicie ignorować bezrobocia, ponieważ bezwzględne dążenie do realizacji celu inflacyjnego kosztem dużego wzrostu bezrobocia może pociągać za sobą tak duże koszty ekonomiczne, że postępowanie to okaże się społecznie nieoptymalne. Rogoff (1985) wykazał, że w celu usunięcia niespójności czasowej wystarczy postawić na czele banku centralnego osobę, która pod względem awersji do inflacji jest bardziej konserwatywna niż reszta społeczeństwa, lecz nie jest też „nieskończenie" konserwatywna, czyli nie ignoruje całkowicie bezrobocia (zob. ramka 5.13). Istotę rozumowania Rogoffa można przedstawić, wykorzystując znany model AS-AD (na rysunku 5.7).

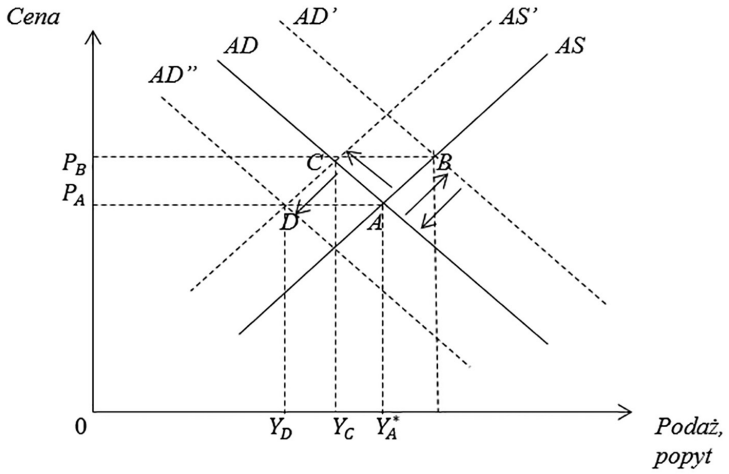
W początkowym punkcie równowagi podaży i popytu gospodarka wytwarza produkt na poziomie potencjalnego PKB równym Y_A^* – załóżmy, że poziom ten nie zmienia się w czasie. Przyjmijmy następnie, że ma miejsce dodatni szok popytowy w postaci egzogenicznego wzrostu łącznego popytu, co wywołuje przesunięcie krzywej *AD* do położenia *AD'*. Nowym punktem równowagi jest *B*. Skutkiem szoku jest wzrost cen z P_A do P_B i wzrost produkcji ponad poziom potencjalny ($Y_B > Y_A^*$). Skłania to bank centralny do podniesienia stóp procentowych, co pozwala jednocześnie obniżyć ceny i obniżyć produkcję do poziomu potencjalnego (powrót z punktu *B* do *A*). Działanie banku centralnego ma podwójny efekt stabilizujący – zarówno w odniesieniu do cen, jak i w odniesieniu do produkcji. W takiej sytuacji nie występuje konflikt celów.

Załóżmy obecnie, że mamy do czynienia z negatywnym szokiem podażowym w postaci egzogenicznego wzrostu kosztów wytwarzania. Krzywa łącznej podaży

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przesuwa się z *AS* do *AS'*, co wywołuje wzrost cen i spadek produkcji poniżej „potencjału" (z Y_A^* do Y_C). Nowym punktem równowagi jest *C*. Skłania to bank centralny do podniesienia stóp procentowych, co pozwala wprawdzie obniżyć ceny, ale jeszcze bardziej ogranicza produkcję, ponieważ pod wpływem wzrostu stóp procentowych krzywa *AD* przesuwa się do położenia *AD"*. Końcowym punktem równowagi jest *D*. Jak widać, w tej sytuacji mamy do czynienia z konfliktem celów: cel stabilizacji cen kłóci się z celem stabilizacji produkcji. Jeśli bank centralny jest bardzo konserwatywny i chce powrócić do poprzedniego poziomu cen P_A (lub poprzedniego poziomu inflacji), spadek produkcji będzie większy (z Y_A^* do Y_D). Ogólnie rzecz biorąc, straty makroekonomiczne są tym większe, im bardziej konserwatywny jest bank centralny.

Rysunek 5.7. Polityka pieniężna, szoki popytowe i podażowe oraz optymalny „konserwatyzm" banku centralnego



Źródło: opracowanie własne.

Dlatego też zbyt rygorystyczne zwalczanie inflacji wiąże się z dużym kosztem makroekonomicznym. Zauważmy, że wielkość utraconej produkcji i zatrudnienia w wyniku stabilizacji cen przez bank centralny jest tym większa, im bardziej płaska staje się krzywa *AS*, czyli im bardziej elastyczna jest podaż względem ceny. W niektórych przypadkach różnica między Y_A^* i Y_D może być na tyle duża, że bezwzględne realizowanie celu w postaci stabilności cen jest społecznie nieopłacalne. Dlatego istnieje pewien optymalny poziom „konserwatyzmu" banku centralnego

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

– jest on wyższy niż średnia opinii publicznej, lecz na pewno nie pozostaje nieskończenie wysoki.

Ramka 5.13. Optymalny poziom konserwatyzmu bankierów centralnych

Kenneth Rogoff (1985) wykazał, że opłaca się postawić na czele banku centralnego osobę, która pod względem awersji do inflacji jest bardziej konserwatywna niż reszta społeczeństwa, lecz jednocześnie nie ignoruje całkowicie bezrobocia. Punktem wyjścia było spostrzeżenie, że nawet gdy bank centralny optymalizuje społeczną funkcję preferencji wobec bezrobocia i inflacji, inflacja będzie systematycznie za wysoka (zob. m.in. Phelps 1967, Kydland i Prescott 1977 oraz Barro i Gordon 1983; zob. też rysunek 5.4). Rogoff wskazuje, że im wyższa waga przypisana jest inflacji w relacji do zatrudnienia (bezrobocia) w funkcji preferencji, tym bardziej bank centralny będzie skłonny troszczyć się o stabilność cen. Lecz jest z tym też związany koszt braku dostatecznie silnej reakcji banku centralnego na szoki podażowe, zmniejszające zatrudnienie. W warunkach sztywnych płac i cen polityka pieniężna może bowiem pełnić w krótkim okresie skutecznie rolę stabilizacyjną. Dlatego korzyść w postaci obniżania inflacji powinna być porównywana z kosztem w postaci braku polityki stabilizacyjnej. Funkcję preferencji banku centralnego można zapisać ogólnie w następujący sposób:

$$U = (n - \tilde{n})^2 + (\chi + \varepsilon)(p - \pi)^2 = min, \tag{5.36}$$

gdzie $(n - \tilde{n})$ jest różnicą między aktualną stopą zatrudnienia a społecznie optymalną stopą zatrudnienia, $(p - \pi)$ stanowi różnicę między aktualną inflacją a celem inflacyjnym, χ jest „społecznie uśrednioną" wagą przypisaną inflacji względem zatrudnienia, a $(\chi + \varepsilon)$ to waga przypisaną inflacji przez „konserwatywnego" bankiera centralnego, przy czym $\varepsilon > 0$. Im większa wartość ε, tym bardziej „konserwatywny" jest bank centralny w porównaniu z resztą społeczeństwa, lecz jednocześnie tym większy staje się koszt szoków podażowych względem korzyści z obniżania inflacji – po przekroczeniu pewnego poziomu ε koszt spadku zatrudnienia na skutek szoku podażowego jest większy niż korzyść ze stabilnej inflacji. Wynika z tego, że optymalna wielkość ε pozostaje dodatnia, ale zawsze mniejsza od nieskończoności, co oznacza, że istnieje również pewien optymalny stopień „konserwatyzmu"

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

bankierów centralnych w polityce pieniężnej, którego przekroczenie wywołuje straty społeczne[61].

Wspomnieliśmy wcześniej, że polityka oparta na publicznie ogłoszonych regułach powoduje spadek niepewności, co zmniejsza premie za ryzyko i obniża rynkowe stopy procentowe. W warunkach awersji do ryzyka spadek niepewności przyczynia się do zwiększenia produkcji, inwestycji i zatrudnienia oraz do poprawy dobrobytu (*welfare*). Lecz skuteczność reguł w realizacji celów polityki zależy od wielu czynników, w tym od wiarygodności banku centralnego. Bank centralny będzie wiarygodny w oczach uczestników rynku, gdy przestrzega własnych reguł i gdy jego polityka jest skuteczna, czyli osiąga postawione cele. Zależy to głównie od dotychczasowych dokonań banku, co oznacza, że wiarygodność buduje się powoli, w długim okresie czasu.

Wiarygodność (*credibility*) ma konkretnie dwa wymiary. Po pierwsze, rozpatrujemy ją z punktu widzenia zgodności działań z wcześniejszymi ich zapowiedziami– jest to wiarygodność *intencji*, czyli miara uczciwości władzy. Po drugie, wiarygodność rozpatrujemy także z punktu widzenia zgodności stopnia wykonania założonych celów – to wiarygodność *polityki*, czyli miara skuteczności władzy. Bank centralny jest oceniany przez uczestników rynku poprzez pryzmat obu wymiarów – powinien działać tak, jak zapowiedział, i osiągać założone cele. W praktyce nie zawsze udaje się zrealizować oba wymiary wiarygodności. Gdy bank postępuje inaczej niż zapowiadał, powstają wątpliwości co do jego prawdziwych intencji; gdy nie realizuje zaś przyjętych celów, powstają wątpliwości co do jego kompetencji.

Wiarygodność w polityce pieniężnej jest ważna nie tylko dlatego, że zwiększa zaufanie, obniża niepewność i, przyczyniając się do wzrostu aktywności gospodarczej, poprawia dobrobyt. Wiarygodność zmniejsza także makroekonomiczne koszty polityki obniżania inflacji do pożądanego poziomu. Jak wspomnieliśmy wyżej, koszt makroekonomiczny polityki dezinflacji to wielkość utraconej produkcji

[61] Wyprowadzając warunek pierwszego rządu dla (5.35) i przekształcając, otrzymujemy:

$$\frac{n-\tilde{n}}{p-\pi}=\chi-\varepsilon,$$

a ponieważ lewa strona jest zawsze dodatnia, bo licznik i mianownik są tego samego znaku, to dla $p \neq \pi$ zachodzi zawsze $0<\varepsilon<\infty$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

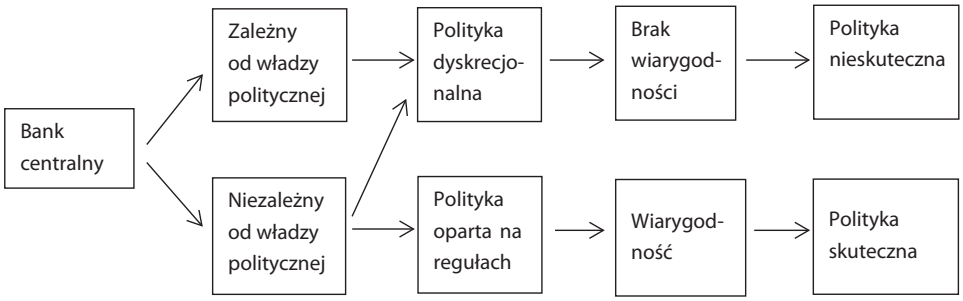
i zatrudnienia, jaka jest spowodowana niezbędnym zaostrzeniem polityki pieniężnej. Gdy bank centralny pozostaje wiarygodny, nawet niewielkie podniesienie oficjalnych (krótkookresowych) stóp procentowych, a nawet sama zapowiedź podwyżki, powoduje spadek oczekiwań inflacyjnych długookresowych rynkowych stóp procentowych, co pozwala uniknąć spadku inwestycji, produkcji i zatrudnienia. Lecz kiedy bank centralny nie jest wiarygodny, podniesienie krótkookresowych stóp procentowych pociąga za sobą spadek oczekiwań dopiero po dłuższym czasie i tylko w ograniczonym stopniu, a w efekcie długookresowe stopy rynkowe nie spadają lub spadają wolniej, co to ogranicza inwestycje, produkcję i zatrudnienie. W tym drugim przypadku w celu zwalczenia inflacji potrzebne jest więc silniejsze podniesienie stóp procentowych i dłuższy okres stosowania restrykcyjnej polityki pieniężnej, tak aby przekonać nieufnych uczestników rynku aby uwierzyli w determinację banku centralnego. Pociąga to za sobą większy koszt makroekonomiczny obniżenia inflacji do pożądanego poziomu.

Warunkiem niezbędnym do zapewnienia wiarygodności banku centralnego jest jego niezależność od władzy politycznej. Jak podkreśliliśmy, władza polityczna w warunkach demokracji pochodzi z wyborów. Okresowe wybory zmuszają polityków do zabiegania o głosy wyborców, co wiąże się ze składaniem obietnic i podejmowaniem działań, które są przychylnie odbierane przez wyborców. W warunkach istnienia krótkookresowej wymienności pomiędzy inflacją a bezrobociem krzywa preferencji władzy politycznej ma zawsze nachylenie ujemne, ponieważ opłaca się obniżać bezrobocie kosztem wzrostu inflacji (przynajmniej do pewnego poziomu). Rozwiązaniem dylematu jest powierzenie kompetencji w zakresie polityki pieniężnej niezależnemu bankowi centralnemu oraz formalne określenie jego mandatu w taki sposób, aby był on konstytucyjnie odpowiedzialny za utrzymanie stabilności cen. Są to dwa niezbędne warunki zapewnienia wiarygodności banku centralnego i skuteczności prowadzonej przez niego polityki pieniężnej. Rysunek 5.8 pokazuje schematycznie zależność skuteczności polityki od wiarygodności i niezależności banku centralnego[62].

[62] Więcej na temat niezależności banku centralnego w systemie demokratycznym zob. Wojtyna (2002).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 5.8. Niezależność, wiarygodność i skuteczność polityki pieniężnej



Źródło: opracowanie własne.

5.6.3. Podstawowe reguły polityki pieniężnej

Najwcześniej stosowaną regułą polityki pieniężnej był XIX-wieczny system waluty opartej na złocie (*gold standard*), w ramach którego podstawowym zadaniem polityki pieniężnej było utrzymanie stabilnej relacji waluty krajowej do jednostki kruszcu, przede wszystkim złota, zgodnie z oficjalnym parytetem. Taki system zapewniał stabilność cen, ponieważ ceny kruszcu w normalnych czasach były dosyć stabilne (ze względu na względnie stabilną podaż kruszców). Poza zabieganiem o utrzymanie stabilnego parytetu złota lub srebra i zapewnieniem niezakłóconej wymiany pieniądza na kruszec banki centralne praktycznie nie zajmowały się niczym innym.

System walutowy oparty na złocie został zawieszony w czasie I wojny światowej, co było jedną z przyczyn hiperinflacji w latach 20. ubiegłego wieku, a następnie załamał się całkowicie w trakcie wielkiego kryzysu lat 1929–1933. Po II wojnie światowej, podczas konferencji w amerykańskim kurorcie Bretton Woods w 1944 roku, kraje rozwinięte postanowiły utworzyć międzynarodowy system walutowy, który miał być pewną imitacją tradycyjnego systemu waluty złotej. W ramach tej strategii poszczególne kraje ustalały parytet swojej waluty do dolara amerykańskiego, z kolei sam dolar miał ustalony parytet do złota i – teoretycznie – pozostawał w pełni wymienialny na złoto. System ten załamał się jednak na początku lat 70., głównie z powodu prowadzenia przez Stany Zjednoczone bardzo ekspansywnej polityki fiskalnej i pieniężnej, co spowodowało przyspieszenie inflacji i zmusiło USA do zawieszenia wymienialności dolara na złoto.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Skłoniło to rządy i banki centralne w krajach rozwiniętych do porzucenia systemu opartego na dolarze amerykańskim i do zmiany sposobu prowadzenia polityki pieniężnej. Oprócz przyjęcia stabilności cen jako celu polityki uznano, że istnieje systematyczny związek między inflacją a wzrostem masy pieniądza w obiegu. Dlatego w kolejnych latach starano się prowadzić politykę pieniężną w taki sposób, aby utrzymywać stałe, umiarkowane tempo wzrostu podaży pieniądza, co miało – zgodnie z postulatem monetarystów – zapewnić stabilność cen. Przy czym ogólną podaż pieniądza w gospodarce można było regulować w sposób bezpośredni, określając wielkość podaży pieniądza rezerwowego w ramach operacji otwartego rynku (regulacja ilościowa), lub w sposób pośredni, ustalając krótkoterminową stopę procentową, jako cenę dostępu do płynności (regulacja cenowa). W ten sposób wykształciły się reguły prowadzenia polityki pieniężnej (*monetary policy rules*) w oparciu o cele pośrednie, takie jak tempo wzrostu podaży pieniądza i poziom stóp procentowych.

W praktyce reguła celu pośredniego w postaci podaży pieniądza (*intermediate monetary target*) polega na tym, że bank centralny najpierw szacuje długookresowy związek między zmianami podaży pieniądza i zmianami poziomu cen, a następnie ogłasza pożądane tempo wzrostu podaży określonego agregatu pieniężnego (M1, M2, M3), spójne z postulatem utrzymania stabilności cen. Przy założeniu, że relacja ta rzeczywiście pozostaje stabilna w czasie, utrzymanie stałego tempa wzrostu podaży pieniądza gwarantuje utrzymanie stabilnego poziomu inflacji. Tempo to może być z reguły nieco wyższe niż tempo wzrostu cen, ponieważ w miarę rozwoju gospodarczego i wzrostu zamożności społeczeństwa rosną zasoby pieniężne ludności, rośnie stopień monetyzacji gospodarki i spada szybkość obiegu pieniężnego. Ogłoszone publicznie tempo wzrostu podaży pieniądza pełni więc funkcję „kotwicy nominalnej" (*nominal anchor*), która – podobnie jak parytet złota w systemie *gold standard* – stabilizowała oczekiwania uczestników rynku[63].

Twórcą reguły celu pośredniego jest Milton Friedman. Krytykując keynesowskie koncepcje wykorzystania zmian podaży pieniądza w polityce stabilizowania koniunktury, podkreślał, że dyskrecjonalne zmiany podaży pieniądza zwiększają niepewność, zaburzają sygnały rynkowe, a przede wszystkim w warunkach istnienia

[63] W przeciwieństwie do system waluty złotej, gdzie złoto miało wartość samoistną, w systemie pieniądza papierowego o wartości umownej potrzebna jest „zewnętrzna" kotwica nominalna podtrzymująca zaufanie do pieniądza papierowego. Może to być albo publicznie ogłoszone ograniczenie nałożone na ilość pieniądza wprowadzaną do obiegu, albo sztywne „przywiązanie" pieniądza papierowego do innego aktywu, na przykład do złota czy do waluty obcej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

racjonalnych oczekiwań sa nieskuteczne. Próby przyspieszenia wzrostu gospodarczego i obniżenia bezrobocia przy pomocy zwiększania nominalnej podaży pieniądza nie przynoszą trwałych rezultatów. Jedynym trwałym skutkiem stosowania takiej polityki jest rosnąca inflacja. Dlatego też Friedman postulował, aby banki centralne zwiększały podaż pieniądza według stałej, umiarkowanej stopy wzrostu, niezależnie od stanu gospodarki realnej. Wówczas inflacja pozostanie pod kontrolą (Friedman 1960).

Reguła Friedmana straciła na popularności w latach 80. i 90., gdy stabilna wcześniej relacja między ilością pieniądza a inflacją uległa silnemu rozregulowaniu. Lecz już w 1976 roku regułę tę podważyli dwaj znani ekonomiści amerykańscy Sargent i Wallace, którzy wykazali, że w warunkach istnienia oczekiwań inflacyjnych (racjonalnych lub autoregresyjnych) zwiększanie podaży pieniądza według stałej stopy wzrostu bez oglądania się na warunki gospodarowania jest polityką suboptymalną. Lepsze rozwiązanie stanowi, ich zdaniem, reguła uwzględniająca sprzężenia zwrotne między zmianami podaży pieniądza a zmianami produkcji. Optymalna reguła celu pośredniego w polityce pieniężnej według Sargenta i Wallace'a omówiona została w ramce 5.14.

Ramka 5.14. Optymalna reguła podaży pieniądza według Sargenta i Wallace'a (1976)*

Sargent i Wallace (1976) zakładają, że rząd chce kontrolować istotną zmienną makroekonomiczną (na przykład stopę bezrobocia lub odchylenia rzeczywistego PKB od potencjalnego PKB) przy pomocy polityki pieniężnej. Oznaczmy wartość tej zmiennej jako y_t. Wówczas kształtowanie się y_t opisuje następujące równanie autoregresyjne:

$$y_t = \alpha + \delta y_{t-1} + \beta m_t + u_t, \tag{5.37}$$

gdzie m_t jest stopą wzrostu podaży pieniądza w okresie t, a u_t składnikiem losowym z wariancją σ_t^2 i średnią równą zeru.

Załóżmy następnie, że rząd ustala stopę wzrostu podaży pieniądza w taki sposób, aby zminimalizować odchylenia zmiennej y_t od pewnego pożądanego poziomu y^*, lecz dostosowując jednocześnie stopę wzrostu podaży pieniądza do

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

rozwoju sytuacji gospodarczej. Wówczas zmiany podaży pieniądza można przedstawić jako następującą funkcję:

$$m_t = g_0 + g_1 y_{t-1}. \tag{5.38}$$

Równanie (5.38) nazwiemy regułą „sprzężenia zwrotnego" (*feedback rule*). Z kolei regułę Friedmana, zakładającą stałą stopę wzrostu podaży pieniądza, można przedstawić w następującej postaci:

$$m_t^F = m_{t-1}^F \tag{5.39}$$

dla wszystkich wartości t.

Podstawiając równanie (5.38) do (5.37) i przekształcając, otrzymujemy:

$$y_t = (\alpha + \beta g_0) + (\delta + \beta g_1) y_{t-1} + u_t. \tag{5.40}$$

Podstawiając w miejsce zmiennych y i u wartości oczekiwane w stanie ustalonym (czyli dla $E(y_t) = y^*$, i $E(u_t) = 0$) i przekształcając, uzyskujemy:

$$E(y_t) = \frac{(\alpha + \beta g_0)}{1 - (\delta + \beta g_1)} = y^*. \tag{5.41}$$

Przyrównując (5.41) do y^* zapewniamy minimalizację wariancji. Wariancja zmiennej y w stanie ustalonym, czyli wokół wartości y^*, jest równa:

$$var(y) = \frac{\sigma_u^2}{1 - (\delta + \beta g_1)^2}. \tag{5.42}$$

Rząd dobiera wartość g_1 w równaniu (5.38) w taki sposób, aby zminimalizować wariancję y, a następnie dobiera wartość g_0 z równania (5.41) w taki sposób, aby zapewnić $E(y_t) = y^*$. Wartość *var* (y), czyli równanie (5.42), osiąga minimum dla $(\delta + \beta g_1) = 0$, a stąd $g_1 = -\frac{\delta}{\beta}$. Wówczas na podstawie równania (5.41) otrzymujemy:

$$(\alpha + \beta g_0) = [1 - (\delta + \beta g_1)] y^*,$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

a stąd

$$g_0 = \frac{\left[1-\left(\delta+\beta g_1\right)\right]y^* - \alpha}{\beta}.$$

Podstawiając $g_1 = -\frac{\delta}{\beta}$ i przekształcając, uzyskujemy:

$$g_0 = \frac{y^* - \alpha}{\beta}.$$

Podstawiając wyliczone wartości g_0 i g_1 do reguły „sprzężenia zwrotnego" (5.38), otrzymujemy zmodyfikowaną postać reguły pieniężnej:

$$m_t = \left(y^* - \alpha\right) - \delta y_{t-1}. \tag{5.43}$$

Równanie (5.43) pokazuje, że zmiany podaży pieniądza nie są stałe w czasie, jak miało to miejsce w regule Friedmana, lecz zależą każdorazowo od zmian PKB w poprzednim okresie. Podstawiając (5.43) do (5.37) i przekształcając, otrzymujemy więc:

$$y_t = y^* + u_t.$$

Co oznacza, że jeśli rząd minimalizuje wariancję zmiennej y i stosuje regułę (5.43), to wartości y są zawsze równe wartości pożądanej y^*, plus minus ewentualne odchylenia losowe. W przeciwieństwie do tego, sztywna reguła Friedmana (5.39) prowadzi do powstawania stałych odchyleń y od y^*.

Alternatywą wobec celu pośredniego w postaci zmian ilości pieniądza w obiegu jest reguła celu pośredniego w postaci krótkookresowej stopy procentowej (zob. Szpunar 2000, rozdział 3). Polega ona na tym, że bank centralny najpierw szacuje związek między poziomem oficjalnej stopy procentowej, określającej cenę płynności w gospodarce, i zmianami ilości pieniądza rezerwowego oraz szerszych miar pieniądza. Następnie ustala stopę procentową na poziomie zapewniającym dopływ pieniądza do gospodarki w takim tempie, które jest spójne z celem inflacyjnym. Zaletą tej reguły pieniężnej jest to, że zmiany stóp procentowych banku centralnego przenoszą się szybko na zmiany rynkowych stóp procentowych, co pozwala kontrolować tempo zwiększania kredytu do gospodarki.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ogólnie rzecz biorąc, wybór celu pośredniego zależy od tego, co łatwiej jest zmierzyć i co łatwiej interpretować w kontekście związku z inflacją. Na przykład przy wysokiej i niestabilnej inflacji trudno szacować oczekiwaną inflację i lepiej stosować regulację ilościową (kontrola agregatu *M*), z kolei w warunkach niskiej inflacji lepiej stosować regulację cenową (kontrola stopy procentowej *r*).

Inaczej natomiast funkcjonuje reguła celu pośredniego w postaci kursu waluty. W tym przypadku bank centralny określa stały (sztywny) kurs walutowy i broni go, skupując i sprzedając waluty obce po tymże stałym kursie. Mówimy w tym przypadku, że bank centralny realizuje strategię celu kursowego (*exchange rate targeting, exchange rate anchoring*). W przeciwieństwie do dwóch reguł omówionych wyżej zadaniem reguły kursowej nie jest – bezpośrednie lub pośrednie – kontrolowanie tempa wzrostu podaży pieniądza. Wpływa ona natomiast na obniżanie oczekiwań inflacyjnych poprzez powiązanie waluty krajowej z wiarygodną i stabilną zagraniczną walutą odniesienia. Reguła ta dostosowuje inflację do inflacji w kraju odniesienia poprzez mechanizm arbitrażu cenowego, który zapewnia możliwości dokonywania zakupów importowych po niezmiennym kursie. Bezwarunkowe zobowiązanie banku centralnego do kupowania i sprzedawania każdej ilości walut obcych w zamian za walutę krajową prowadzi do ciągłych zmian poziomu oficjalnych rezerw walutowych, co z kolei wywołuje ciągłe zmiany ilości pieniądza krajowego w obiegu. Podaż pieniądza krajowego staje się w ten sposób zmienną endogeniczną.

Chcąc utrzymać pewien zakres autonomii w polityce pieniężnej, bank centralny może jednak do pewnego stopnia kontrolować podaż pieniądza krajowego albo nakładając ograniczenia na transgraniczny przepływ kapitału, albo przy pomocy tzw. operacji sterylizacyjnych, czyli podejmując działania neutralizujące pieniężne skutki zmiany rezerw walutowych. Operacje te polegają na tym, że równolegle do interwencji walutowej bazującej na zakupie lub sprzedaży walut obcych za walutę krajową, bank centralny dokonuje operacji otwartego rynku w postaci, odpowiednio, *reverse repo* lub *repo* na tę samą kwotę środków krajowych, tak aby ilość pieniądza krajowego w obiegu nie ulegała zmianie. Połączenie interwencji walutowej z operacjami otwartego rynku nazywamy interwencją sterylizowaną (szerzej o interwencjach walutowych piszemy w następnym rozdziale).

Gdy mamy do czynienia z napływem kapitału zagranicznego, bank centralny skupuje waluty obce, co wywołuje wzrost bazy monetarnej i wzrost podaży pieniądza krajowego. Jeśli bank centralny chce utrzymać podaż pieniądza krajowego na stałym poziomie (np. w obawie przed wzrostem inflacji), musi jednocześnie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ściągać z rynku nadwyżki pieniądza krajowego poprzez operacje rynku otwartego – na tym polega sterylizacja pieniężnych efektów interwencji walutowej.

Sterylizacja pozostaje jednak instrumentem niedoskonałym. Po pierwsze, jest kosztowna, ponieważ bank centralny musi zachęcać banki komercyjne do zwiększonego zakupu papierów wartościowych, podnosząc stopy procentowe. To z kolei oznacza wzrost kosztów operacji *reverse repo*. Przy sterylizacji na dużą skalę może dojść do pojawienia się strat w bilansie banku, które będą musiały zostać pokryte przez budżet państwa. Po drugie, sterylizacja może okazać się na dłuższą metę nieskuteczna. Wzrost stóp procentowych przyciąga bowiem dodatkowo kapitał zagraniczny i w ten sposób przyczynia się do zwiększenia skali operacji sterylizacyjnych. Relację wielkości tego dodatkowego przypływu kapitału do wielkości sterylizacji mierzy tzw. współczynnik offsetu (*offset coefficient*). Jeśli jest on większy od jedności, oznacza to, że w wyniku sterylizacji wzrost podaży pieniądza rezerwowego na skutek dodatkowego napływu kapitału okazał się większy niż spadek podaży pieniądza rezerwowego na skutek operacji sterylizacyjnych. W efekcie podaż pieniądza rezerwowego rośnie zamiast maleć i sterylizacja przestaje mieć sens (Rosati 1999). W takiej sytuacji bank centralny prędzej czy później jest zmuszony do rewaluacji kursu własnej waluty, aby zmniejszyć napływ kapitału. Autonomia banku centralnego w ustalaniu stóp procentowych i regulowania podaży pieniądza jest w tym modelu ograniczona – im większa staje się elastyczność kapitału względem zmian stopy procentowej i im bardziej będzie otwarty rachunek kapitałowy, tym mniejsza okaże się autonomia.

Jak wspomnieliśmy, generalnie lepiej jest prowadzić politykę pieniężną według ogłoszonych reguł niż w sposób dyskrecjonalny. Lecz polityka oparta na regułach również posiada swoje mankamenty. Najpoważniejszym z nich jest nieuchronna sztywność systemu podejmowania decyzji. Mechaniczne stosowanie reguł oznacza w praktyce brak możliwości szybkiego reagowania na nowe, nieprzewidziane wydarzenia, w tym zjawiska kryzysowe, utrudnia też dostosowanie reakcji banku centralnego do specyficznych warunków. Ponieważ życie jest nieporównanie bogatsze od nawet najlepszych modeli, przyjęte reguły nie są w stanie uwzględnić wszystkich możliwych scenariuszy wydarzeń. W efekcie ich rygorystyczne stosowanie ogranicza niezbędną elastyczność i swobodę polityki. Dlatego też często mieliśmy do czynienia z naruszaniem reguł lub z ich mniej lub bardziej uzasadnionymi zmianami, w obliczu nowych sytuacji. Przykładem są stałe kursy walutowe, które w momencie wprowadzania na rynek miały pozostać na zawsze niezmienne, a w praktyce ulegały częstym zmianom. W związku z tym, zdaniem niektórych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

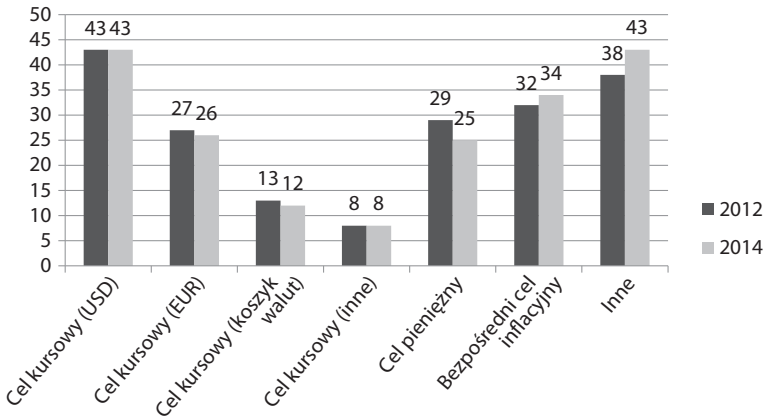
autorów, wszystkie stosowane w praktyce reguły są okazjonalnie naruszane, co oznacza, że polityka pieniężna jest zawsze w jakimś zakresie prowadzona w sposób dyskrecjonalny (Bernanke i in. 1999).

Warunkiem skuteczności polityki pieniężnej opartej na celach pośrednich jest stabilność relacji między zmienną opisującą cel pośredni (podaż pieniądza, stopy procentowe, kurs walutowy) a inflacją. Gdy relacja ta jest stabilna, kontrola celu pośredniego pozwala również kontrolować inflację. Lecz w ostatnich dziesięcioleciach doszło do zaburzeń tych zależności. Przyczynami naruszenia stabilnego związku między zmiennymi celami pośrednimi a inflacją była rosnąca skala przepływów kapitału w skali międzynarodowej, spowodowana rozwojem technologii informatycznych i telekomunikacyjnych, oraz usuwanie barier ograniczających swobodną wymienialność walut. Dlatego też coraz więcej państw (między innymi Polska) odeszło od wykorzystywania reguły celów pośrednich w polityce pieniężnej i zdecydowało się na stosowanie strategii bezpośredniego celu inflacyjnego (*direct inflation targeting*). Strategia ta zostanie szczegółowo omówiona w punkcie 5.7.

5.6.4. Reguły polityki pieniężnej stosowane na świecie

Międzynarodowy Fundusz Walutowy gromadzi dane na temat reguł polityki pieniężnej stosowanych przez poszczególne kraje. Wykres 5.15 prezentuje rozkład głównych strategii pieniężnych z punktu widzenia przyjętych celów pośrednich lub ich braku.

Wykres 5.15. Reguły polityki pieniężnej na świecie a liczba krajów w roku 2012 i 2014



Źródło: IMF 2012, 2014.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Jak widać, na 191 krajów objętych badaniem w 2014 roku prawie połowa (89) stosowała różne odmiany reguły kursu walutowego, przy czym najczęściej kurs waluty krajowej był ustalany w stosunku do dolara amerykańskiego. Trzydzieści cztery kraje (w tym większość krajów rozwiniętych) stosowały strategię bezpośredniego celu inflacyjnego (BCI), z kolei pozostałe kraje zdecydowały się albo na tradycyjne reguły pośredniego celu pieniężnego (*monetary intermediate target*) albo na strategie mieszane, z pewnym zakresem dyskrecjonalności. W porównaniu z 2012 rokiem, nastąpił niewielki spadek liczby krajów stosujących regułę kursu walutowego i pośredniego celu pieniężnego, a jednocześnie wzrosła liczba krajów stosujących strategie BCI i strategie mieszane.

5.7. Strategia bezpośredniego celu inflacyjnego

5.7.1. Istota i elementy składowe strategii BCI

Jak wyżej wspomniano, reguły celów pośrednich są mało skuteczne, gdy parametry funkcji opisującej związek celu pośredniego z celem ostatecznym (inflacja) są niestabilne. Ma to miejsce wówczas, gdy relacja między inflacją a podażą pieniądza ulega zaburzeniom, czyli gdy zmieniają się wartości mnożników pieniężnych lub wartości parametrów funkcji popytu na pieniądz. Zjawisko to związane jest ze zmianami szybkości obiegu pieniężnego (*velocity instability problem* – Bernanke i Mishkin 1997). Przyczyną tych zmian mogą być innowacje technologiczne w dziedzinie finansów (np. rozwój elektronicznych form pieniądza) lub szoki zewnętrzne i zjawiska kryzysowe. Lecz bardziej fundamentalną przyczyną osłabienia skuteczności reguły pieniężnej opartej na celu pośrednim w postaci podaży pieniądza jest stopniowa utrata kontroli nad podażą pieniądza krajowego przez banki centralne w wielu państwach. Jest to spowodowane liberalizacją przepływów kapitału w skali międzynarodowej w okresie ostatniego ćwierćwiecza. Dotyczy to głównie kapitału finansowego, który przepływa z kraju do kraju w poszukiwaniu najkorzystniejszych stóp przychodu. Wzrost stóp procentowych w danym kraju wywołuje napływ kapitału z zagranicy (przy danym poziomie ryzyka), co przekłada się na wzrost podaży pieniądza krajowego. Z kolei odpływ kapitału ogranicza podaż pieniądza krajowego. W efekcie mamy do czynienia z egzogenicznymi zmianami wielkości obiegu pieniężnego, często wbrew woli banku centralnego. Transgraniczna mobilność kapitału ogranicza więc autonomię krajowej polityki pieniężnej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W latach 90. komplikacje te stały się na tyle poważne, że niektóre banki centralne zdecydowały się na zmianę sposobu prowadzenia polityki pieniężnej, a w szczególności na rezygnację z celu pośredniego, jakim jest podaż pieniądza. Wobec niechęci (zresztą uzasadnionej, jak przekonamy się o tym w następnym rozdziale) do powrotu do reguły kursu walutowego powstała potrzeba wypracowania nowych zasad. Wśród nich wielką popularność zyskała tzw. reguła bezpośredniego celu inflacyjnego (BCI), wprowadzona w życie po raz pierwszy przez bank centralny Nowej Zelandii na początku lat 90. Od tamtej pory wiele banków centralnych w różnych krajach zdecydowało się porzucić tradycyjne reguły polityki pieniężnej i zaczęło stosować strategię BCI. Również Narodowy Bank Polski od 1999 roku prowadzi politykę pieniężną na podstawie tej reguły.

Reguła – albo strategia – BCI polega na tym, że bank centralny ustala pożądany poziom inflacji w średnim okresie (czyli ustala tzw. cel inflacyjny – *inflation target*), a następnie prognozuje inflację i dostosowuje politykę pieniężną do różnicy między celem inflacyjnym a przewidywaną inflacją. Innymi słowy, bank centralny zaostrza politykę pieniężną, gdy z prognozy wynika, że inflacja będzie wyższa od przyjętego celu, i luzuje politykę pieniężną, kiedy prognoza wskazuje, że inflacja będzie niższa od celu. W praktyce oznacza to, że w pierwszym przypadku bank centralny podnosi oficjalne stopy procentowe, a w drugim je obniża.

W rzeczywistości strategia BCI jest czymś znacznie szerszym i bardziej kompleksowym niż tylko wąsko zdefiniowaną regułą prowadzenia polityki pieniężnej (Szpunar 2000, s. 183–196). Stanowi w istocie pewien całościowy system, w ramach którego bank centralny prowadzi politykę pieniężną w taki sposób, aby utrzymać stabilność cen w różnych warunkach. Strategia BCI nie jest tak sztywna czy wręcz mechaniczna jak reguła celu pośredniego, ale nie jest też dowolna jak pełna dyskrecjonalność. Może ona przybierać różne formy, uwzględniające specyfikę poszczególnych krajów. Jest to więc system zapewniający niezbędną elastyczność, pozwalający bankom centralnym reagować na nieoczekiwane wydarzenia, zachowując jednocześnie jasny cel działania (Bernanke i in. 1999, Wojtyna 2012).

Rezygnacja z celu pośredniego w postaci podaży pieniądza oznaczała daleko idącą zmianę w sposobie prowadzenia polityki pieniężnej. Oznaczała w praktyce zerwanie z zasadą Friedmana, zgodnie z którą banki centralne powinny starać się utrzymywać stabilne, umiarkowane tempo wzrostu podaży pieniądza w długim okresie. Parametry polityki pieniężnej miały teraz dostosowywać się do przewidywanego poziomu inflacji. Ta orientacja „ku przyszłości" (*forward-looking policy*) jest najbardziej charakterystyczną cechą nowej koncepcji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Reguła oparta na praktycznie nieobserwowalnej zmiennej, jaką jest prognozowana inflacja, niosła jednak ryzyko osłabienia wiarygodności polityki pieniężnej. Uczestnicy rynku mogli obawiać się, że prognozy będą sporządzane w sposób mało wiarygodny, a niepewna przyszłość stanie źródłem częstych błędów w polityce pieniężnej. Co więcej, rezygnacja z nominalnej kotwicy, jaką stanowił parytet złota w systemie waluty złotej czy ustalone tempo wzrostu podaży pieniądza w systemie pieniądza papierowego, niosły ryzyko wzrostu oczekiwań inflacyjnych i dewaluacyjnych. Dlatego rozwijając koncepcję BCI, szczególną uwagę poświęcono tym jej elementom, które wzmacniały jej wiarygodność i stabilizowały oczekiwania uczestników rynku.

Przyjmuje się, że w operacyjnym zastosowaniu strategii BCI banki centralne muszą rozstrzygnąć następujące kwestie (zob. m.in. Bernanke i Mishkin 1997; Bernanke i in. 1999, s. 27–38):

a) wybór miernika inflacji – czy ma nim być wskaźnik CPI, PPI, HICP, wskaźnik inflacji bazowej (*core inflation*), czy też jakaś inna miara zmian cen;
b) ustalenie celu inflacyjnego w perspektywie średniookresowej, co jest równoznaczne z przyjęciem określonej definicji „stabilności cen". Wyznaczenie celu może przyjąć postać wartości punktowej, przedziału wartości bądź wartości punktowej z określonymi limitami odchyleń (na przykład w Polsce 2,5% +/–1 punkt procentowy, a w Unii Gospodarczo-Walutowej „poniżej 2%, ale blisko 2%");
c) przyjęcie systemu systematycznych (powtarzalnych), niezależnych i wiarygodnych prognoz inflacji, oraz przyjęcie zasady „patrzenia do przodu" (*forward-looking*) zamiast „patrzenia do tyłu" (*backward-looking*);
d) przyjęcie i ogłoszenie „funkcji reakcji", rozumianej jako sposób powiązania decyzji polityki pieniężnej z wynikami prognoz, uwzględniający jednocześnie dopuszczalne opóźnienia w reakcji, możliwy okres odchylania się inflacji od celu oraz sposób reagowania na jednorazowe szoki podażowe i zewnętrzne (*external shocks*);
e) strategia komunikowania się z rynkami (wyjaśnianie decyzji, udostępnianie stosowanych modeli, wyjaśnianie błędów i odchyleń).

W dalszej części tego podrozdziału omówimy najważniejsze elementy składowe kompleksowego systemu BCI.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.7.2. Wybór miernika inflacji

Przyjęcie właściwego miernika inflacji to głównie dokonanie wyboru pomiędzy CPI, który jest wskaźnikiem powszechnie znanym i najszerzej wykorzystywanym (*headline inflation*), zrozumiałym dla szerokiej publiczności i najbardziej przejrzystym, choć mniej podatnym na politykę pieniężną, a wskaźnikami inflacji bazowej (*core inflation*), które lepiej reagują na politykę pieniężną, lecz są mniej znane i mniej „przejrzyste" w odbiorze publicznym.

Wybrany miernik powinien być dostatecznie reprezentatywny i obejmować dostatecznie szeroki „koszyk" dóbr i usług (*broad-based*), musi być też dokładny i szybko podawany do wiadomości. Konieczne jest wyjaśnienie konstrukcji wskaźnika oraz informowanie o zmianach w sposobie liczenia, tak aby uczestnicy rynku – jeśli zechcą – byli w stanie zweryfikować poprawność obliczeń. Wreszcie wskaźnik inflacji powinien być liczony przez odrębną instytucję niezależną od rządu i od banku centralnego, aby uniknąć obaw o manipulacje i zapewnić wysoką wiarygodność polityki pieniężnej. Dlatego też wskaźniki cen w większości krajów są obliczane przez niezależne urzędy statystyczne.

Warunki reprezentatywności i przejrzystości spełnia wskaźnik CPI, choć zarazem jest on podatny na zmiany cen wywołane czynnikami egzogenicznymi. Jego odmianą w Unii Europejskiej jest tzw. wskaźnik zharmonizowany HICP (*Harmonized Index of Consumer Prices*), który obliczany jest przez Europejski Urząd Statystyczny (Eurostat) na podstawie ujednoliconej struktury konsumpcji w odniesieniu do wszystkich państw członkowskich. Dzięki harmonizacji uzyskuje się porównywalność wskaźników między poszczególnymi krajami. Niektóre z nich stosują w ramach strategii BCI wskaźniki inflacji bazowej, których zadaniem jest wyłączenie z zakresu miernika inflacji jednorazowych skoków cenowych oraz ruchów cen wywołanych odziaływaniem czynników zewnętrznych i niezależnych od polityki pieniężnej (zob. ramka 5.15).

Ramka 5.15. Wskaźniki inflacji bazowej w Polsce

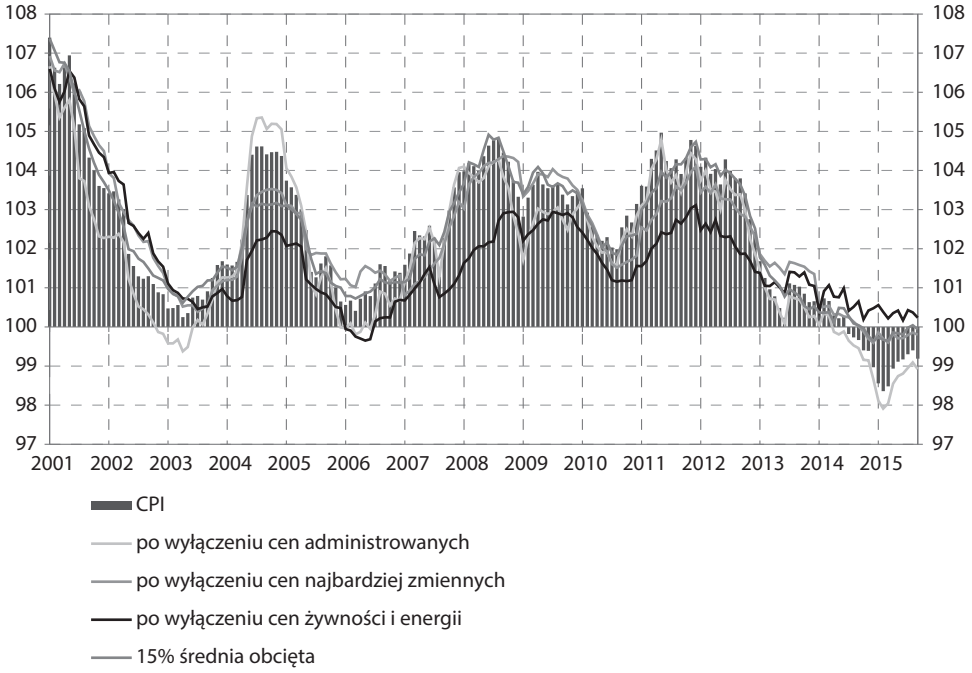
Wskaźniki inflacji bazowej (*core inflation*) mają na celu „oczyszczanie" wskaźnika CPI z wpływu czynników niezależnych od polityki pieniężnej (takich jak zewnętrzne szoki podażowe, zmiany cen administrowanych, zmiany stawek

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

podatków pośrednich wpływających na poziom cen). W Polsce, poczynając od marca 2009 roku, Narodowy Bank Polski oblicza systematycznie cztery wskaźniki inflacji bazowej:

a) inflacja bazowa po wyłączeniu cen administrowanych,
b) inflacja bazowa po wyłączeniu cen najbardziej zmiennych,
c) inflacja po wyłączeniu cen żywności i energii – tzw. inflacja „netto",
d) 15% średnia obcięta (symetrycznie obcinająca po 15% z lewej i prawej strony rozkładu wskaźniki cen o największej i najmniejszej dynamice w porównaniu z poprzednim okresem).

Wykres 5.16. Wskaźniki CPI oraz wskaźnik inflacji bazowej w latach 2001–2015 (analogiczny miesiąc poprzedniego roku = 100)



Źródło: NBP.

Wśród analityków rynkowych najbardziej popularną miarą inflacji bazowej jest wskaźnik inflacji „netto". Wyłączenie cen żywności w konstrukcji tego wskaźnika wynika stąd, że ich zmiany w przeważającej mierze zależą od zmian warunków produkcji i podaży (warunki pogodowe wpływające na poziom urodzaju), podczas gdy popyt na żywność jest względnie nieelastyczny. Z kolei wyłączenie cen energii

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jest uzasadnione tym, że są one kształtowane przez zmiany cen nośników energii na rynkach światowych (ropa naftowa, gaz ziemny), a krajowa polityka pieniężna nie ma na nie wpływu.

Wykres 5.16 przedstawia kształtowanie się wskaźników inflacji bazowej w Polsce w latach 2001–2015 na tle zmian wskaźnika CPI. Można zauważyć, że spora część inflacji w Polsce była wywołana niekorzystnymi zmianami cen żywności i paliw, co oznacza, że przez większą część analizowanego okresu wskaźnik CPI generalnie zawyżał rzeczywistą inflację. Jedynie w latach 2002–2003 oraz 2013–2015 stabilne (lub spadające) ceny paliw i żywności wpływały na spadek inflacji, a wskaźnik inflacji „netto" był wyższy niż wskaźnik CPI.

5.7.3. Wybór celu inflacyjnego

Wybór celu inflacyjnego, czyli określonej liczbowej wartości wybranego miernika inflacji, powinien być spójny z postulatem zapewnienia stabilności cen. Wyjaśniliśmy już wcześniej, dlaczego wybór celu inflacyjnego na poziomie zero lub bliskim zeru nie znajduje ekonomicznego uzasadnienia. Banki centralne wybierają więc jako cel inflacyjny inflację na poziomie większym od zera, lecz na tyle niskim, aby nie powodował on konfliktu z dążeniem do szeroko pojmowanej stabilności cen. Ogólnie przyjmuje się, że warunki te spełnia inflacja na poziomie między 1% a 3% rocznie (Bernanke i in. 1999, s. 30).

Wybierając cel inflacyjny, trzeba także zdecydować, czy ma on przyjąć postać określonej liczby, czyli wartość punktową, czy raczej formę określonego przedziału wartości. Zaletą celu punktowego jest bardzo jasne i precyzyjne wskazanie pożądanego poziomu inflacji, co pozwala zakotwiczyć oczekiwania inflacyjne i zmniejsza niepewność co do przyszłych działań banku centralnego. Z kolei wadą pozostaje praktyczna trudność z wykonaniem tak ustalonego celu, ponieważ, formalnie rzecz biorąc, prawdopodobieństwo „trafienia" w wartość punktową jest nieskończenie małe. W efekcie niezdolność do precyzyjnego „trafiania w cel" mogłaby podważać wiarygodność polityki pieniężnej. Dlatego też część banków centralnych ustala cel w postaci przedziału dopuszczalnych wartości. Mankamentem tego rozwiązania jest jednak to, że uczestnicy rynku nie bardzo wiedzą, jaki konkretnie poziom inflacji bank centralny uznaje za najbardziej pożądany. W szczególności, gdy przyjęty przedział okazuje się stosunkowo szeroki, prognozy rynkowe dotyczące przyszłej inflacji są nieuchronnie obciążone większym błędem. Uczestnicy rynku nie wiedzą,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

czy w oczach banku centralnego inflacja bliska dolnej granicy przedziału jest równie pożądana jak inflacja w środku przedziału lub inflacja bliska górnej granicy.

Możliwym sposobem rozwiązania tej trudności jest ustalenie celu punktowego wraz z przedziałem dopuszczalnych odchyleń. Osiąga się wówczas zarówno stabilizację oczekiwań, jak i dopuszcza się niewielkie odchylenia inflacji od celu. Nie są one wówczas traktowane przez uczestników rynku jako błąd w realizacji polityki. Takie rozwiązanie przyjęła Rada Polityki Pieniężnej w Polsce, gdzie od 2004 roku cel inflacyjny jest określony jako inflacja CPI na poziomie 2,5%, z dopuszczalnym przedziałem odchyleń plus/minus jeden punkt procentowy. Również Europejski Bank Centralny, ustalając cel inflacyjny na poziomie „poniżej 2%, ale blisko 2%", dopuszcza pewną elastyczność w realizacji celu inflacyjnego, choć jej zakres jest znacznie mniejszy niż w Polsce.

Cel inflacyjny może być też określony w odniesieniu do poziomu cen, a nie do tempa wzrostu cen. Tak zwany „bezpośredni cel cenowy" (*price-level targeting*) nie musi oznaczać utrzymywania stałego poziomu cen – byłoby to równoznaczne z przyjęciem celu zerowej inflacji, co jest niewskazane z powodów podanych wyżej. Cel w postaci określonego poziomu cen może być podnoszony w czasie. Zaletą takiego rozwiązania jest to, że pozwala ono uniknąć większych zmian cen w dłuższym okresie. Bezpośredni cel inflacyjny nie kompensuje mianowicie szoków cenowych, których występowanie podnosi poziom cen w sposób trwały. Zastosowanie w tej sytuacji bezpośredniego celu cenowego pozwala korygować te odchylenia, a w długim okresie umożliwia większą stabilizację cen.

Wprawdzie przyjęty cel inflacyjny obowiązuje z reguły w długim okresie czasu (chyba że mamy do czynienia z hamowaniem wysokiej inflacji), lecz banki centralne zdają sobie sprawę, że ich oddziaływanie na gospodarkę i ceny obejmuje zasadniczo krótki i średni okres. W praktyce cele inflacyjne ustalane na okresy krótsze niż rok są trudne lub wręcz niemożliwie do zrealizowania ze względu na nieuniknione opóźnienia w mechanizmie transmisji (zob. punkt 5.5.4). Z kolei w okresie dłuższym niż 3–4 lata oddziaływanie impulsów polityki pieniężnej na zmienne realne stopniowo wygasa; co więcej, trudno jest wpływać na oczekiwania w tak długim horyzoncie czasowym. Dlatego też banki centralne koncentrują się zasadniczo na horyzoncie średniookresowym – starają się przewidywać rozwój inflacji w perspektywie 2–3 lat – i dostosowują do tych prognoz swoje decyzje z zakresu polityki pieniężnej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.7.4. System prognoz inflacji

Właśnie dlatego, że to przewidywana – a nie aktualna – inflacja jest podstawą podejmowania bieżących decyzji, kluczowy składnik reguły BCI stanowi system wiarygodnych, niezależnych i systematycznie powtarzalnych prognoz. Oczywiście, wszystkie szacunki dotyczące nieznanej przyszłości są z natury rzeczy obarczone niepewnością, co oznacza, że zawsze istnieje prawdopodobieństwo popełnienia błędu w decyzjach podejmowanych na ich podstawie. W szczególności zaś złe prognozy prowadzą do podejmowania złych decyzji. Tej ich wady nie da się całkowicie wyeliminować. Można jednak starać się, aby prognozy były maksymalnie wiarygodne i obciążone możliwie najmniejszym błędem. Ich trafność zależy od kompetencji autorów, stosowanej metodologii oraz jakości, dostępności i zakresu danych wykorzystywanych w prognozowaniu. Dlatego też sporządzaniem prognoz służących realizacji celów polityki pieniężnej zajmują się zespoły specjalistów, stosujących nowoczesne techniki i metody prognostyczne. Często stosowane jest podejście dwuetapowe, obejmujące najpierw przygotowanie prognozy inflacji na podstawie ekonometrycznych modeli makroekonomicznych, a następnie weryfikowanie otrzymanych wyników metodą ocen ekspertów. Procedura taka pozwala na uwzględnienie nowych czynników i zmian w dotychczasowych tendencji.

Prognozy powinny być sporządzane przez instytucje niezależne od władz politycznych i od banku centralnego. Chodzi o uniknięcie pokusy manipulowania prognozami w celu osiągnięcia doraźnych celów politycznych i zapewnienie im wysokiej wiarygodności. Idealnym rozwiązaniem byłoby powierzenie tego zadania niezależnemu ośrodkowi naukowo-badawczemu. W wielu krajach brakuje jednak takich instytucji, które dysponowałyby zarówno dostępem do niezbędnych danych, jak i profesjonalnym zespołem analityków, zapewniających wiarygodność i bezstronność sporządzanych prognoz. Dlatego często zadanie to wykonuje sam bank centralny, który z reguły dysponuje nie tylko dostępem do danych, lecz także odpowiednimi zasobami ludzkimi i finansowymi. Aby jednak w takiej sytuacji zapewnić wysoką wiarygodność ogłaszanych prognoz i uniknąć podejrzeń o manipulację danymi, bank centralny powinien udostępnić uczestnikom rynku zastosowaną metodologię i dane statystyczne wykorzystywane w procesie przygotowania prognoz. Analitycy rynkowi mogą wówczas weryfikować poprawność prognoz banku centralnego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wreszcie prognozy powinny być powtarzalne, czyli sporządzane w sposób systematyczny, zarówno z wykorzystaniem nowych danych napływających z gospodarki, jak i korektą dotychczas popełnionych błędów. Prognozy takie mają charakter kroczący, co oznacza, że, choć ich horyzont jest stały – na przykład 3 lata, to w miarę upływu czasu systematycznie obejmuje kolejne okresy. W praktyce banki centralne sporządzają takie prognozy co kilka miesięcy, najczęściej 3–4 razy w ciągu roku. Taka częstotliwość pozwala na uwzględnianie nowych tendencji i jednoczesne odsiewanie zmian przypadkowych i jednorazowych.

Dostosowanie decyzji polityki pieniężnej do różnicy między przewidywaną inflacją a celem inflacyjnym stanowi fundamentalną zmianę w prowadzeniu polityki pieniężnej. W ramach tradycyjnej reguły celu pośredniego bank centralny regulował podaż pieniądza w obiegu poprzez dostosowanie oficjalnych stóp procentowych (i ewentualnie skali operacji otwartego rynku) do pewnych wybranych wskaźników, takich jak łączna wielkość obiegu pieniężnego (M2, M3) czy poziom krótkookresowych rynkowych stóp procentowych. Ten sposób prowadzenia polityki pieniężnej oparty jest na zasadzie „patrzenia wstecz" (*backward looking*), ponieważ podstawę decyzji bieżących stanowią zawsze dane z przeszłości, choć nieodległej. Można to przyrównać do jazdy samochodem ze wzrokiem utkwionym w lusterko wsteczne – korekty kierunku jazdy są dokonywane dopiero wówczas, gdy widzimy, że samochód już zjechał z drogi. Uwzględnianie nieuchronnych opóźnień w transmisji impulsów pieniężnych do gospodarki sprawia, że taki system polityki pieniężnej ma charakter reaktywny – decyzje banku centralnego są z definicji zawsze spóźnione. W przeciwieństwie do tego system BCI jest *expressis verbis* oparty na zasadzie „patrzenia do przodu" (*forward looking*), ponieważ podstawą podejmowania decyzji są odchylenia *przewidywanej* inflacji od przyjętego celu.

5.7.5. Funkcja reakcji banku centralnego. Reguła Taylora

Gdy przyjęty system prognoz zapewnia ich wiarygodność, niezależność i powtarzalność, zasada „patrzenia do przodu" wymaga dostosowywania decyzji polityki pieniężnej do odchyleń przewidywanej inflacji od ustalonego celu inflacyjnego. W tym zakresie banki centralne mogą określić i stosować pewne reguły reagowania na tak zidentyfikowane odchylenia. Tak zwana „funkcja reakcji" (*reaction function*) banku centralnego informuje uczestników rynku o tym, jak zachowa się on w sytuacji, gdy przewidywana inflacja zacznie odchylać się od celu. Z punktu widzenia rynków byłoby pożądane, gdyby bank centralny stosował się rygorystycznie do konkretnej,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

publicznie ogłoszonej funkcji reakcji. Na przykład można sobie wyobrazić funkcję reakcji, która zakłada, że jeśli przewidywana inflacja odchyla się od celu inflacyjnego w górę o 1 punkt procentowy, bank centralny podnosi stopy procentowe, powiedzmy, o 0,5 punktu procentowego. Tego rodzaju rygorystyczne zobowiązanie, choć niewątpliwie zmniejsza niepewność co do przyszłości, nie jest jednak wskazane. Po pierwsze, prognozy same w sobie są obarczone niepewnością. Po drugie – i co ważniejsze, polityka pieniężna nie może być prowadzona na podstawie jednej mechanicznie funkcjonującej reguły, ponieważ rzeczywistość gospodarcza jest nieskończenie złożona i nawet najbardziej rozbudowane modele prognostyczne nie są w stanie uwzględnić wszystkich zależności. Bank centralny powinien więc zawsze zachowywać pewną swobodę działań w zakresie polityki pieniężnej, aby właściwie reagować na nieoczekiwane czynniki jednorazowe i systematyczne, nowe tendencje oraz zmienne niemierzalne. W polityce potrzebna jest zawsze pewna elastyczność, a wiązanie sobie rąk jedną mechaniczną regułą byłoby bardzo ryzykowne.

Dlatego też funkcje reakcji, nawet jeśli są ogłaszane w postaci określonej zależności (co zdarza się jednak rzadko), służą raczej jako wskazówki kierunkowe polityki banku centralnego. Określają ogólne zasady reagowania banku na pojawiające się – *przewidywane* – odchylenia inflacji od celu. Bank centralny ogłasza jedynie, że jeśli przewidywana inflacja stanie się w horyzoncie średniookresowym wyższa od celu, będzie zaostrzał politykę pieniężną, jeśli przewidywana inflacja okaże się natomiast niższa od celu, będzie luzował politykę pieniężną. Bank jednocześnie unika konkretnych zobowiązań, zwłaszcza tych dotyczących określonej skali podwyżek (lub obniżek) stóp i ich związku ze skalą odchyleń, a także momentu ich dokonywania, pozostawiając sobie pewien zakres dyskrecjonalności. Aby uniknąć jednak niepożądanego wzrostu niepewności i spadku wiarygodności polityki, bank centralny powinien w miarę możliwości jasno określić, kiedy i w jakich okolicznościach należy spodziewać się zmian w polityce pieniężnej.

W bardziej rozbudowanej postaci funkcja reakcji banku centralnego danego kraju może uwzględniać szereg najważniejszych czynników mających wpływ na poziom oficjalnych stóp procentowych. Formalnie można ją przedstawić w następującej postaci ogólnej:

$$i = i^* + \left(r - r^*\right) + \left(p_e - p_e^*\right) + \alpha\left(y - \hat{y}\right) + \beta\left(p_e - \pi\right), \tag{5.44}$$

gdzie: i to oficjalna stopa procentowa banku centralnego, r oznacza realną stopę przychodu z kapitału w kraju, p_e to oczekiwany poziom inflacji w kraju, $\left(y - \hat{y}\right)$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

to różnica między rzeczywistym a potencjalnym tempem wzrostu PKB, czyli tzw. luka podażowa (*output gap*), $(p_e - \pi)$ oznacza różnicę między prognozowaną inflacją a celem inflacyjnym, znak (*) symbolizuje odpowiednie zmienne w przypadku „zagranicy". Pierwsze trzy elementy w formule (5.44) mierzą wpływ czynników egzogenicznych i strukturalnych (i^*, r-r^*, p_e^*), niezależnych w krótkim okresie od polityki pieniężnej, na stopę procentową banku centralnego. Zagraniczna stopa procentowa i^* jest punktem odniesienia dla stóp krajowych w warunkach swobody przepływu kapitału i umożliwia arbitraż. Wyższa (niższa) realna stopa przychodu z kapitału w kraju niż za granicą uzasadnia wyższą (niższą) krajową stopę procentową przy danych oszczędnościach krajowych. Podobnie też wyższa (niższa) stopa inflacji w kraju niż za granicą wywołuje wzrost (spadek) krajowych stóp procentowych. Dwa ostatnie składniki równania (5.37) odnoszą się do polityki pieniężnej. Gdy rzeczywiste tempo wzrostu krajowego PKB – y – przekracza potencjalne tempo wzrostu $\hat{y}$, bank centralny powinien podnieść stopy procentowe, aby przywrócić równowagę w gospodarce. Parametr $\alpha > 0$ mierzy stopień awersji banku centralnego wobec odchyleń rzeczywistego tempa wzrostu PKB od tempa potencjalnego. Im bardziej bank centralny jest przekonany o tym, że szybki wzrost gospodarczy spowoduje przyspieszenie inflacji, tym wyższa będzie wartość współczynnika α. Podobnie, gdy przewidywana inflacja p_e jest wyższa niż cel inflacyjny π, bank centralny powinien podnieść stopy procentowe, aby obniżyć inflację do pożądanego poziomu. Parametr $\beta > 0$ mierzy stopień awersji banku centralnego wobec odchylenia inflacji od celu. W im mniejszym stopniu bank centralny jest skłonny tolerować odchylenie inflacji od celu, tym wyższa będzie wartość współczynnika β (zob. ramka 5.16).

W skrajnym przypadku, dla $\alpha = 0$, i przy pozostałych składnikach strukturalnych stałych równanie (5.37) redukuje się do najprostszej postaci reguły BCI:

$$i = \Delta + \left(p_e - p_e^*\right) + \beta\left(p_e - \pi\right). \tag{5.45}$$

Ramka 5.16. Reguła Taylora

Amerykański ekonomista John Taylor wykazał, że zmiany oficjalnych stóp procentowych można modelować jako reakcję banku centralnego na odchylenia rzeczywistej inflacji od celu inflacyjnego i rzeczywistego tempa wzrostu gospodarczego

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

od tempa potencjalnego[64]. Budując funkcję reakcji banku centralnego, Taylor brał przede wszystkim pod uwagę realne warunki gospodarki USA, w przypadku której (podobnie jak wśród innych dużych gospodarek o dominującej roli rynku wewnętrznego) wartości zmiennych zagranicznych są mniej istotne w określaniu krajowej stopy procentowej. Wówczas równanie (5.44) można zapisać w postaci tzw. reguły Taylora:

$$i = r^n + p_e + \alpha\left(y - \hat{y}\right) + \beta\left(p_e - \pi\right), \tag{5.46}$$

gdzie r^n jest „neutralną" (*neutral*) realną stopą procentową równą potencjalnej realnej stopie wzrostu PKB[65]. Równanie (5.46) pokazuje funkcję reakcji banku centralnego, który stara się prowadzić politykę pieniężną w taki sposób, aby realizować jednocześnie dwa cele makroekonomiczne – utrzymanie tempa wzrostu PKB na poziomie potencjalnym i utrzymanie inflacji na poziomie celu inflacyjnego. Stosując regułę Taylora w praktyce, bank centralny ustala stopy procentowe w reakcji na odchylenia rzeczywistego tempa wzrostu PKB – y – od potencjalnego tempa wzrostu PKB – $\hat{y}$ – oraz na odchylenia prognozowanej inflacji p_e od celu inflacyjnego π. Wzajemna relacja parametrów α/β ukazuje relatywną wagę obu celów gospodarczych – tempa wzrostu i stabilności cen – w polityce pieniężnej. Im niższa relacja α/β, tym ważniejszy w polityce pieniężnej jest cel w postaci stabilności cen. Gdy tempo wzrostu PKB jest równe potencjalnemu i gdy inflacja pozostaje równa celowi, stopa procentowa banku centralnego powinna być utrzymywana na poziomie $i = r^n + \pi$.

Zgodnie z regułą Taylora, stopy procentowe powinny być relatywnie wysokie, gdy inflacja znajduje się powyżej celu i/lub tempo wzrostu PKB kształtuje się powyżej potencjalnego tempa wzrostu. Również odwrotnie, stopy procentowe powinny być relatywnie niskie, gdy inflacja znajduje się poniżej celu i/lub wzrost gospodarczy jest niższy od „potencjału". Reguła ta jest szczególnie przydatna w ocenie efektywności polityki pieniężnej w USA, gdzie Fed realizuje kilka celów jednocześnie, w tym utrzymanie wysokiego poziomu zatrudnienia (co jest funkcją tempa wzrostu PKB) i utrzymanie stabilności cen. Porównanie wartości „teoretycznych" stóp procentowych wynikających z reguły Taylora (zakładając, że $r^n = 2{,}0$, oraz $\alpha, \beta = 0{,}5$)

[64] Taylor (1993a), (1993b).

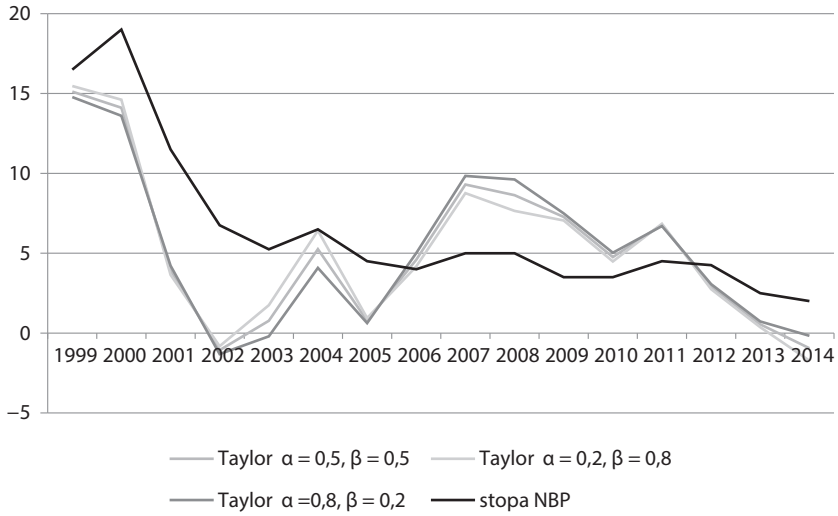
[65] „Neutralna" stopa procentowa (nie należy jej mylić ze stopą „naturalną") może być zdefiniowana jako stopa procentowa równa stopie wzrostu PKB, która maksymalizuje konsumpcję *per capita* w długim okresie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i rzeczywistych wartości stóp funduszy federalnych (*federal funds rate*) w latach 1987–1992 wykazało dużą zbieżność obu zmiennych.

Na wykresie 5.17 przedstawiono wartości hipotetycznych stóp procentowych wynikające z reguły Taylora dla Polski oraz wartości stóp procentowych NBP w okresie 1999–2014. Stopy „Taylora" zostały wyliczone na podstawie równania (5.44). Obszarem odniesienia jest strefa euro, przy czym dane dotyczące inflacji w Polsce i w strefie euro nawiązują do bieżących (a nie prognozowanych) odczytów z grudnia każdego roku (rok do roku), różnicę ($r - r^*$) ustalono na poziomie 2 punktów procentowych, a wielkości luki podażowej w przypadku Polski zaczerpnięto z bazy danych AMECO.

Wykres 5.17. Oficjalne stopy procentowe NBP i stopy wyliczone według reguły Taylora w okresie 1999–2014



Źródło: opracowanie własne.

Z wykresu wynika, że w analizowanym okresie stopy „Taylora" w przypadku różnych wartości parametrów α i β praktycznie nie różniły się od siebie. Z kolei stopy procentowe NBP zdecydowanie odbiegały od wartości zalecanych regułą Taylora – były wyraźnie wyższe w latach silnej dezinflacji 1999–2005 oraz wyraźnie niższe w latach 2006–2011. Te odchylenia mogą sugerować, że polityka pieniężna NBP była zbyt restrykcyjna w pierwszym okresie, a zbyt ekspansywna w drugim. O poprawności tej oceny może świadczyć silne spowolnienie polskiej gospodarki w latach 2000–2003 oraz spekulacyjny wzrost cen aktywów w okresie 2006–2008.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W ostatnich latach (2012–2014) stopy NBP ponownie wzrosły ponad poziom wynikający z reguły Taylora, co sugeruje nadmiernie restrykcyjną politykę pieniężną.

Równanie (5.45) można przekształcić następująco:

$$i = p_e(1+\beta) - \beta\pi + \Delta, \tag{5.47}$$

gdzie:

$$\Delta = i^* + (r - r^*) + \alpha(y - \hat{y}).$$

Przy założeniu, że $\Delta = const$, równanie (5.47) pokazuje, że wzrost oczekiwanej inflacji o jeden punkt procentowy wymaga, aby bank centralny – *ceteris paribus* – podniósł nominalne stopy procentowe o więcej niż jeden punkt procentowy, ponieważ $(1 + \beta) > 1$. Oznacza to, że wzrost inflacji wymaga wzrostu *realnych* stóp procentowych, aby sprowadzić inflację z powrotem do celu. Ten wymóg bywa niekiedy nazywany zasadą Taylora (*Taylor principle* – zob. Davig i Leeper 2007).

5.7.6. Inflacja jako skutek szoków zewnętrznych i efekty „drugiej rundy"

Jak już wspomnieliśmy, przyjęcie jakiekolwiek formalnej funkcji reakcji jako rygorystycznego zobowiązania banku centralnego do podjęcia określonych działań mogłoby nadmiernie usztywnić politykę pieniężną i w rezultacie obniżałoby jej skuteczność. Mogłoby też przyczyniać się do wzmacniania szoków w gospodarce. Automatyzm reakcji byłby szkodliwy zwłaszcza w dwóch sytuacjach. Pierwsza z nich ma miejsce, gdy mechanizm transmisji nie zapewnia stabilnej relacji pomiędzy instrumentami a celami polityki pieniężnej. Tak będzie w warunkach występowania zmiennych w czasie opóźnień między impulsem polityki pieniężnej a efektami oraz gdy występują czynniki zakłócające prawidłowe funkcjonowanie kanałów transmisji (nadpłynność sektora bankowego, podwyższone ryzyko suwerenne, ryzyko kontrahenta). Druga sytuacja ma miejsce wówczas, gdy inflacja jest wynikiem działania czynników niezależnych od polityki pieniężnej (czyli niezależnych od zmian ilości pieniądza w obiegu), takich jak wstrząsy podażowe (*supply-side shocks*) czy inne zdarzenia jednorazowe i przejściowe.

Kiedy inflacja jest wynikiem działania czynników niezależnych od polityki pieniężnej, bank centralny staje przed poważnym dylematem i nie może reagować

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w sposób mechaniczny. Takie zjawiska jak wzrost cen paliw i surowców na rynkach światowych czy nieurodzaj w rolnictwie oznaczają skokową zmianę kosztów wytwarzania w przypadku producentów krajowych. Są to negatywne szoki zewnętrzne (*external shocks*), których skutkiem jest wymuszony wzrostem kosztów wzrost cen wytwarzanych w kraju produktów[66]. Podobny wpływ na ceny mają także zdarzenia jednorazowe, takie jak katastrofy naturalne, wzrost podatków pośrednich, podniesienie cen administrowanych czy embargo handlowe. Ponieważ wzrost cen wywołany tymi zjawiskami nie jest jednak spowodowany polityką pieniężną, w tym w szczególności wzrostem podaży pieniądza, nie jesteśmy w stanie zlikwidować źródeł tego wzrostu cen przy pomocy polityki pieniężnej. Podniesienie (lub obniżenie) oficjalnych stóp procentowych przez krajowy bank centralny nie wywoła ani spadku cen ropy naftowej na rynkach międzynarodowych, ani spadku cen zbóż, o których decyduje przede wszystkim wielkość zbiorów, czyli podaż. W związku z tym ekonomiści są zdania, że na bezpośredni wzrost cen wywołany szokami podażowymi lub zdarzeniami jednorazowymi polityka pieniężna nie powinna reagować. W takiej sytuacji jakakolwiek sztywna funkcja reakcji nie może być stosowana. Reakcja w postaci zaostrzenia polityki pieniężnej jest potrzebna dopiero wówczas, gdy zaczynają pojawiać się tzw. efekty drugiej rundy (*second-round effects*), czyli wtórne efekty pieniężne pierwotnego szoku podażowego, a więc kiedy rośnie podaż pieniądza przy akomodacji początkowego wzrostu cen, np. w formie wzrostu kredytów finansujących wzrost kosztu zakupu surowców i paliw lub wzrost wynagrodzeń.

Efekty drugiej rundy można zilustrować na przykładzie negatywnego szoku zewnętrznego w postaci wzrostu cen ropy naftowej na rynkach międzynarodowych. Wywołuje on wzrost krajowych cen paliw, przekładający się bezpośrednio na wzrost inflacji. Wzrost ten z reguły nie powoduje jeszcze zaostrzenia polityki pieniężnej przez bank centralny, lecz jego konsekwencją pozostaje wzrost kosztów transportu. Gdy firmy transportowe podniosą ceny swoich usług, pozostałe przedsiębiorstwa będą zmuszone albo ten wzrost cen zaakceptować, albo zrezygnować ze współpracy. Gdy zdecydują się jednak płacić więcej za usługi transportowe, spróbują podnieść ceny swoich wyrobów, aby przerzucić wzrost kosztów na potencjalnych klientów. Pojawia się więc tendencja do ogólnego wzrostu cen, który zmusza przynajmniej część firm (lub końcowych odbiorców) do sięgnięcia po dodatkowe

[66] Możemy też mieć do czynienia z szokami pozytywnymi, takimi jak pojawienie się nowych wynalazków i zastosowanie nowych technologii, które powodują spadek cen.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kredyty obrotowe, aby pokryć zwiększone wydatki. Wzrost kredytów wywołuje jednak wzrost podaży pieniądza, co dla banku centralnego stanowi sygnał informujący o tym, że pojawiły się efekty drugiej rundy. W obliczu tej sytuacji bank zaostrza z reguły politykę pieniężną, czyli ogranicza wzrost podaży pieniądza, aby zahamować wzrost inflacji.

Inflacyjne skutki szoków zewnętrznych stawiają przed bankiem centralnym poważne wyzwanie. Zbyt szybkie podniesienie stóp procentowych może zepchnąć gospodarkę w recesję, a zarazem nie spowoduje zapobiegnie wzrostowi cen wywołanemu szokiem zewnętrznym. Z kolei zwlekanie z decyzją o podniesieniu stóp zbyt długo grozi nadmiernym wzrostem podaży pieniądza i wzrostem inflacji, co może powodować utrwalenie się oczekiwań inflacyjnych i obniży wiarygodność banku centralnego. Decyzji nie ułatwia też to, że dane o zmianach obiegu pieniężnego docierają do banku centralnego z pewnym opóźnieniem. Bank centralny staje więc zawsze przed dylematem: czy czekać z decyzją do momentu aż pojawią się efekty drugiej rundy, ryzykując, że reakcja będzie spóźniona, czy raczej reagować z wyprzedzeniem, ryzykując nadmierną restrykcyjność polityki pieniężnej.

5.7.7. Rola niepewności

Dylemat związany z podjęciem prawidłowej decyzji w obliczu szoku zewnętrznego przypomina po raz kolejny, że bank centralny działa w warunkach niepewności. Wynika ona z tego, że polityka pieniężna dotyczy głównie nieznanej przyszłości. W takich warunkach łatwo o pomyłkę, a te z kolei osłabiają skuteczność polityki pieniężnej i obniżają jej wiarygodność. Dlatego też banki centralne, świadome ryzyka popełnienia błędu, podejmują szereg działań mających na celu ograniczenie negatywnych skutków swoich ewentualnych błędnych decyzji.

Po pierwsze, w większości banków centralnych decyzje z zakresu polityki pieniężnej są podejmowane przez ciała kolegialne, a nie jednoosobowo. Organy te, takie jak Rada Prezesów (*Governing Council*) w Europejskim Banku Centralnym, Federalny Komitet Otwartego Rynku (*Federal Open Market Committee*) w amerykańskiej Rezerwie Federalnej, Komitet Polityki Pieniężnej (*Monetary Policy Committee*) w Banku Anglii czy Rada Polityki Pieniężnej w Narodowym Banku Polskim, składają się z kilku lub kilkunastu osób i podejmują decyzje większością głosów. Uważa się, że prawdopodobieństwo popełnienia błędu przez większość członków danego ciała jest mniejsze, niż gdyby decyzję miała podejmować jednak osoba. Ponadto, ciała kolegialne pozwalają na większe wewnętrzne zróżnicowanie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

poglądów, w szczególności umożliwiają przełamywanie ograniczeń związanych z „myśleniem grupowym" (*groupthink*), które pojawia się wśród przedstawicieli jednej grupy zawodowej reprezentujących ten sam lub zbliżony punkt widzenia (Haldane 2014). Dlatego też w skład ciał kolegialnych decydujących o polityce pieniężnej powinny wchodzić osoby z różnych środowisk, w tym spoza banku centralnego.

Po drugie, wypowiadając się na temat planowanej polityki, przedstawiciele banków centralnych unikają z reguły bardzo jasnych i precyzyjnych zapowiedzi co do swoich przyszłych decyzji. Wiedzą wprawdzie doskonale, że uczestnicy rynku cenią sobie przewidywalność polityki, lecz zdają sobie też sprawę z własnych ograniczeń. Posługując się bardziej niejednoznacznym i warunkowym językiem, bankierzy centralni pozostawiają sobie otwarte różne opcje polityki, co pozwala na ewentualne wycofanie się z wcześniejszych zapowiedzi bez uszczerbku dla wiarygodności[67]. Przykładem takiego „ezopowego" języka, który pozostawia wszystkie opcje otwarte, jest chociażby końcowy fragment Komunikatu ogłoszonego przez Radę Polityki Pieniężnej po posiedzeniu odbywającym w dniach 1–2 lipca 2014 roku: „Rada postanowiła utrzymać stopy procentowe NBP na niezmienionym poziomie. Decyzje Rady w kolejnych miesiącach będą uwarunkowane napływającymi informacjami wpływającymi na ocenę perspektyw wzrostu gospodarczego i inflacji w średnim okresie".

Innym jeszcze sposobem ograniczania ujemnych skutków ewentualnych błędnych decyzji jest dokonywanie zmian parametrów polityki pieniężnej stopniowo, metodą „małych kroków". W normalnych warunkach banki centralne zmieniają poziom stóp procentowych o ułamki punktu procentowego, z reguły o 25 punktów bazowych (0,25 punktu procentowego) w ramach serii podwyżek lub obniżek. Na przykład, kiedy gdy bank centralny widzi przestrzeń do znacznej obniżki stóp procentowych, rzędu 0,5–1,0 punktu procentowego, zastosuje najczęściej cykl mniejszych obniżek następujących po sobie w kilku kolejnych miesiącach – każda o 0,25 punktu – zamiast jednorazowo ciąć stopy o jeden czy pół punktu. Sens stopniowania polega na tym, że każdy miesiąc przynosi nową porcję danych statystycznych, które pozwalają lepiej ocenić, czy obrany kierunek polityki jest prawidłowy

[67] Były wieloletni prezes Rezerwy Federalnej Alan Greenspan słynął z tego rodzaju dość pokrętnych wypowiedzi w czasie swoich publicznych wystąpień. Przypisuje mu się m.in. następujące zdanie, które miało paść podczas jednego z przesłuchań w Senacie USA: „Wiem, że wydaje się wam, że rozumiecie to, co mówię, ale wydaje mi się, że nie zdajecie sobie sprawy, że to, co usłyszeliście, nie jest tym, co miałem na myśli (I know you think you understand what you thought I said but I'm not sure you realize that what you heard is not what I meant)".

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i czy kolejne obniżki okażą się uzasadnione. Takie konserwatywne zachowanie pozwala wprawdzie uniknąć nadmiernych wstrząsów w gospodarce, lecz zwiększa też ryzyko spóźnienia się z podjęciem decyzji.

5.7.8. Strategia komunikowania się z rynkami

Komunikacja z rynkami jest ważnym elementem działalności banków centralnych służącym zwiększaniu przewidywalności polityki pieniężnej i umacnianiu jej wiarygodności. Banki centralne i ich szefowie komunikują się z rynkiem na wiele sposobów: ogłaszając komunikaty, organizując konferencje prasowe, udzielając wywiadów, publikując raporty i analizy. Element komunikacji stanowi także ogłaszanie określonego „nastawienia" w polityce pieniężnej i posługiwanie się tzw. *forward guidance* (zob. punkt 5.4.4).

Rola komunikowania się z rynkami nabrała szczególnego znaczenia w ramach strategii bezpośredniego celu inflacyjnego. Wcześniej, gdy banki centralne skupiały się na utrzymywaniu określonej stopy wzrostu podaży pieniądza, uczestnicy rynku wiedzieli, czego można się spodziewać ze strony banku centralnego. Przyjęte reguły, w ramach których bank reagował na zmiany podaży pieniądza, zapewniały przejrzystość polityki pieniężnej, choć już niekoniecznie jej skuteczność. Ze strategią BCI jest inaczej. Wymaga ona „patrzenia do przodu", przewidywania przyszłości i wpływania na oczekiwania. Decyzje banku centralnego zależą od tego, jak postrzega on rozwój sytuacji gospodarczej i kształtowanie się procesów inflacyjnych w perspektywie średniookresowej. Aby uczynić politykę pieniężną dostatecznie przewidywalną i przejrzystą, bank centralny musi dzielić się z uczestnikami rynku swoją wiedzą na temat przyszłości, tłumaczyć dlaczego podjął taką a nie inną decyzję, a także wyjaśniać przyczyny popełnianych błędów.

Jak wiemy, nadrzędnym celem banku centralnego jest zapewnienie stabilności cen, co umożliwia utrzymanie gospodarki na ścieżce zrównoważonego wzrostu – gdy inwestorzy inwestują za dużo lub za mało, gospodarka może się „przegrzać" albo wpaść w recesję, a wówczas inwestorzy poniosą straty. Dlatego też operacyjnym celem komunikowania się z rynkami jest chęć zapewnienia większej przejrzystości i przewidywalności polityki pieniężnej, tak by ograniczyć niepotrzebne straty uczestników rynku. Na przykład, gdy inwestorzy przewidują spadek stóp procentowych, kupują obligacje, licząc na wzrost ich cen. Jeśli bank centralny nie ma jednak zamiaru obniżać stóp procentowych, ceny obligacji nie wzrosną, a inwestorzy nic nie zarobią, co więcej nawet – mogą ponieść straty. Gdy z kolei bank

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

centralny zapowiada, że zaostrzy politykę pieniężną (czyli podniesie stopy procentowe), inwestorzy zaczynają sprzedawać obligacje, bo spodziewają się, że ich cena zacznie spadać, a to wywołuje wzrost ich rentowności oraz wzrost rynkowych stóp procentowych. W ten sposób bank centralny osiąga więc cel, jakim jest zaostrzenie polityki pieniężnej, nawet bez podnoszenia oficjalnych stóp procentowych.

Drugim celem operacyjnym komunikowania się z rynkami jest wpływanie na oczekiwania inflacyjne i wzmacnianie w ten sposób mechanizmu transmisji polityki pieniężnej. Przykładowo, gdy inflacja utrzymuje się powyżej celu, ale wynika to z przejściowych szoków zewnętrznych, bank centralny może przeciwdziałać utrwalaniu się oczekiwań inflacyjnych poprzez informowanie, że wzrost inflacji ma charakter przejściowy; nie ma wówczas potrzeby podnoszenia stóp procentowych. Jeśli zarazem bank centralny zapowie, że nie dopuści do trwałego wzrostu inflacji, oczekiwania inflacyjne zaczną spadać, a to uruchomi mechanizm „samospełniającej się nadziei" i wywoła spadek rzeczywistej inflacji. Kluczem pozostaje wiarygodność banku centralnego – tylko wtedy bank centralny będzie w stanie skłaniać uczestników rynku do określonych zachowań bez konieczności podnoszenia oficjalnych stóp procentowych (Rozkrut 2011).

Polityka informacyjna jest szczególnie ważna, gdy dochodzi do długotrwałego odchylania się inflacji od celu na skutek działania czynników niezależnych od polityki pieniężnej (szoki zewnętrzne). Wówczas komunikaty banku centralnego stabilizują oczekiwania inflacyjne poprzez utrwalanie przekonania o przejściowym charakterze wzrostu inflacji i pozwalają wyjaśnić brak reakcji banku na utrzymującą się wyższą inflację.

Czynnikiem zwiększającym przejrzystość i przewidywalność polityki pieniężnej jest także publikowanie zapisów z przebiegu dyskusji na posiedzeniach organów decyzyjnych banku centralnego (tzw. *minutes*) oraz ogłaszanie wyników głosowań przeprowadzonych podczas posiedzeń. Pozwala to rozpoznać rozkład opinii wśród członków ciała decyzyjnego, dzięki czemu uczestnicy rynku mogą lepiej przewidywać przyszłe decyzje z zakresu polityki pieniężnej. Publikowanie *minutes* i wyników głosowania ma miejsce z reguły po upływie pewnego czasu od terminu posiedzenia ciała decyzyjnego[68], aby umożliwić opinii publicznej pełniejszą ocenę podjętych decyzji.

[68] W Polsce wyniki głosowań w Radzie Polityki Pieniężnej publikowane są po upływie 6 tygodni od posiedzenia.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.8. Polityka pieniężna w alternatywnych modelach makroekonomicznych

5.8.1. Model syntezy neoklasycznej

W rozdziale 2 opisaliśmy pokrótce główne koncepcje makroekonomiczne, które rozwinęły się w okresie ostatnich kilkudziesięciu lat po załamaniu się konsensu wokół syntezy neoklasycznej. Obecnie pokażemy, jaką rolę w ramach poszczególnych modeli odgrywa polityka pieniężna.

Podstawowym modelem makroekonomicznym wykorzystywanym w ramach *syntezy neoklasycznej* był omówiony wcześniej model IS-LM, uzupełniony kluczowym składnikiem w postaci krzywej Phillipsa, która zakładała krótkookresową wymienność pomiędzy inflacją i bezrobociem. Ta ostatnia zależność wskazywała na możliwość zarządzania popytem zagregowanym przy pomocy zmian podaży pieniądza (lub zmian poziomu wydatków publicznych), tak aby stabilizować produkcję i utrzymywać bezrobocie na niskim poziomie. Jednocześnie zwolennicy syntezy neoklasycznej dość sceptycznie odnosili się do możliwości skutecznego kontrolowania inflacji przy pomocy polityki pieniężnej, uznając, że inflacja zależy w dużym stopniu od czynników psychologicznych i cechuje się generalnie dużą inercją. Same zmiany podaży pieniądza były w jej ramach mniej istotnym czynnikiem inflacji niż zmiany w dostępności kredytu bankowego. Ogólnie rzecz biorąc, choć zwolennicy syntezy klasycznej uznawali politykę pieniężną za potencjalnie istotny instrument polityki stabilizacyjnej, to zarazem uważali, że w praktyce jej stosowanie jest mało skuteczne, i większą wagę przypisywali w tej dziedzinie polityce fiskalnej.

5.8.2. Monetaryzm i teoria realnego cyklu koniunkturalnego

Z kolei *monetaryzm*, stojąc na gruncie klasycznej dychotomii, jednoznacznie odrzucał wykorzystanie polityki pieniężnej do stymulowania gospodarki i postulował, aby ograniczyć jej zadanie do utrzymywania stabilności cen. Monetaryści wierzyli, że zmiany w zakresie podaży pieniądza mają wpływ na produkcję jedynie w krótkim okresie, z kolei w dłuższym okresie wpływ ten zanika i zmiany podaży pieniądza oddziałują jedynie na zmiany poziomu cen. W związku z tym cel polityki ma stanowić stabilizowanie poziomu cen, czyli walka z inflacją. Jednocześnie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

monetaryści uważali, że zależność między zmianami ilości pieniądza w obiegu a zmianami inflacji w krótkim okresie jest wysoce niestabilna i zatem trudna do wiarygodnego oszacowania. Dlatego też dyskrecjonalna polityka pieniężna nie przynosi oczekiwanych rezultatów. Biorąc to pod uwagę, Milton Friedman (1960, 1968) zalecał, aby bank centralny zwiększał ilość pieniądza w gospodarce według stałej stopy wzrostu, równej w przybliżeniu stopie wzrostu efektywności czynników produkcji, utrzymując jednocześnie nominalne stopy procentowe na poziomie bliskim zeru.

W ramach modeli *realnego cyklu koniunkturalnego* (RBC) polityka pieniężna jest zasadniczo neutralna, i to nawet w krótkim okresie. Dzieje się tak dlatego, że produkcja w ramach cyklu zmienia się głównie w reakcji na zmiany produktywności, będące z kolei efektem wprowadzania innowacji i zmian w technologiach wytwarzania. W rezultacie przekształcenia w zakresie podaży pieniądza mają charakter w pełni endogeniczny, stanowiąc reakcję na zmiany popytu na pieniądz wywołane zmianami produkcji (King i Plosser 1984). Innymi słowy, o podaży pieniądza decydują zasadniczo zmiany w sektorze realnym (stąd nazwa teorii). Teza o całkowitej nieistotności polityki pieniężnej w ramach teorii RBC była jednak krytykowana przez wielu ekonomistów (zob. np. Summers 1986, Mankiw 1989).

5.8.3. Modele nowej syntezy neoklasycznej

Teoria *nowej syntezy neoklasycznej* (NNS) łagodzi to skrajne stanowisko i pozwala sformułować szereg ważnych zaleceń pod adresem polityki pieniężnej. Po pierwsze – zgodnie z tradycją keynesowską – uznaje się, że polityka pieniężna wywiera istotny wpływ na zmienne realne, nawet w horyzoncie do kilku lat, ponieważ dostosowania cen przebiegają z opóźnieniem i stopniowo (zjawisko tzw. *staggered prices*). Po drugie – zgodnie z postulatem monetarystów – w dłuższym okresie skuteczność polityki pieniężnej maleje i ostatecznie zanika. Oznacza to, że wymienność między inflacją a zmiennymi realnymi w dłuższym okresie nie istnieje, podczas gdy ekspansja pieniężna wywołuje jedynie inflację. Po trzecie, obniżenie inflacji i utrzymywanie jej na niskim poziomie niesie istotne korzyści ekonomiczne w postaci poprawy efektywności transakcji i ograniczenia skutków zniekształceń cenowych. Po czwarte – zgodnie z koncepcją racjonalnych oczekiwań – podkreśla się znaczenie wiarygodności polityki dla jej skuteczności. Ogólnie rzecz biorąc, polityka pieniężna może wywierać wpływ na zmienne realne, ale

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

tylko wtedy, gdy zmiany w zakresie podaży pieniądza nie są antycypowane przez uczestników rynku. W przeciwnym razie polityka pieniężna staje się nieistotna.

Hipoteza o „nieistotności" polityki pieniężnej (*policy irrelevance*), nawiązująca do wcześniejszych prac Friedmana i Phelpsa, dotyczących naturalnej stopy bezrobocia, została w najpełniejszej formie sformułowana przez Sargenta i Wallace'a (1975). Zgodnie z ich argumentacją, jeśli tylko oczekiwania cenowe mają charakter racjonalny, zmiany poziomu produkcji nie zależą od reguły polityki pieniężnej wybranej przez władzę gospodarczą. Jak pamiętamy z rozdziału 2 (punkt 2.11.3), hipoteza o „nieistotności" polityki pieniężnej głosi, że tylko nieantycypowane zmiany w polityce pieniężnej wywierają wpływ na zmienne realne. Ponieważ uczestnicy rynku pozostają jednak racjonalni, są również w stanie w pełni przewidzieć skutki stosowania jakiejkolwiek reguły polityki pieniężnej, czyli polityki, która prowadzona jest systematycznie (w sposób powtarzalny). Zmiany cen okazują się więc w pełni przewidywalne, a to oznacza, że systematycznie prowadzona polityka nie jest w stanie wpływać na zmienne realne, takie jak produkcja i zatrudnienie[69].

Szczególnie interesujący w ramach teorii NNS jest mechanizm wpływu polityki pieniężnej na sektor realny w krótkim okresie. Ponieważ uwzględnia się istnienie niedoskonałej konkurencji na rynkach (oligopol, monopol), ceny rynkowe przewyższają koszty o wielkość marży netto. Jednocześnie, ze względu na to, że każdorazowa zmiana cen kosztuje, w danym okresie tylko część firm decyduje się na dostosowanie (podwyżkę) ceny (zgodnie z mechanizmem dostosowawczym zaproponowanym przez Calvo 1983)[70] i ustala nową marżę. Pozostałe firmy utrzymują ceny bez zmian. Zasadniczo w ramach teorii NNS ceny ustalane są przez monopolistyczne firmy w ramach dynamicznego procesu uwzględniającego historyczny poziom cen (*dynamic price setting*) oraz oczekiwane zmiany w poziomie obecnego i przyszłego kosztu krańcowego (Goodfriend i King 1997).

[69] Hipotezę o nieistotności polityki pieniężnej podważył Stanley Fischer. Wskazał on, że jeśli sztywność płac wynika z kontraktów płacowych zawieranych na okresy dłuższe niż częstotliwość podejmowania decyzji przez władze monetarne, polityka pieniężna może wpływać na zmiany produkcji i zatrudnienia, nawet gdy oczekiwania są racjonalne (Fischer 1977).

[70] Mechanizm zaproponowany przez Calvo (1983) zakłada, że prawdopodobieństwo, iż w danym okresie cena pozostanie bez zmian, wynosi α, a prawdopodobieństwo, że cena zostanie podniesiona, wynosi $(1 - \alpha)$. Wówczas średni czas trwania danej ceny jest równy $\frac{1}{(1-\alpha)}$. Na przykład, jeśli $\alpha = 0{,}75$, oznacza to, że w danym okresie, na przykład w ciągu kwartału, średnio jedna czwarta firm podnosi ceny. Wówczas średni czas utrzymywania się sztywnej ceny wynosi 4 kwartały.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Oznaczmy *krańcową* marżę brutto (*marginal markup*) w firmach, które dostosowują ceny do poziomu P^* w okresie *t*, jako:

$$\mu_t^* = \frac{P_t^*}{\Psi_t}, \tag{5.48}$$

a *średnią* marżę brutto (*average markup*) w całej gospodarce jako:

$$\mu_t = \frac{P_t}{\Psi_t}, \tag{5.49}$$

gdzie Ψ_t jest kosztem krańcowym w okresie *t*, i oczywiście μ_t^*, $\mu_t > 1$. Na podstawie równania (5.49) otrzymujemy:

$$\frac{\Psi_t}{P_t} = \frac{1}{\mu_t} = \gamma_t, \tag{5.50}$$

gdzie γ_t jest *realnym* kosztem krańcowym w czasie *t*. Równanie (5.50) pokazuje, że gdy w wyniku wzrostu popytu realny koszt krańcowy γ rośnie (bo część firm nie podnosi cen), to jednocześnie średnia marża brutto μ_t maleje. Skłoni to firmy, które nie podniosły cen w okresie *t*, do ich podnoszenia w okresach następnych. Dlatego też teoria NNS postuluje, aby polityka pieniężna dążyła do stabilizowania marż na możliwie niskim poziomie, co zapewni niską inflację. Polityka pieniężna powinna więc wygładzać skutki wstrząsów popytowych, a zarazem – zgodnie z postulatem teorii RBC –dostosowywać się (akomodować) do wstrząsów wynikających ze zmian produktywności czynników wytwórczych (wstrząsy podażowe).

5.9. Polityka pieniężna w nowych modelach keynesowskich*

5.9.1. Wprowadzenie

Większość ekonomistów akceptuje pogląd, że sposób prowadzenia polityki pieniężnej ma istotne znaczenie w przypadku działalności gospodarczej. Postęp badań nad różnymi odmianami strategii polityki, w tym uwzględnienie optymalizacji wielookresowej przez uczestników rynku, roli oczekiwań i problemu niespójności czasowej, doprowadził do wyłonienia się nowej klasy modeli makroekonomicznych, uwzględniających wyniki tych analiz. Współczesne podejście polega coraz częściej na wykorzystaniu dynamicznych stochastycznych modeli równowagi ogólnej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(*dynamic stochastic equilibrium models* – DSGE), które uwzględniają mikroekonomiczne podstawy zachowania producentów i konsumentów, a także biorą pod uwagę niektóre standardowe założenia keynesowskie, takie jak sztywność cen i płac nominalnych (Bernanke, Gertler i Watson 1997; Goodfriend i King 1997; Gali 2008) i wynikającą z niej krótkookresową wymienność (*trade-off*) między produkcją a inflacją. Dlatego też tę nową klasę modeli nazywa się nowymi modelami keynesowskimi (*New Kenyesian Models* – NKM)[71]. Jednym z pierwszych tego typu układów jest model Barro-Gordona, opisany wcześniej w punkcie 5.6.1. Uwzględnia on częściową sztywność płac oraz wpływ oczekiwań na zachowania gospodarstw domowych. Poniżej omówimy model bardziej rozbudowany, w którym dodatkowo rozpatruje się w sposób bezpośredni konkretną funkcję preferencji banku centralnego uwzględniającą dwa cele – inflację i poziom produkcji – oraz przyjmuje się, że firmy i gospodarstwa domowe optymalizują swoje zachowania w perspektywie wielookresowej.

5.9.2. Model Clarida, Galì i Gertner: polityka dyskrecjonalna

Jednym z najbardziej znanych nowych modeli keynesowskich jest model rozwinięty przez Richarda Claridę, Jordi Galí i Marka Gertlera (Clarida, Gali i Gertler, 1999), dalej zwany jako model CGG, który jest omówiony poniżej. Instrumentem polityki pieniężnej w ramach tego modelu jest krótkookresowa nominalna stopa procentowa ustalana przez bank centralny, a zmiennymi celu są poziom inflacji i luka podażowa (*output gap*). Główne pytanie dotyczy tego, na jakim poziomie należy ustalić stopę procentową, aby wybrane zmienne celu przyjęły pożądane wartości. Ważną cechą modelu jest uwzględnienie częściowo racjonalnych oczekiwań. Zakłada się, że zachowanie sektora prywatnego (firmy i gospodarstwa domowe) zależy nie tylko od bieżącej polityki pieniężnej, lecz także od oczekiwań co do przyszłej polityki pieniężnej. Oznacza to, że zapowiedzi banku centralnego co do przyszłych działań oraz wiarygodność tych zapowiedzi stają się ważnym czynnikiem określającym zachowanie sektora realnego (Persson i Tabellini 1997). Im bardziej wiarygodne są te zapowiedzi, tym silniejszy stabilizujący wpływ bank centralny może wywierać na gospodarkę. Wynika z tego, że przyjęcie wiążących zobowiązań co do stosowanej reguły polityki pieniężnej w przyszłości albo wdrożenie budzących

[71] Zob. np. Blanchard i Galí (2005, 2010). W polskiej literaturze Rzońca (2014) posługuje się pojęciem nowego keynesowskiego schematu analitycznego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zaufanie rozwiązań instytucjonalnych (jak na przykład zapewnienie niezależności banku centralnego), może być źródłem istotnych korzyści makroekonomicznych.

W punkcie wyjścia w wersji bazowej modelu CGG zakłada się brak zobowiązań banku centralnego co do przyszłej polityki pieniężnej – bank centralny ogranicza się jedynie do ogłoszenia deklaracji, że będzie dążył do utrzymania niskiej inflacji, przy utrzymaniu produkcji na poziomie „naturalnym" (czyli na poziomie potencjalnego PKB), nie ogłasza natomiast żadnej określonej funkcji reakcji. Taki stan rzeczy był dość typowy dla polityki banków centralnych w większości krajów w latach 80. i 90. W wielu przypadkach bank centralny nie wyjaśniał nawet, co dokładnie miała oznaczać „niska inflacja" i jaki poziom produkcji uznawano za naturalny (potencjalny). W ramach tego modelu polityka pieniężna oddziałuje w krótkim okresie na gospodarkę w sposób podobny jak w modelu IS-LM (czyli poprzez wpływ bieżących realnych stóp procentowych na popyt i produkcję). Istnieją jednak dwie zasadnicze różnice: po pierwsze, funkcje popytu i podaży są wynikiem optymalizacyjnych zachowań podmiotów gospodarczych[72], a – po drugie – zachowania podmiotów zależą także od oczekiwań co do przyszłych wartości stóp procentowych.

Zdefiniujemy lukę podażową x_t w czasie t jako różnicę między produktem rzeczywistym y_t a produktem potencjalnym („naturalnym") z_t, czyli $x_t = y_t - z_t$. Oznaczmy następnie stopę inflacji jako π, a nominalną stopę procentową jako i. Wówczas stronę popytu w modelu określa równanie (5.51), a stronę podaży równanie (5.52):

$$x_t = -\varphi\left(i_t - E_t\pi_{t+1}\right) + E_t x_{t+1} + g_t \tag{5.51}$$

$$\pi_t = \delta x_t + \beta E_t \pi_{t+1} + u_t . \tag{5.52}$$

Przy czym:

$$g_t = \mu g_{t-1} + \hat{g}_t$$

$$u_t = \rho u_{t-1} + \hat{u}_t,$$

gdzie g_t reprezentuje losowy szok popytowy, u_t wyznacza losowy szok podażowy, $\hat{g}_t$ i $\hat{u}_t$ są zmiennymi losowymi o średniej 0 i stałej wariancji σ_g^2 i σ_u^2, a operator E

[72] To znaczy, że firmy ustalają ceny, dążąc do maksymalizacji zysku, a gospodarstwa domowe dostosowują bieżące wydatki, zmierzając do optymalnego wygładzania konsumpcji w czasie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

oznacza wartość oczekiwaną. Pierwsze równanie (5.51) jest zbliżone do tradycyjnej funkcji *IS*, ale różni się od niej tym, że bieżący produkt zależy nie tylko od obecnej realnej stopy procentowej, lecz także od oczekiwanej wartości produktu w przyszłości. Wyższa oczekiwana wartość produktu $E_t x_{t+1}$ wpływa na wzrost bieżącej produkcji, ponieważ ludzie oczekują wyższych dochodów i, dążąc do wygładzania konsumpcji w czasie, zwiększają swój bieżący popyt i wydatki. Wysokość realnej stopy procentowej $\left(i_t - E_t \pi_{t+1}\right)$ wpływa na międzyokresową substytucję konsumpcji, a znak „minus" stojący przy współczynniku międzyokresowej elastyczności substytucji φ względem realnej stopy procentowej pokazuje, że wzrost realnej stopy procentowej skłania do ograniczania wydatków i obniża produkcję. Wreszcie, dodatni szok popytowy g_t wywołuje proporcjonalny wzrost produkcji.

Drugie równanie (5.52) opisuje zmodyfikowaną krzywą Phillipsa. Wzrost luki podażowej pociąga za sobą wzrost inflacji, ponieważ rośnie zatrudnienie i poziom wynagrodzenia. Lecz bieżąca inflacja zależy także od oczekiwanej inflacji w przyszłości, zgodnie z mechanizmem wpływu oczekiwań inflacyjnych na bieżące wynagrodzenia, czyli $\beta > 0$ (*expectations-augmented Phillips curve* – zob. Blanchard 1997). W modelu przyjmuje się założenie, że firmy dostosowują ceny, dążąc do maksymalizacji zysku, lecz częstotliwość tych zmian uwzględnia ograniczenia wynikające z istnienia określonych kosztów tych dostosowań, zgodnie z mechanizmem zaproponowanym przez Calvo (1983)[73]. To, że niektóre ceny pozostają stałe z okresu na okres, pozwala uwzględnić w modelu częściową sztywność cen w krótkim okresie. Współczynnik β mierzy siłę tego związku (więcej na temat zmodyfikowanej krzywej Philipsa zob. ramka 5.15).

Rozwijając iteracyjnie równanie (5.51) na przyszłe okresy, otrzymujemy:

$$x_t = E_t \sum_{k=0}^{\infty} \left[-\varphi\left(i_{t+k} - \pi_{t+1+k}\right) \right] + g_{t+k}. \tag{5.53}$$

Jak widać, obecna luka podażowa zależy nie tylko od bieżącej realnej stopy procentowej i szoku popytowego, ale także od oczekiwanej ścieżki wartości tych zmiennych w przyszłości. W takim stopniu, w jakim polityka pieniężna ma wpływ

[73] Koszty związane ze zmianami cen (tzw. koszt „menu", koszt zmian kontraktów, koszt pogorszenia relacji z klientami itp.) powodują, że w każdym okresie tylko część firm dostosowuje ceny swoich produktów. Można więc przyjąć, że w danym okresie prawdopodobieństwo, iż dana firma pozostawi ceny bez zmian jest równe θ, a prawdopodobieństwo zmiany pozostaje równe $(1 - \theta)$. Zakładając niezależność rozkładu prawdopodobieństwa w każdym okresie, średnia długość okresu, w którym ceny pozostają bez zmian, jest równa $\left(\frac{1}{1-\theta}\right)$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

na stopy realne dzięki istnieniu nominalnych sztywności, wpływa ona na zagregowany popyt i w konsekwencji na produkcję. Rozwijając iteracyjnie równanie (5.52) na przyszłe okresy, otrzymujemy:

$$\pi_t = E_t \sum_{k=0}^{\infty} \beta^k \left(\delta x_{t+k} + u_{t+k} \right). \tag{5.54}$$

W przeciwieństwie do standardowej funkcji Phillipsa uwzględniającej przyszłe oczekiwania, współczesna inflacja π_t nie zależy inercyjnie od swoich wartości z przeszłości, lecz całkowicie od obecnych i przyszłych wartości produktu oraz szoków podażowych. Firmy ustalają ceny na podstawie oczekiwanych kosztów krańcowych, które z kolei zależą od wielkości nadwyżkowego popytu (gdy popyt rośnie, rośnie produkcja, i w krótkim okresie wzrastają koszty krańcowe produkcji). Z kolei szok podażowy u_t reprezentuje zmiany kosztów krańcowych wywołane innymi czynnikami niż nadwyżkowy popyt.

Instrumentem polityki pieniężnej jest krótkookresowa nominalna stopa procentowa *i*. Zmiana *i*, przy istniejących sztywnościach nominalnych (inflacja zależy od oczekiwanych wartości luki podażowej), pozwala wpływać na stopę realną, a w konsekwencji a także na produkcję. Ponieważ zarówno gospodarstwa domowe, jak i firmy kierują się oczekiwaniami (firmy są *forward-looking*), będą starały się jednak przewidzieć, w jaki sposób polityka pieniężna zareaguje w przyszłości na pojawiające się szoki.

Aby zamknąć model, przyjmijmy, że bank centralny dąży do minimalizacji sumy kwadratów odchyleń luki podażowej oraz inflacji od założonych celów[74, 75]. Wówczas funkcja wyboru banku centralnego przyjmie następującą postać:

$$L = \frac{1}{2} E_t [\sum_{k=0}^{\infty} \beta^k \left(\alpha x_{t+k}^2 + \pi_{t+k}^2 \right)] = min, \tag{5.55}$$

gdzie parametr α mierzy relatywną wagę odchylenia produkcji od poziomu potencjalnego względem odchylenia inflacji od celu inflacyjnego. Oczywiście, funkcja (5.55) osiąga bezwzględne minimum dla $x_t = 0$ oraz $\pi_t = 0$. O ile tempo wzrostu produktu potencjalnego jest różne w przypadku poszczególnych krajów, o tyle

[74] W przypadku luki podażowej tym celem najczęściej będzie $x_t = 0$, z kolei odnośnie inflacji można przyjąć, że celem inflacyjnym jest pewna dodatnia, lecz niska stopa inflacji π^*; wówczas π można interpretować jako odchylenie inflacji od celu.

[75] Uzasadnienie w przypadku wykorzystania funkcji minimalizującej sumę kwadratów odchyleń od zadanych celów jako funkcji preferencji można znaleźć m.in. w pracy Rottemberga i Woodforda (1997).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przyjmuje się, że w gospodarkach rozwiniętych cel inflacyjny powinien zawierać się w przedziale od 1% do 3% (Bernanke i Mishkin 1997).

Funkcja wyboru (5.55) oraz dwa równania (5.51) i (5.52) tworzą razem problem optymalizacyjny, którego rozwiązanie pozwala określić ścieżkę nominalnych stóp procentowych i_t, zapewniającą minimum funkcji wyboru. Główna różnica w stosunku do klasycznych zagadnień optymalizacyjnych polega na tym, że w tym przypadku zmienne celu (inflacja oraz produkcja) zależą nie tylko od bieżącej polityki (stopa procentowa banku centralnego), lecz także wprost od oczekiwań co do przyszłej polityki. Ponieważ luka podażowa zależy od przyszłych stóp procentowych (im niższe stopy, tym wyższa produkcja – równanie (5.53)), a inflacja od bieżącej i przyszłych wartości luki podażowej (im wyższa luka, tym wyższa inflacja – równanie (5.54)), przewidywania co do przyszłych stóp procentowych wpływają na obecną inflację.

W tym kontekście wiarygodność zobowiązań banku centralnego nabiera szczególnego znaczenia. Na przykład bank centralny, który zobowiąże się w sposób *wiarygodny* do utrzymania niskiej inflacji w przyszłości, może obniżyć bieżącą inflację mniejszym kosztem makroekonomicznym w postaci utraconej produkcji niż inny bank, centralny, który takiego zobowiązania nie podejmie. W takiej sytuacji polityka oparta na wiarygodnych zobowiązaniach – na przykład na ogłoszonych wcześniej regułach – może przynosić wymierne korzyści makroekonomiczne.

Kluczowa różnica między polityką dyskrecjonalną a polityką opartą na zobowiązaniach nie polega jednak na tym, czy bank centralny podejmuje jakieś zobowiązanie, czy też nie, lecz na tym, czy podjęte zobowiązanie w sposób wiarygodny ogranicza swobodę banku centralnego w prowadzeniu polityki pieniężnej w przyszłości. Tylko wtedy, gdy uczestnicy rynku wierzą bowiem, że bank centralny będzie postępował zgodnie z przyjętymi regułami, sami będą także dostosowywać swoje obecne zachowania do oczekiwanych decyzji.

Aby ukazać korzyści wynikające z prowadzenia polityki opartej na zobowiązaniach (regułach), autorzy modelu CGG najpierw rozpatrują – jako punkt odniesienia – sytuację, w której bank centralny nie podejmuje żadnych zobowiązań. W takim przypadku określenie optymalnej polityki pieniężnej polega na wyborze w każdym okresie trzech zmiennych – x_t, π_t, i_t – tak, aby zminimalizować funkcję wyboru *L*. Można ten proces optymalizacji polityki podzielić na dwa etapy. Najpierw bank centralny określa takie wartości x_t i π_t, które zapewniają minimum funkcji *L*, a następnie dla tych wartości zmiennych celu bank określa taką wysokość stopy procentowej *i*, która spełnia warunek krzywej *IS*, czyli równanie (5.51).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Zakładając, że sektor prywatny formułuje oczekiwania w sposób racjonalny, bank centralny w każdym okresie minimalizuje funkcję:

$$L = \frac{1}{2}\left(\alpha x_t^2 + \pi_t^2\right) + F_t. \tag{5.56}$$

Przy ograniczeniu:

$$\pi_t = \delta x_t + f_t \tag{5.57}$$

dla danego F_t i f_t, przy czym:

$$F_t = -\frac{1}{2} E_t [\sum_{k=0}^{\infty} \beta^k \left(\alpha x_{t+k}^2 + \pi_{t+k}^2\right)]$$

$$f_t = \beta E_t \pi_{t+1} + u_t.$$

Równania (5.56) i (5.57) to przekształcone równania (5.51) i (5.52). Pokazują one, że przy polityce dyskrecjonalnej przyszła inflacja i produkcja nie zależą od obecnej polityki oraz że bank centralny nie wpływa na oczekiwania (ponieważ jego ewentualne zapowiedzi nie są wiarygodne, co oznacza, że oczekiwania sektora prywatnego muszą być traktowane przez bank centralny jako zmienne egzogeniczne).

Rozwiązaniem problemu (5.55) ÷ (5.57) jest następujący warunek optymalności[76]:

$$x_t = -\frac{\delta}{\alpha}\pi_t, \tag{5.58}$$

z którego wynika, że gdy inflacja wzrasta powyżej celu ($\pi_t > 0$), trzeba ograniczać popyt, aby obniżyć lukę podażową x_t. Tempo spowalniania popytu jest proporcjonalne do współczynnika δ, który mierzy skalę obniżenia inflacji na jednostkę

[76] Aby otrzymać warunek (5.58), budujemy funkcję Lagrange'a dla układu (5.56) ÷ (5.57) o postaci:

$$L = \frac{1}{2}\left(\alpha x_t^2 + \pi_t^2\right) + F_t - \lambda\left(\pi_t - \delta x_t + f_t\right),$$

gdzie λ jest mnożnikiem Lagrange'a. Różniczkując względem zmiennych celu i przyrównując do zera, otrzymujemy (pomijamy dla uproszczenia zapisu indeks czasu t):

$$\frac{\delta L}{\delta x} = \alpha x + \lambda\delta = 0$$

$$\frac{\delta L}{\delta \pi} = \pi - \lambda = 0.$$

Z drugiego równania otrzymujemy $\pi = \lambda$. Podstawiając do pierwszego z równań, uzyskujemy (5.58).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

utraconej produkcji, oraz odwrotnie proporcjonalne do relatywnej wagi utraconej produkcji wobec obniżenia inflacji (współczynnik α).

Aby otrzymać formę zredukowaną w przypadku x_t i π_t, podstawiamy warunek optymalności (5.58) do równania zagregowanej podaży (5.52) i, zakładając nadal, że oczekiwania sektora prywatnego są racjonalne, otrzymujemy w rezultacie optymalne funkcje reakcji w ramach polityki dyskrecjonalnej (po przekształceniach):

$$x_t = -\delta q u_t \tag{5.59}$$

$$\pi_t = \alpha q u_t, \tag{5.60}$$

gdzie

$$q = \frac{1}{\delta^2 + \alpha(1 - \beta\rho)}.$$

Bank centralny dostosowuje więc zmienne celu x i π do szoków podażowych u, przy danych wartościach parametrów ρ, α, β, δ.

Przechodząc do drugiego etapu procedury, określamy optymalną funkcję reakcji banku centralnego, czyli wysokość wymaganej stopy procentowej, zapewniającą spełnienie równania krzywej *IS*. Wstawiając wartość x_t z równania (5.59) do równania (5.52) i przekształcając, otrzymujemy:

$$E_t \pi_{t+1} = \rho \pi_t = \rho \alpha q u_t.$$

I ostatecznie uzyskujemy:

$$i_t = \gamma_\pi E_t \pi_{t+1} + \frac{1}{\varphi} g_t, \tag{5.61}$$

gdzie:

$$\gamma_\pi = 1 + \frac{(1-\rho)\delta}{\rho\varphi\alpha} > 1.$$

Ponieważ $\gamma_\pi > 1$, to równanie (5.61) pokazuje, że optymalną polityką banku centralnego przy braku zobowiązań (w ramach dyskrecjonalności) jest podnoszenie stopy procentowej o więcej niż jeden punkt procentowy na każdy punkt procentowy wzrostu inflacji powyżej celu i obniżanie stóp procentowych o więcej niż punkt procentowy na każdy punkt procentowy spadku inflacji poniżej celu. Jak już wiadomo,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

taką regułę reagowania nazywa się także zasadą Taylora (zob. punkt 5.7.5). Dodatkowo wysokość stopy procentowej powinna uwzględniać obecność szoku popytowego g_t.

Ramka 5.17. Neokeynesowska krzywa Phillipsa, konflikt celów w polityce pieniężnej i „niebiańska zbieżność"

W ramach standardowego nowego modelu keynesowskiego (*New Keynesian Model* – NKM) przyjmuje się założenie, że ceny i płace są sztywne w krótkim okresie (zgodnie z mechanizmem opisanym przez Calvo 1983) po stronie podaży, oraz założenie, że gospodarstwa domowe optymalizują konsumpcję w ujęciu wielookresowym po stronie popytu. Zakłada się też, że bank centralny prowadzi politykę pieniężną zgodnie z regułą Taylora, co oznacza, że minimalizuje odchylenia produkcji od produktu potencjalnego oraz inflacji od celu inflacyjnego. Wówczas neokeynesowska krzywa Phillipsa może być opisana następującym równaniem, stanowiącym wersję równania (5.52):

$$\pi_t = \alpha\left(y_t - z_t\right) + \beta E_t\left(\pi_{t+1}\right), \tag{5.62}$$

gdzie oznaczenia są wyjaśnione w tekście powyżej. Taki sposób opisu polityki pieniężnej jest często krytykowany ze względu na trudności z określeniem rozmiaru luki podażowej $\left(y_t - z_t\right)$. Trudności te wynikają z dwóch powodów. Po pierwsze, produkt potencjalny jest trudny do oszacowania, a na dodatek podlega zmianom w czasie. Po drugie, dane dotyczące rzeczywistego PKB spływają z dużym opóźnieniem i z reguły nie są dostępne z częstotliwością większą niż raz na kwartał. Uzależnianie decyzji polityki pieniężnej od kształtowania się luki popytowej w sytuacji, gdy dotyczące jej dane są tak trudno dostępne, jest więc dość ryzykowne (zob. np. Hall 2005, Mankiw 2005).

Zauważmy jednak, że z równania (5.62) pośrednio wynika, że dwa cele realizowane przez bank centralny nie muszą być wzajemnie sprzeczne. Przyjmijmy, że produkcja znajduje się na poziomie potencjalnym a inflacja równa się celowi, i załóżmy dalej, że negatywny szok podażowy (na przykład wzrost cen ropy naftowej) obniża produkt potencjalny z_t. Pociąga to za sobą wzrost wartości różnicy $\left(y_t - z_t\right)$ oraz wzrost prawej strony równania (5.62). W rezultacie inflacja rośnie. Gdy bank centralny podniesie stopy procentowe, inflacja powróci do celu, a zarazem rzeczywista produkcja obniży się do nowego poziomu potencjalnego. Polityka

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pieniężna, nakierowana wyłącznie na stabilizowanie cen, osiąga również, niejako mimochodem, także i cel drugi w postaci stabilizowania produkcji na poziomie potencjalnym. Olivier Blanchard nazwał tę właściwość dość patetycznie „niebiańską zbieżnością" (*divine coincidence*). W takiej sytuacji w polityce banku centralnego znika wymienność między stabilizowaniem inflacji a stabilizowaniem luki podażowej, w związku z czym nie występuje konflikt celów polityki.

Ten brak wymienności wynika jednak z przyjęcia założenia o endogeniczności produktu potencjalnego oraz założenia o braku realnych zniekształceń rynkowych, takich zwłaszcza jak sztywność płac realnych (*real wage rigidities*). Gdy jednak uchyli się te założenia, „niebiańska zbieżność" znika i bank centralny ponownie staje w obliczu konieczności dokonywania wyboru między stabilizacją inflacji a stabilizacją luki podażowej (Blanchard i Galí 2005).

5.9.3. Model CGG: polityka pieniężna oparta na wiążących regułach

Podkreśliliśmy wyżej, że wiarygodność polityki pieniężnej jest istotna dla utrzymania niskiej inflacji. Z jednej strony chodzi o potencjalną skłonność banku centralnego do tolerowania wyższej inflacji w dążeniu do utrzymania produktu powyżej poziomu potencjalnego (zob. model Barro-Gordona, punkt 5.6.1). W takiej sytuacji uczestnicy rynku nie dają wiary zapewnieniom banku centralnego, że będzie starał się on utrzymać niską inflację w przyszłości. Z drugiej strony brak zaufania do banku centralnego może sprawić, że obniżenie inflacji będzie wiązać się z większym kosztem makroekonomicznym w postaci utraconej produkcji. W takiej sytuacji uczestnicy rynku wątpią w trwałość polityki antyinflacyjnej. W obu przypadkach można jednak uniknąć tych problemów dzięki zwiększonej wiarygodności banku centralnego. Mianowicie, gdy proces ustalania cen zależy od oczekiwań co do przyszłej polityki pieniężnej (tak jak zakłada się w modelu CGG), można dzięki zwiększonej wiarygodności obniżyć bieżącą inflację.

Załóżmy najpierw, że mamy do czynienia z pierwszym przypadkiem, czyli że celem banku centralnego jest utrzymanie luki podażowej na poziomie wyższym od zera, na przykład $x^* - x = h > 0$, gdzie x^* jest pożądanym poziomem luki podażowej. Wówczas funkcja wyboru przyjmuje postać[77]:

[77] Zauważmy, że wówczas społecznie optymalny poziom luki podażowej $x + h$ jest wyższy od poziomu x zgodnego z produktem potencjalnym. Wskazana sytuacja może wystąpić w warunkach istnienia zniekształceń rynkowych, takich jak konkurencja monopolistyczna lub podatki.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$L = \frac{1}{2} E_t [\sum_{k=0}^{\infty} \beta^k \left(\alpha \left(x_{t+k} - h \right)^2 + \pi_{t+k}^2 \right) \Big] = min. \tag{5.63}$$

Warunek optymalności (5.58) ulega modyfikacji i przyjmuje postać następującą:

$$x_t^h = -\frac{\delta}{\alpha} \pi_t^h + h, \tag{5.64}$$

gdzie x_t^h i π_t^h oznaczają wartości optymalne zmiennych celu dla $h > 0$. Podstawiając (5.64) do równania krzywej *IS* (5.51) i do równania zmodyfikowanej krzywej Phillipsa (5.52), otrzymujemy:

$$x_t^h = x_t \tag{5.65}$$

oraz

$$\pi_t^h = \pi_t + \frac{\alpha}{\delta} h. \tag{5.66}$$

Równanie (5.65) pokazuje, że dla $h > 0$, luka podażowa jest ostatecznie taka sama, czyli nie ma trwałego przyrostu produktu potencjalnego. Z kolei równanie (5.66) pokazuje, że inflacja jest systematycznie wyższa, ponieważ drugi składnik prawej strony równania jest zawsze większy od zera. Oznacza to, że jeśli celem banku centralnego jest utrzymanie produktu powyżej poziomu potencjalnego, to skutkiem takiej polityki będzie permanentnie wyższa inflacja, przy utrzymaniu produktu na poziomie potencjalnym. Jest to więc rozwiązanie bezwarunkowo gorsze od poprzedniego, i to także w przypadku gdy bank centralny ukrywa swoje prawdziwe intencje (czyli gdy $h > 0$). Wobec braku zobowiązania co do reguły reagowania na inflację, racjonalni uczestnicy rynku przewidują, że bank będzie starał się prowadzić politykę proinflacyjną, która pozwoli zwiększyć produkcję ponad poziom potencjalny aż do momentu, w którym dalsze zwiększanie inflacji nie zapewnia żadnych krótkookresowych przyrostów produkcji (jest to punkt *G* na rysunku 5.6)[78].

[78] Niektórzy autorzy zwracają jednak uwagę na to, że racjonalny bank centralny nie będzie „oszukiwał" rynków, bo zdaje sobie sprawę z tego, że w warunkach istnienia racjonalnych oczekiwań takie działanie byłoby nieskuteczne i szkodliwe dla reputacji banku (McCallum 1997a). Także Alan Blinder, na podstawie własnych doświadczeń jako gubernatora Fed, (Blinder, 1997) potwierdza, że w praktyce Fed nigdy nie próbował „oszukiwać" rynków.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Gdyby bank centralny zadeklarował jednak w sposób wiarygodny, że dąży do $h = 0$, osiągnąłby ten sam poziom produkcji, lecz przy niższej inflacji. Lecz jak bank centralny może przekonać uczestników rynku, że nie ma zamiaru ich „oszukiwać"? Jedno z możliwych rozwiązań sugeruje Rogoff (1985), proponując, aby politycy mianowali szefem banku centralnego osobę „konserwatywną", która ma większą niż przeciętna w danym społeczeństwie awersję do inflacji (zob. ramka 5.12). Oznacza to, że w równaniu (5.60) współczynnik α musi być odpowiednio mniejszy, aby ograniczyć wpływ h na tolerowany poziom inflacji. Lecz całkowite wyeliminowanie wpływu h na poziom inflacji ma miejsce dopiero wtedy, gdy $\alpha = 0$, czyli gdy bankiem centralnym kieruje „ekstremista" nieprzywiązujący w ogóle wagi do poziomu produkcji[79].

W warunkach równowagi optymalną polityką jest dostosowywanie wielkości luki podażowej do szoków podażowych, zgodnie z równaniem (5.60). Załóżmy obecnie, że bank centralny zobowiązuje się publicznie do reagowania wyłącznie na szoki podażowe u i że jego funkcja reakcji ma postać następującą (dla wszystkich wartości t):

$$x_t^c = -\omega u_t, \tag{5.67}$$

gdzie x_t^c jest *warunkową* wartością luki podażowej, czyli pod warunkiem podjęcia przez bank centralny określonego zobowiązania co do funkcji reakcji, a ω oznacza współczynnik reakcji (*feedback rule*). Zauważmy, że optymalna polityka w ramach dyskrecjonalności jest specjalnym przypadkiem funkcji reakcji (5.67) dla $\omega = \delta q$. Podstawiając (5.67) do równania krzywej Phillipsa (5.52), otrzymujemy wzór na warunkową wartość inflacji π_t^c (czyli pod warunkiem podjęcia przez bank centralny zobowiązania co do funkcji reakcji):

$$\pi_t^c = \left(\frac{1-\delta\omega}{1-\beta\rho}\right) u_t. \tag{5.68}$$

[79] Zarówno Rogoff (1985), jak i Clarida, Galí i Gertler (1999) podkreślają, że „ekstremistyczna" polityka dla $\alpha = 0$ byłaby szkodliwa z punktu widzenia dobrobytu społecznego i że istnieje pewien optymalny poziom konserwatyzmu bankierów centralnych, wyraźnie niższy od nieskończoności, który zapewnia optymalizację funkcji dobrobytu społecznego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Jak łatwo zauważyć, podobnie jak x_t^c, także π_t^c jest funkcją szoku podażowego u_t[80].

Jaka jest optymalna wartość parametru reakcji ω? Podstawiając ω z równania (5.67) do równania (5.68) i przekształcając, otrzymujemy:

$$\pi_t^c = \frac{\delta}{1-\beta\rho} x_t^c + \frac{1}{1-\beta\rho} u_t. \tag{5.69}$$

Równanie to pokazuje, że spadek wartości x_t^c o jeden procent wywołuje spadek π_t^c o $(\frac{\delta}{1-\beta\rho})$ procent. Przypomnijmy, że w warunkach polityki dyskrecjonalnej i przy oczekiwaniach egzogenicznych, na które bank centralny nie ma wpływu, równanie krzywej Phillipsa (5.52) wskazuje, że spadek luki podażowej x_t o jeden procent powoduje spadek inflacji tylko o δ procent. Lecz, ponieważ $\delta < \frac{\delta}{1-\beta\rho}$, to spadek luki podażowej o jeden procent w ramach polityki polegającej na podjętych zobowiązaniach wywołuje większy spadek inflacji niż w warunkach polityki dyskrecjonalnej. Oznacza to, że, podejmując wiarygodne zobowiązanie, bank centralny może obniżać inflację do pożądanego poziomu mniejszym kosztem w postaci utraconej produkcji.

Dzieje się tak dlatego, że wybór parametru ω wpływa nie tylko na bieżącą wartość luki podażowej x_t, lecz także na wszystkie przyszłe wartości x_{t+k}^c dla $k = 1, 2, 3, \ldots$[81]. Jednocześnie, ponieważ inflacja zależy od przyszłych wartości nad-

[80] Równanie (5.68) otrzymuje się, przekształcając funkcję krzywej Phillipsa (5.52) w jej iteracyjne rozwinięcie (5.54), podstawiając następnie (5.67) oraz wykorzystując wzór na sumę ciągu:

$$\pi_t^c = \delta x_t^c + \beta E_t \pi_{t+1}^c + u_t \tag{5.68a}$$

$$\pi_t^c = E_t \sum_{k=0}^{\infty} \beta^k \left(\delta x_{t+k}^c + u_{t+k}\right) \tag{5.68b}$$

$$\pi_t^c = E_t \sum_{k=0}^{\infty} \beta^k \left(-\delta\omega u_{t+k} + u_{t+k}\right). \tag{5.68c}$$

Uwzględniając, że $E_t x_t^c = -\omega u_t$ dla każdego t, można zapisać prawą stronę (5.68c) jako ciąg o postaci:

$$(-\delta\omega+1)u_t + \beta\rho(-\delta\omega+1)u_t + \beta^2\rho^2(-\delta\omega+1)u_t + \beta^3\rho^3(-\delta\omega+1)u_t + \ldots,$$

gdzie $(-\delta\omega+1)u_t$ jest wyrazem stałym. Wzorem na sumę tego ciągu jest (5.68).

[81] Przyjęcie przez bank centralny wiarygodnego zobowiązania, że $x_t^c = -\omega u_t$, oznacza, że bank będzie agresywnie reagował na szok podażowy także w przyszłości, czyli $E_t x_{t+k}^c = -\omega u_t$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wyżkowego popytu, zobowiązanie do stosowania twardej i agresywnej zasady reagowania pociga za sobą silniejszy spadek inflacji na jednostkę spadku produkcji niż w ramach polityki dyskrecjonalnej. Korzyść z podjęcia wiążącego zobowiązania przez bank centralny wynika z możliwości wpływania na skierowane w przyszłość (*forward-looking*) oczekiwania inflacyjne.

5.10. Konflikt celów w polityce pieniężnej i „trójkąt niemożności"

5.10.1. Parytet stóp procentowych

Przy swobodnym transgranicznym przepływie kapitału (pełna mobilność między krajem i zagranicą), stopy procentowe stają się w dużym stopniu niezależne od banku centralnego i są określone przez tzw. parytet stóp procentowych bez pokrycia (*uncovered interest parity* – UIP). Hipoteza UIP mówi o tym, że przy doskonałej mobilności kapitału, braku kosztów transakcyjnych oraz innych ograniczeń, a także przy założeniu, że inwestorzy są neutralni względem ryzyka, oczekiwany przychód z takich samych lokat w różnych walutach musi być taki sam:

$$(1+i)=\left(1+i^{*}\right)\frac{e^{p}}{e}, \tag{5.70}$$

gdzie i jest krótkoterminową stopą procentową w kraju, i^* oznacza krótkoterminową stopę procentową za granicą, e oznacza kurs walutowy (kurs waluty obcej wyrażony w walucie krajowej), a e^p jest przewidywanym kursem walutowym za rok. Równanie (5.70) zostaje spełnione dzięki pełnej swobodzie arbitrażu rynkowego. Przybliżeniem tej formuły jest następujące równanie[82]:

$$i=i^{*}+\frac{e^{p}-e}{e}, \tag{5.71}$$

[82] Logarytmując obie strony równania (5.71), otrzymujemy:

$$\ln(1+i)=\ln\left(1+i^{*}\right)+\ln\left[1+\frac{e^{p}-e}{e}\right].$$

Uwzględniając następnie, że $\ln(1+i) \approx i$ dla małych wartości i, otrzymujemy równanie (5.72).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

czyli:

$$i = i^{*} + \frac{de}{e}, \tag{5.72}$$

gdzie *de/e* jest stopą oczekiwanej zmiany kursu walutowego w ciągu roku[83].

Logika parytetu stóp procentowych jest prosta. Załóżmy, że występuje doskonała mobilność kapitału w skali międzynarodowej, inwestorzy pozostają neutralni względem ryzyka (*risk-neutral*)[84], a krajowe i zagraniczne aktywa są dobrymi substytutami[85]. Dany inwestor może ulokować swój kapitał w walucie krajowej po stopie procentowej i lub w walucie zagranicznej po stopie procentowej i^{*}. Aby zainwestować jednak za granicą, inwestor musi najpierw zamienić posiadane środki w walucie krajowej na walutę zagraniczną po kursie $1/e$. Po roku dochód z tej inwestycji wyniesie i^{*}, przy czym, jeśli inwestor chce odzyskać swój kapitał, będzie musiał zamienić posiadaną środki w walucie zagranicznej na walutę krajową po kursie e^{p}. Łączny przychód z tej inwestycji będzie więc równy sumie stopy i^{*} i względnej zmiany kursu $\frac{de}{e} * 100$.

Ogólnie rzecz biorąc, inwestor wybierze lokatę zapewniającą wyższą stopę przychodu, po uwzględnieniu ryzyka zmiany kursu. Branie pod uwagę oczekiwanych zmian kursu okazuje się ważne, ponieważ, nawet jeśli stopa procentowa w kraju jest wyższa niż za granicą, w ciągu roku może nastąpić deprecjacja kursu waluty krajowej, która obniży całkowitą dochodowość lokat krajowych. Gdy skala deprecjacji w ciągu roku jest większa niż różnica stóp procentowych, dochodowość lokaty w kraju staje się *per saldo* niższa niż za granicą. Gdy dochodowość lokat w kraju, po uwzględnieniu oczekiwanej zmiany kursu, jest wyższa niż za granicą, kapitał będzie napływał z zagranicy do kraju, powodując spadek stóp procentowych. Napływ ten będzie trwał do momentu, aż stopy przychodu w kraju i za

[83] Zob. Burda i Wyplosz (2000), s. 592–594 oraz Bénassy-Quéré i in. (2011), s. 295–298.

[84] Inwestor neutralny względem ryzyka traktuje jednakowo inwestycje, które zapewniają taki sam oczekiwany zysk, nawet jeśli jedna z nich obciążona jest ryzykiem. Na przykład traktuje jednakowo inwestycje, z których jedna zapewnia mu zysk 5 milionów EUR z prawdopodobieństwem 100%, a druga daje 50% szans na zysk 10 milionów EUR i 50% szans na brak zysku w ogóle. Z kolei inwestor unikający ryzyka (*risk-averse*) wybierze w tej sytuacji pierwszą opcję inwestycji, podczas gdy inwestor skłonny do ryzyka *(risk-seeking*) zdecyduje się na drugą opcję.

[85] Oznacza to, że ryzyko obciążające oba rodzaje aktywów jest jednakowe.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

granicą wyrównają się. Również odwrotnie, kapitał będzie z kraju odpływał, gdy stopa przychodu po uwzględnieniu zmiany kursu stanie się niższa niż za granicą.

Parytet stóp procentowych silnie ogranicza autonomię polityki pieniężnej. Załóżmy, że zagraniczna stopa procentowa wynosi $i^* = 2\%$, stopa krajowa $i = 4\%$, a inwestorzy oczekują osłabienia kursu waluty krajowej w ciągu roku o $\frac{e - e^p}{e} =$ $= 2\%$. Parytet stóp procentowych jest więc spełniony, bo dochodowość lokaty w walucie krajowej po uwzględnieniu oczekiwanej zmiany kursu jest taka sama jak dochodowość lokaty w walucie zagranicznej. Załóżmy teraz, że krajowy bank centralny obniża stopę procentową o 1 punkt procentowy, do poziomu $i = 3\%$, z kolei oczekiwany w przyszłości poziom kursu e^p nie zmienia się. Parytet zostaje naruszony, lewa strona równania (5.71) jest teraz niższa od prawej strony. W tej sytuacji kapitał będzie odpływał z kraju za granicę. W systemie kursu stałego oznacza to spadek podaży pieniądza i wzrost krajowej stopy procentowej i. Lewa strona równania (5.71) wzrośnie, co przywróci równowagę. W systemie kursu płynnego nastąpi deprecjacja kursu e o 1%, co obniży oczekiwania dalszej deprecjacji. Wzrost e przy niezmienionym e^p zmniejszy różnicę $(e - e^p)$ i w rezultacie prawa strona równania (5.71) spadnie, co przywróci równowagę.

Zauważmy, że w systemie płynnego kursu kurs nominalny reaguje natychmiast na zmianę stóp procentowych, a skala tej reakcji zależy od tego, jak długo krajowa stopa procentowa będzie odchylona w dół od poziomu parytetu. Jeśli rynki oczekują, że obniżona stopa utrzyma się przez okres dwóch lat, początkowa deprecjacja kursu wyniesie 2%. Ogólnie, wahania kursu walutowego będą z reguły większe niż zmiany stóp procentowych (zob. ramka 5.18). To właśnie duże zmiany kursu – zwłaszcza jego osłabienie – stanowią ograniczenie polityki pieniężnej w warunkach swobody przepływu kapitału. Banki centralne nie mogą dowolnie obniżać stóp, ponieważ grozi to masowym odpływem kapitału, którego zahamowanie wymagałoby bardzo głębokiej deprecjacji waluty, a to z kolei przyczyniłoby się do gwałtownego przyspieszenia inflacji. Z drugiej strony banki centralne nie mogą też dowolnie podnosić stóp procentowych, gdyż grozi to masowym napływem kapitału i niekontrolowanym wzrostem podaży pieniądza, a w konsekwencji także wzrostem inflacji[86].

[86] W pewnym zakresie można neutralizować pieniężne skutki napływu kapitału zagranicznego poprzez operacje sterylizacyjne, ale są one kosztowne i na dłuższą metę nieskuteczne (zob. następny rozdział).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Krótkookresowa reakcja kursu walutowego na zmiany stóp procentowych jest tym silniejsza, im dłuższy staje się czas inwestowania brany pod uwagę przez inwestorów. Uwzględniając okres dłuższy niż jeden rok, równanie parytetu UIP (5.71) można przekształcić następująco:

$$e_t = e_{t+T}^p + \sum_{\tau=0}^{T-1} E\left(i_{t+\tau} - i_{t+\tau}^*\right), \tag{5.73}$$

gdzie wyrażenie w nawiasie jest sumą oczekiwanych różnic pomiędzy krajową i zagraniczną stopą procentową w okresie od t do T. Powyższe równanie pozwala zauważyć, że bieżący kurs reaguje natychmiast na oczekiwane przyszłe zmiany stóp procentowych. Ponieważ horyzont analityczny inwestorów rozciąga się z reguły poza okres jednego roku, zmiany kursu walutowego będą bardziej gwałtowne niż zmiany stóp procentowych.

Ramka 5.18. Niepokryty i pokryty parytet stóp procentowych*

W rzeczywistości, równanie parytetu UIP rzadko jest spełnione. O braku równowagi świadczy chociażby fakt, że w praktyce nieustannie mamy do czynienia z przepływami kapitału między krajem a zagranicą. Na przykład w okresie od stycznia 2012 roku do lipca 2014 roku przepływy kapitału miedzy Polską i zagranicą, dotyczące tylko zakupu i sprzedaży dłużnych instrumentów finansowych, czyli instrumentów zależnych od stopy procentowej (obligacje, bony, lokaty, instrumenty pochodne), wynosiły przeciętnie 1733 mln USD w skali miesiąca, czyli około 80 mln USD w ciągu jednego dnia roboczego, przy czym był to zarówno napływ, jak i odpływ kapitału. Świadczy to o tym, że na rynku walutowym następuje nieustanne dostosowywanie przepływów kapitału do zmieniających się warunków i aktualizowanych prognoz. Ponieważ jednocześnie nie mamy jednak do czynienia z systematyczną tendencją do stałego odpływu lub stałego napływu kapitału, przepływy te można interpretować jako krótkookresowe odchylenia od stanu równowagi określonej równaniem niepokrytego parytetu stóp procentowych.

Te odchylenia od równowagi są spowodowane niepewnością co do przyszłości. Inwestorzy nie wiedzą, jaki będzie kurs walutowy za rok, za dwa lat lub jakikolwiek inny kurs w przyszłości, i dlatego przychód z inwestycji dokonanej w innych walutach jest zawsze niepewny. Obciąża go ryzyko kursowe. Inwestorzy mogą jednak

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zabezpieczyć się przed ryzykiem zmiany kursu walutowego, zawierając transakcję terminową na rynku walutowym. Spójrzmy ponownie na równanie niepokrytego parytetu (5.71), w którym występuje przewidywany – ale dziś nieznany – kurs walutowy za rok e^p. Zamiast przewidywać poziom kursu walutowego za rok, inwestor może już dziś kupić po kursie terminowym określoną ilość walut obcych z dostawą za rok, pozbywając się w ten sposób ryzyka kursowego. Taka transakcja nosi nazwę transakcji z pokryciem (*hedging*). Wówczas równanie parytetu stóp procentowych z pokryciem (*covered interest parity* – CIP) ma postać:

$$(1+i)=\left(1+i^*\right)\frac{e_T}{e}, \tag{5.74}$$

gdzie e_T oznacza kurs terminowy waluty obcej za rok. Przybliżeniem równania parytetu CIP (dla małych wartości i) jest wzór (5.75):

$$i=i^*+\frac{e_T-e}{e}. \tag{5.75}$$

Gdy $e_T > e$, rynek oczekuje deprecjacji waluty krajowej, co oznacza, że, dokonując inwestycji za granicą na okres jednego roku, inwestor osiągnie dochód składający się z odsetek i^* oraz – dodatkowo – dochód z tytułu aprecjacji waluty obcej wobec waluty krajowej. Drugi składnik prawej strony równania (5.75) nazywa się wówczas terminowym dyskontem. Z kolei gdy $e_T < e$, mamy do czynienia z terminową premią.

Rozpatrzmy następujący przykład liczbowy. Polski inwestor chce zainwestować środki w wysokości 4 mln PLN na okres jednego roku. Ma do wyboru zakup obligacji polskich denominowanych w złotych o oprocentowaniu rocznym 4% lub obligacji zagranicznych denominowanych w euro o oprocentowaniu rocznym 2%. Bieżący kurs euro wynosi 4 PLN/EUR, z kolei kurs terminowy z dostawą waluty za rok 4,20 PLN/EUR. Zakładając brak ograniczeń w przepływie kapitału i pomijając koszty transakcyjne, prawa strona równania (5.75) przyjmuje wartość następującą: 0,02 + (4,20–4,00)/4,00 = 0,07. Jak widać, równanie parytetu CIP nie jest spełnione, ponieważ prawa strona (7%) jest większa od lewej strony (4%). W tej sytuacji będzie następował odpływ kapitału, powodujący osłabienie bieżącego kursu złotego do poziomu około 4,1176 PLN/EUR, czyli do momentu, gdy kurs bieżący e osłabi się na tyle, że różnica ($e_T - e$) zostanie obniżona do poziomu 2%. Wówczas obie strony formuły parytetu wyrównają się i sytuacja powróci do równowagi.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Równanie bez pokrycia oraz równanie z pokryciem są do siebie bardzo zbliżone, ponieważ w rzeczywistości kurs terminowy jest właśnie kursem przewidywanym przez rynki na dany okres. Zasadnicza różnica polega na tym, że, stosując kurs terminowy, inwestor pozbywa się ryzyka kursowego, które przejmuje – za określoną opłatą – pośrednik finansowy, którym jest najczęściej bank komercyjny.

Jak wspomnieliśmy, w praktyce równanie parytetu jest spełnione jedynie w tendencji, ponieważ aktywa krajowe i zagraniczne nie są doskonałymi substytutami, a inwestorzy pozostają z reguły wrażliwi na ryzyko. Oznacza to, że w krótkim okresie krajowa stopa procentowa może odchylać się od prawej strony równania (5.71) ze względu na niepewność co do przyszłych wartości kursu. Im wyższa niepewność, tym wyższa wariancja zmiennej losowej e^p, co sprawia, że inwestorzy, unikający ryzyka, będą wówczas wymagać dodatkowej premii za ryzyko związane z inwestowaniem w obcej walucie. Zabezpieczanie się przed ryzykiem walutowym przy wykorzystaniu transakcji terminowej kosztuje; ponadto występują także koszty związane z wymianą waluty krajowej na walutę obcą i z powrotem. Uwzględniając te dwa dodatkowe składniki kosztów, równanie parytetu UIP można zapisać w następujący sposób:

$$i = i^* + \frac{e^p - e}{e} + k_T + k_W, \tag{5.76}$$

gdzie k_T jest procentowym kosztem transakcji terminowej w odniesieniu do wartości inwestycji, a k_W stanowi łączny procentowy koszt wymiany walut na początku i na końcu okresu inwestycji. Oznacza to, że krajowe stopy procentowe będą w rzeczywistości nieco wyższe niż te określone równaniem (5.72).

5.10.2. Konflikt celów i „trójkąt niemożności" (*impossible trinity*)

W poprzednim punkcie pokazaliśmy, że w warunkach swobody przepływu kapitału bank centralny ma ograniczoną autonomię w ustalaniu stóp procentowych. Ponieważ krajowe rynkowe stopy procentowe kształtują się w okolicach parytetu stóp procentowych zgodnie z równaniami (5.46) i (5.72), bank centralny nie może wyznaczyć oficjalnej stopy procentowej na poziomie znacząco wyższym lub znacząco niższym od stopy wyznaczonej przez równanie parytetu, ponieważ doszłoby wówczas do systematycznego napływu lub odpływu kapitału. Spowoduje

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

to albo odpływ rezerw walutowych (w systemie kursu stałego), albo spadek kursu waluty krajowej (w systemie kursu płynnego). Na dłuższą metę jest to sytuacja nie do utrzymania, ponieważ rezerwy walutowe zawsze są ograniczone, a nadmierna deprecjacja waluty grozi wysoką inflacją i kryzysem walutowym.

Z drugiej jednak strony wiemy już, że, ustalając oficjalne stopy procentowe, bank centralny w dążeniu do zapewnienia stabilności cen kieruje się różnicą między przewidywaną inflacją a celem inflacyjnym w ramach strategii BCI (ewentualnie zmianami podaży pieniądza, jeśli stosuje regułę celu pośredniego). W szczególności funkcja reakcji banku centralnego w postaci (5.45) pokazuje, że bank centralny ustala oficjalną krótkookresową stopę procentową w taki sposób, aby zminimalizować odchylenia rzeczywistego PKB od PKB potencjalnego i odchylenia prognozowanej inflacji od celu inflacyjnego.

Oznacza to, że ta sama zmienna zależna – oficjalna krótkookresowa stopa procentowa banku centralnego – jest określona przez dwa różne równania, w których występują dwie różne grupy zmiennych niezależnych. W przypadku równania parytetu stóp procentowych krajowa oficjalna stopa procentowa zależy od stopy procentowej za granicą oraz od prognozowanej zmiany kursu walutowego, z kolei w przypadku równania funkcji reakcji ta sama krajowa stopa procentowa zależy od stopy procentowej za granicą oraz od odchyleń PKB od potencjału i prognozowanej inflacji od celu. Nie istnieje żaden powód *ex ante*, aby stopy procentowe wynikające z równania parytetu i z funkcji reakcji były sobie równe, a wręcz odwrotnie – będą się z reguły różnić od siebie. W tej sytuacji bank centralny staje wobec poważnego dylematu – które równanie uznać za decydujące dla wyliczenia właściwej stopy procentowej?

Sytuacja ta została przedstawiona na rysunku 5.9. Po prawej stronie ukazano funkcję parytetu stóp procentowych, czyli kształtowanie się krajowej stopy procentowej i (oś pionowa) jako funkcji oczekiwanej zmiany kursu e^p (oś pozioma). Im wyższa staje się wartość e^p, tym wyższa będzie oczekiwana deprecjacja waluty krajowej i tym wyższa musi być stopa procentowa, aby powstrzymać odpływ kapitału. W lewej części pokazano funkcję reakcji, czyli kształtowanie się stopy procentowej i (oś pionowa) jako funkcji różnicy między prognozowaną inflacją a celem inflacyjnym (oś pozioma)[87]. Im wyższy staje się cel inflacyjny π, tym mniejsze będzie odchylenie prognozowanej inflacji p_e od celu i tym niższa może być stopa procentowa

[87] W celu uproszczenia pominięto różnicę między PKB rzeczywistym i PKB potencjalnym oraz różnicę w stopach inflacji w kraju i za granicą.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

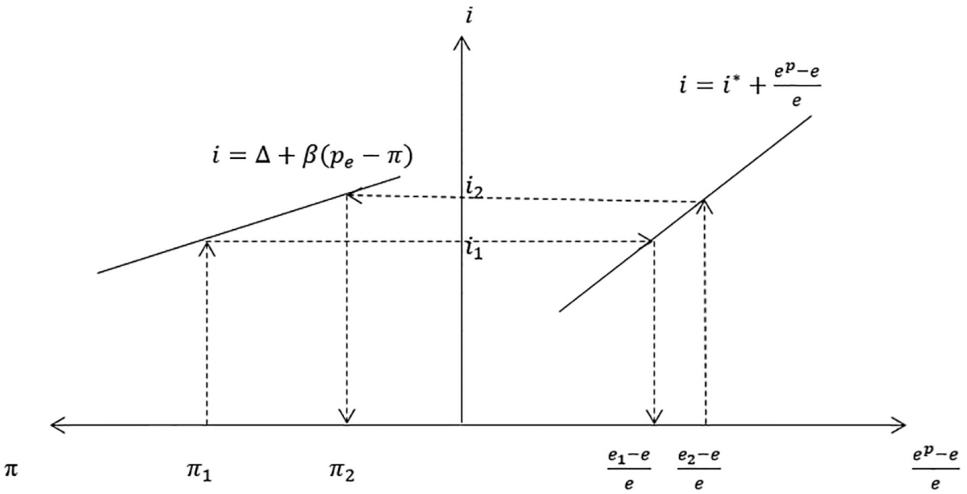
spójna z przyjętym celem. Jeśli bank centralny ustali cel inflacyjny na poziomie π_1, cel ten „narzuci" wymaganą stopę procentową i_1, a w rezultacie kurs walutowy ukształtuje się samorzutnie na poziomie e_1, spójnym ze stopą procentową i_1. Jeśli bank centralny pragnie jednak kontrolować kurs walutowy, na przykład ustalając stały kurs e_2, to wyznaczy on spójną z tym kursem stopę i_2 oraz spójny z tą stopą cel inflacyjny π_2. Jak widać, wybór celu inflacyjnego przesądza o poziomie stopy procentowej i w konsekwencji o poziomie kursu walutowego, a wybór kursu walutowego przesądza o poziomie stopy procentowej i w konsekwencji o poziomie inflacji. Mamy więc do czynienia z konfliktem celów – nie ma bowiem możliwości, aby jednocześnie i niezależnie ustalać wszystkie trzy zmienne: poziom inflacji, poziom stopy procentowej i poziom kursu walutowego.

Ta dwukierunkowa zależność nosi nazwę „trylematu polityki pieniężnej", „trójkąta niemożności" lub „niemożliwej trójcy" (*impossible trinity*). W warunkach gospodarki otwartej na przypływy kapitałowe nie można jednocześnie ustalać wszystkich trzech zmiennych. Możliwe staje się jedynie ustalanie dwóch spośród nich, podczas gdy trzecia dostosowuje się sama. Bank centralny może więc wyznaczyć cel inflacyjny i stosownie do tego ustalić spójną z tym celem stopę procentową (zgodnie z równaniem (5.45)), lecz wówczas kurs walutowy ukształtuje się na poziomie równowagi parytetu niezależnie od decyzji banku centralnego. Alternatywnie bank centralny może wyznaczyć jakiś pożądany poziom kursu walutowego (lub tempo jego zmian) i spójny z tym kursem poziom stopy procentowej, lecz wówczas inflacja w średnim okresie ukształtuje się autonomicznie na poziomie określonym funkcją reakcji, niezależnie od decyzji banku centralnego.

Trylemat polityki pieniężnej wyjaśnia, dlaczego w krajach, w których stosuje się strategię BCI (lub regułę celu pośredniego w postaci podaży pieniądza), obowiązują systemy płynnego kursu walutowego. Po prostu jakikolwiek system stałego (kontrolowanego) kursu byłby niespójny z przyjętą strategią polityki pieniężnej i uniemożliwiałby w dłuższym czasie kontrolę inflacji. Dlatego też system płynnego kursu obowiązuje między innymi w takich krajach jak USA, Wielka Brytania, Japonia, Kanada, Australia, a także Polska i strefa euro. Trylemat wyjaśnia także, dlaczego w krajach, które stosują stałe kursy walutowe, dochodzi do trwałych odchyleń inflacji od poziomu obserwowanego w krajach sąsiednich. Na przykład w krajach bałtyckich (Estonia, Litwa i Łotwa), które stosowały sztywne kursy walutowe, inflacja w okresie szybkiego wzrostu gospodarczego (2003–2008) była wyższa o kilka punktów procentowych niż wśród pozostałych państw Europy Środkowo-Wschodniej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 5.9. Konflikt celów w polityce pieniężnej



Źródło: opracowanie własne (Rosati 2000).

W praktyce okazuje się, że ograniczenia płynące z trylematu nie są bezwzględnie sztywne. Doświadczenia wielu krajów dowodzą, że udaje im się utrzymywać w pewnym zakresie autonomię polityki pieniężnej nawet w warunkach swobody przepływu kapitału. Niektóre kraje stosują także różne formy pośredniej lub ukrytej kontroli kursu walutowego. Świadczy o tym utrzymywanie się dość trwałych odchyleń od parytetu stóp procentowych. Jest to możliwe, ponieważ prognozy kursu są zawsze obarczone sporym ryzykiem błędu, a przepływy kapitału wiążą się w praktyce z dodatkowymi kosztami transakcyjnymi. Dlatego też trylemat polityki pieniężnej nie jest bezwzględnie wiążącym ograniczeniem w przypadku krótkookresowej polityki pieniężnej. Oczywiście, najbardziej radykalnym sposobem rozwiązania trylematu jest po prostu eliminacja swobody transgranicznego przepływu kapitału. Pozwala to wprawdzie odzyskać pełną autonomię polityki pieniężnej, lecz wiąże się jednocześnie z utratą korzyści z integracji finansowej i brakiem możliwości korzystania z zagranicznych zasobów kapitału.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.11. Zależność między polityką pieniężną a polityką fiskalną

5.11.1. Czy polityka pieniężna i polityka fiskalna są od siebie zależne?

Poglądy na temat zależności występujących między polityką fiskalną i polityką pieniężną ulegały istotnym zmianom na przestrzeni ostatniego stulecia. W tradycyjnym systemie waluty złotej obie polityki były traktowane jako niezależne od siebie – polityka fiskalna realizowała ograniczone funkcje alokacyjne i dystrybucyjne, a polityka pieniężna dążyła do zapewnienia stabilnej wartości pieniądza krajowego w relacji do kruszcu. Uważano, że polityka fiskalna może mieć wpływ na spadek wartości pieniądza oraz inflację tylko w przypadku finansowania deficytu budżetowego dodrukiem pieniądza (monetyzacja deficytu). Gdy jednak deficyt jest finansowany emisją długu (sprzedaż obligacji), podaż pieniądza i ceny nie ulegają zmianie i obie polityki pozostają zasadniczo niezależne (Wojtyna 1998, Marszałek 2003). Wydarzenia lat 30. XX wieku, w tym odejście od systemu *gold standard* i rewolucja keynesowska, otworzyły drogę do wykorzystania polityki pieniężnej jako instrumentu polityki stabilizacyjnej. Doprowadziło to do uzależnienia polityki pieniężnej od polityki fiskalnej w ramach prowadzonej przez rządy antycyklicznej polityki makroekonomicznej. Banki centralne pożyczały pieniądze rządom w celu finansowania deficytów budżetowych w okresach dekoniunktury.

Skutkiem podporządkowania bankierów centralnych ministrom finansów była jednak rosnąca inflacja. Monetyzacja deficytów i tolerowanie wyższej inflacji w dążeniu do ograniczenia bezrobocia doprowadziły do utrwalenia oczekiwań inflacyjnych i przyspieszenia wzrostu cen. W szeregu prac opublikowanych w latach 60. i 70. Milton Friedman dowodził, że obarczanie polityki pieniężnej dodatkowymi zadaniami, takimi jak walka z bezrobociem i przyspieszanie wzrostu gospodarczego jest nieuzasadnione i szkodliwe, a także że polityka pieniężna powinna zająć się tym, na co ma rzeczywisty wpływ, czyli walką z inflacją. Friedman krytykował próby pobudzania gospodarki przy pomocy polityki pieniężnej, wskazując, że zakłócają one sygnały płynące z rynku, a na dłuższą metę powodują jedynie wzrost inflacji. Dlatego też uważał, że banki centralne powinny być w swej polityce niezależne od rządów, a nominalna podaż pieniądza w gospodarce powinna rosnąć w stałym, umiarkowanym tempie, niezależnie od rzeczywistego stanu koniunktury gospodarczej. Miało to stabilizować oczekiwania i gwarantować, że w dłuższym okresie inflacja zawsze pozostanie pod kontrolą.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Stanowisko Friedmana zostało zakwestionowane przez dwóch znanych amerykańskich ekonomistów Sargenta i Wallace'a, i to od razu na dwóch frontach. Po pierwsze, w swojej słynnej pracy na temat „nieprzyjemnej arytmetyki monetarnej" Sargent i Wallace (1981) wykazali, że oddzielenie polityki pieniężnej od polityki fiskalnej wcale nie gwarantuje tej pierwszej skutecznej kontroli nad inflacją. Ekspansywna polityka fiskalna rządu i systematycznie rosnący dług publiczny mogą w pewnym momencie zmusić bank centralny do monetyzacji długu i dopuszczenia wysokiej inflacji, tak aby uniknąć bankructwa państwa. Sargent i Wallace określili tę zależność mianem „nieprzyjemnej arytmetyki monetarnej" – omawiamy ją w dalszej części rozdziału. Pokazuje ona, że w długim okresie występuje zależność między polityką fiskalną a polityka pieniężną.

Po drugie, w innej, równie ważnej pracy ci sami autorzy dowodzą, że tzw. reguła Friedmana, polegająca na utrzymywaniu przez bank centralny stałej stopy wzrostu podaży pieniądza niezależnie od okoliczności jest suboptymalna (Sargent i Wallace 1976). W warunkach istnienia oczekiwań inflacyjnych (racjonalnych lub adaptacyjnych) bardziej efektywną regułą polityki pieniężnej jest reguła uwzględniająca sprzężenia zwrotne między zmiennymi wpływającymi na inflację, wykorzystująca najnowsze informacje napływające z gospodarki, w tym informacje dotyczące realizowanej przez rząd polityki fiskalnej. Wniosek ten otworzył drogę do stopniowego odchodzenia od strategii stosowania pośrednich celów w polityce pieniężnej i zastępowania ich strategią bezpośredniego celu inflacyjnego.

Rozwój badań nad polityką pieniężną w latach 70. i 80. doprowadził do pojawienia się dwóch ważnych innowacji. Po pierwsze, kierując się ustaleniami literatury na temat racjonalnych oczekiwań, banki centralne w większości państw skupiły się na walce z inflacją, przyjmując stabilność cen jako główny, a niekiedy jedyny, cel polityki pieniężnej. Po drugie, pod wpływem badań dotyczących zjawiska niespójności czasowej polityki uznano, że niezależność banków centralnych jest niezbędnym warunkiem skuteczności polityki pieniężnej. Oba te nowe zjawiska doprowadziły do ponownej separacji polityki pieniężnej i polityki fiskalnej. Na wyraźne rozdzielenie obszarów działania (polityka pieniężna oddziałuje na zmienne nominalne, a polityka fiskalna na zmienne realne) nałożyła się dodatkowo instytucjonalna emancypacja banków centralnych. W tych warunkach tak długo, jak rząd nie prowadzi nadmiernie ekspansywnej polityki fiskalnej grożącej bankructwem kraju, bank centralny nie otrzymuje żadnego formalnego bodźca, aby koordynować swą politykę z polityką fiskalną.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Pozostaje jednak zdrowy rozsądek. Niezależnie od braku instytucjonalnych kanałów współpracy między polityką fiskalną i polityką pieniężną istnieją wielorakie powiązania ekonomiczne, które mają charakter efektów zewnętrznych. Najbardziej znanym przykładem jest wspomniana „nieprzyjemna arytmetyka monetarna", lecz występują także inne następstwa. Na przykład ekspansywna polityka fiskalna finansowana napływem kapitału z zagranicy może być źródłem inflacji i w ten sposób będzie zmuszać bank centralny do zaostrzania polityki pieniężnej. W tej sytuacji realizacja przez politykę fiskalną celu w postaci pobudzania koniunktury przynosi określone skutki dla polityki pieniężnej. Jak poucza „nieprzyjemna arytmetyka monetarna", w skrajnym przypadku bank centralny może okazać się niezdolny do zahamowania inflacji wywołanej przez nieodpowiedzialną politykę fiskalną. Mówimy wówczas o „dominacji" polityki fiskalnej nad polityką pieniężną. Z drugiej strony restrykcyjna polityka pieniężna może wpływać na wzrost rynkowych stóp procentowych, co zwiększa koszty obsługi długu publicznego i ogranicza pozostałe wydatki budżetowe lub zmusza do podnoszenia podatków. W tym przypadku to realizacja przez politykę pieniężną celu w postaci ograniczenia inflacji niesie określone skutki dla polityki fiskalnej. Mówimy wówczas o „dominacji" polityki pieniężnej nad polityką fiskalną. Argument na rzecz separacji obu polityk jest zatem uzasadniony tylko wtedy, gdy wzajemne efekty zewnętrzne między dwoma modelami polityki są nieznaczne. W przeciwnym razie występuje obiektywna konieczność koordynacji obu polityk (Niemann i von Hagen 2008)[88].

W praktyce rzadko mamy jednak do czynienia z taką koordynacją. Jej zwolennikami są o wiele częściej przedstawiciele strony fiskalnej, czyli rządu. Często słyszymy apele przedstawicieli rządu skierowane do banków centralnych o poluzowanie polityki pieniężnej. Motyw stanowi w tym przypadku dążenie do stymulowania gospodarki poprzez wzrost zagregowanego popytu i do ograniczania bezrobocia. Z kolei bankierzy centralni są z reguły znacznie bardziej sceptyczni, a niektórzy w ogóle wykluczają jakąkolwiek aktywną koordynację, obawiając się ograniczenia własnej niezależności[89].

[88] Brak koordynacji może doprowadzić do nieskuteczności obu odmian polityki. Wykorzystując aparat teorii gier, Dixit i Lambertini (2003) wykazują, że taka sytuacja ma charakter gry niekooperacyjnej, w której gracze podejmują decyzje, uwzględniając sposób reakcji drugiego gracza. Rozwiązaniem takiej niekooperacyjnej gry jest równowaga w sensie Nasha, w przypadku której polityka pieniężna jest nadmiernie restrykcyjna, a polityka fiskalna nadmiernie ekspansywna.

[89] Na przykład pierwszy prezes Europejskiego Banku Centralnego Wim Duisenberg uważał, że polityka fiskalna jest jednym z egzogenicznych czynników określających warunki prowadzenia polityki pieniężnej i nie znajdował żadnego uzasadnienia dla koordynacji obu polityk (zob. Duisenberg 2003).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.11.2. „Nieprzyjemna arytmetyka monetarna"

Jak już wspomnieliśmy, w latach 70. i 80. teoria oraz praktyka polityki pieniężnej uległy dwóm istotnym zmianom. Pierwsza z nich polegała na uznaniu, że głównym celem polityki pieniężnej powinno być utrzymanie stabilności cen, a druga na zagwarantowaniu niezależności banków centralnych, zwłaszcza w sprawach związanych z prowadzeniem polityki pieniężnej. Obie te zmiany doprowadziły do upowszechnienia się przekonania, że polityka pieniężna stała się skutecznym instrumentem kontrolowania inflacji, a zarazem pozostała niezależna od polityki fiskalnej.

Przekonanie to podważyli w 1981 roku dwaj amerykańscy ekonomiści Thomas Sargent i Neil Wallace. W swej słynnej pracy na temat „nieprzyjemnej arytmetyki monetarnej" („*Some Unpleasant Monetarist Arithmetic*") wykazali, że – przy pewnych założeniach i w pewnych szczególnych okolicznosciach – sama polityka pieniężna może nie być zdolna do trwałej kontroli poziomu inflacji. W szczególności, gdy rząd finansuje deficyt budżetowy jednocześnie na dwa sposoby – sprzedając obligacje sektorowi prywatnemu i zwiększając podaż pieniądza rezerwowego – oraz gdy w sposób niezależny określa ścieżkę deficytu w czasie, bank centralny jest zmuszony dostosować podaż pieniądza do wielkości deficytu przy danym popycie na obligacje rządowe zgłaszanym przez sektor prywatny. W ramach takiego mechanizmu koordynowania polityki makroekonomicznej, w którym to polityka fiskalna dominuje nad polityką pieniężną, bank centralny może nie być w stanie utrzymać podaży pieniądza pod kontrolą i w efekcie nie uda mu się utrzymać pożądanej stabilności cen (Woodford 2001, Jakóbik 2008).

Wykorzystując model makroekonomiczny z wielookresowym ograniczeniem budżetowym, Sargent i Wallace wykazali, że gdy polityka fiskalna jest dostatecznie ekspansywna, a przy tym przynajmniej część deficytu budżetowego jest finansowana przyrostem pieniądza rezerwowego, bank centralny może stracić kontrolę nad inflacją. Ograniczenie zdolności banku centralnego do kontrolowania inflacji wynika z prostej arytmetyki opisującej zależność między wielkością długu publicznego, poziomem stóp procentowych, tempem wzrostu gospodarczego i tempem wzrostu podaży pieniądza rezerwowego. Załóżmy, że rząd systematycznie zwiększa deficyt, bank centralny, dążąc do zachowania stabilności cen, stara się zaś utrzymać stałe i umiarkowane tempo wzrostu podaży pieniądza. Systematyczny wzrost deficytu wymaga w tych warunkach systematycznego wzrostu emisji obligacji, co wywołuje wzrost stóp procentowych powyżej tempa wzrostu gospodarczego,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

a w konsekwencji prowadzi do wzrostu zadłużenia. Dług nie może jednak rosnąć w nieskończoność, ponieważ popyt sektora prywatnego na obligacje jest ograniczony w relacji do wielkości produkcji w danym kraju. W takiej sytuacji – prędzej czy później – musi dojść do zwiększenia bazy monetarnej w celu sfinansowania zapadających obligacji i odsetek. To z kolei wywołuje ostatecznie wzrost inflacji (zob. ramka 5.19) powyżej poziomu uznanego za pożądany przez bank centralny.

Ramka 5.19. „Nieprzyjemna arytmetyka monetarna" według Sargenta i Wallace'a*

Sargent i Wallace (1981) rozpatrują przypadek gospodarki spełniającej założenia „monetarystyczne", gdzie tempo wzrostu produkcji oraz tempo wzrostu ludności jest jednakowe i wynosi n, realna stopa procentowa na obligacjach rządowych osiąga poziom r, przy czym $r > n$, a funkcja popytu na pieniądz jest zgodna z ilościową teorią pieniądza przy stałej wartości szybkości obiegu $1/h$. W takiej gospodarce polityka pieniężna nie wpływa na wielkości realne, w tym na realne tempo wzrostu i na realną stopę procentową. Polityka fiskalna jest określona przez sekwencję *realnych* deficytów budżetowych w kolejnych okresach, $D_1, D_2, \ldots, D_t$, a polityka pieniężna przez sekwencję wielkości obiegu pieniądza rezerwowego w kolejnych okresach, $M_1, M_2, \ldots, M_t$. Zakładając, że obligacje rządowe zapadają po upływie jednego okresu, ograniczenie budżetowe skonsolidowanego sektora rządowego (rządu i banku centralnego) można zapisać następująco:

$$D_t = \frac{\left[M_t - M_{t-1}\right]}{p_t} + \left[B_t - B_{t-1}\left(1 + r_{t-1}\right)\right], \tag{5.77}$$

gdzie p_t jest poziomem cen w okresie t, $B_{t-1}\left(1 + r_{t-1}\right)$ stanowi realną wielkość zasobu obligacji w rękach sektora prywatnego w okresie t, a r_{t-1} oznacza stopę oprocentowania obligacji pomiędzy okresem t–1 i t. Równanie (5.77) stanowi, że deficyt budżetowy musi być sfinansowany w określonych proporcjach przez emisję nowego pieniądza rezerwowego i przez przyrost długu. Oznaczając wielkość ludności przez N_t, gdzie $N_t = (1+n)N_{t-1}$, i dzieląc równanie (5.77) obustronnie przez N_t, otrzymujemy równanie ograniczenia budżetowego w przeliczeniu na jednego mieszkańca:

$$\frac{B_t}{N_t} = \frac{B_{t-1}\left(1 + r_{t-1}\right)}{N_{t-1}(1+n)} + \frac{D_t}{N_t} + \frac{M_t - M_{t-1}}{N_t p_t}. \tag{5.78}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Sargent i Wallace przyjmują, że polityka fiskalna dominuje nad polityką pieniężną, co oznacza, że ścieżka deficytów D_t jest określona egzogenicznie dla wszystkich t. Wówczas alternatywne ścieżki polityki pieniężnej polegają na wybraniu pewnej stałej stopy wzrostu podaży pieniądza rezerwowego θ dla M_t dla $t = 2, 3, \ldots, T$, gdzie T jest pewnym okresem w przyszłości, w którym realny dług publiczny na mieszkańca przestaje rosnąć ze względu na nasycenie rynku obligacjami i zanik popytu na obligacje. Poza okresem T, przyrost podaży pieniądza rezerwowego staje się jedynym źródłem finansowania deficytu. Jeśli początkowy stan podaży pieniądza M_1 jest dany, możemy zapisać, że w okresach $t = 2, 3, \ldots, T$, podaż pieniądza rośnie zgodnie z formułą:

$$M_t = (1+\theta) M_{t-1}. \tag{5.79}$$

Na podstawie równania ilościowego pieniądza poziom cen w okresie t jest równy:

$$p_t = \left(\frac{1}{h}\right)\left(\frac{M_t}{N_t}\right). \tag{5.80}$$

W związku z czym stopa inflacji w okresie $t = 2, 3, \ldots, T$ wynosi:

$$\frac{p_t}{p_{t-1}} = (M_t / N_t) / \left(\frac{M_{t-1}}{N_{t-1}}\right) = (1+\theta)/(1+n). \tag{5.81}$$

Równanie (5.81) pokazuje, że jeśli bank centralny wybierze pewną politykę pieniężną o określonych parametrach θ i T, to – przy danym n – decyzja ta przesądza również o określonej stopie inflacji w okresie $t = 2, 3, \ldots, T$.

Sargent i Wallace wykazują jednak, że przy tych założeniach inflacja przed okresem T określa poziom inflacji także po okresie T. Oznaczmy poziom długu rządowego na mieszkańca osiągnięty w okresie T przy danej polityce pieniężnej opisanej parametrem θ jako b_T^θ. Zauważmy, że im niższa wartość θ, tym bardziej restrykcyjna jest polityka pieniężna, i tym większy udział przyrostu długu w finansowaniu deficytu. Aby określić zależność inflacji poza okresem T od wielkości b_T^θ, wykorzystamy równanie (5.80) i równość $\frac{B_t}{N_t} = \frac{B_{t-1}}{N_{t-1}} = b_T^\theta$. Wówczas równanie (5.78) można – po stosownych przekształceniach – zapisać w następujący sposób:

$$b_T^\theta = \left[\frac{1-r_{t-1)}}{1+n}\right] b_T^\theta + \frac{D_t}{N_t} - h\left[1 - \frac{N_{t-1}P_{t-1}}{N_t P_t}\right],$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

czyli:

$$1-\frac{\left(\frac{1}{1+n}\right)p_{t-1}}{p_t}=\frac{b_T^{\theta}}{h}\left[(r_{t-1}-n)/(1+n)\right]+\left(D_t/N_t\right). \tag{5.82}$$

Jeśli – zgodnie z założeniem – $r_{t-1}>n$, prawa strona równania (5.82) rośnie wraz ze wzrostem b_T^{θ}. To z kolei oznacza, że wraz ze wzrostem b_T^{θ} rośnie także lewa strona równania, a zatem – przy danym n – spada wartość $\left(\frac{p_{t-1}}{p_t}\right)$, czyli rośnie inflacja $\left(\frac{p_t}{p_{t-1}}\right)$. Im wyższy więc jest poziom zadłużenia na mieszkańca w momencie T, tym większa będzie inflacja po okresie T.

Pozostaje teraz znaleźć zależność między poziomem zadłużenia na mieszkańca w okresie T, b_T^{θ} a wartością θ, czyli kluczowym parametrem polityki pieniężnej w okresie przed momentem T. Zacznijmy od ustalenia wartości początkowej b_1. Na podstawie równania (5.78) otrzymujemy, dla $t = 1$:

$$b_1=\left[\frac{B_0^n}{N_1p_1}\right]+\left[\frac{M_1-M_0}{N_1p_1}\right]+\left[\frac{D_1}{N_1}\right], \tag{5.83}$$

gdzie B_0^n oznacza nominalną wartość w okresie $t=1$ długu wyemitowanego w okresie $t=0$, a B_0^n/p_1 stanowi przybliżenie wartości $B_0(1+r_0)$[90]. W kolejnym kroku znajdujemy wartości $b_2^{\theta}, b_3^{\theta}, \ldots, b_T^{\theta}$. W tym celu ponownie wykorzystujemy równanie (5.80), zapisując dla $t = 2, 3, \ldots T$:

$$b_t=\left[\frac{1+r_{t-1}}{1+n}\right]b_{t-1}+\left(\frac{D_t}{N_t}\right)-\left(\frac{h\theta}{1+\theta}\right). \tag{5.84}$$

Rozwijając iteracyjnie równanie (5.84) na przyszłe okresy, otrzymujemy ostatecznie dla wartości t poza okresem T:

$$b_t^{\theta}=\varphi_t^1b_1+\left(\sum_{s=2}^{t}\varphi_t^s\left(\frac{D_s}{N_s}\right)\right)-\left(\frac{h\theta}{1+\theta}\right)\sum_{s}^{t}\varphi_t^s, \tag{5.85}$$

[90] W podstawieniu chodzi o to, aby nie przesądzać na temat początkowej wysokości stopy procentowej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gdzie $\varphi_t^{s=t} = 1$, dla $t > s$ natomiast:

$$\varphi_t^s = \left(\prod_{j=s}^{t-1}\left(1+r_j\right)\right) / \left(1+n\right)^{t-s}. \tag{5.86}$$

Ponieważ wartość wyrażenia $\left(\frac{h\theta}{1+\theta}\right)$ po prawej stronie równania (5.85) spada w wartościach bezwzględnych wraz ze spadkiem wartości θ, to ze względu na stojący przed wyrażeniem znak „minus", prawa strona równania *rośnie* w miarę wzrostu wartości θ. Oznacza to, że bardziej restrykcyjna polityka pieniężna w okresie $t = 1, 2, 3, \ldots, T$, prowadzi do wzrostu wartości b_T^θ, a to z kolei – na podstawie równania (5.82) – wywołuje wyższą inflację poza okresem T. Intuicyjnie otrzymany wynik jest łatwy do zinterpretowania. Im ostrzejsza polityka pieniężna w czasie przed okresem T, tym większy przyrost zadłużenia z tytułu emisji obligacji w celu sfinansowania ustalonej egzogenicznie ścieżki deficytu. Z kolei im wyższy poziom zadłużenia w okresie T, tym większa musi być podaż pieniądza rezerwowego po okresie T, a zatem tym wyższa będzie inflacja. „Nieprzyjemna arytmetyka monetarna" pokazuje, że bardziej restrykcyjna polityka pieniężna – a zatem niższa inflacja – obecnie, może prowadzić w określonych okolicznościach do wyższej inflacji w przyszłości.

„Nieprzyjemna arytmetyka monetarna" odzwierciedla pieniężne skutki deficytu budżetowego w ujęciu wielookresowym, gdy jest on finansowany – choćby częściowo – kreacją pieniądza. Uświadamia to, że osiągnięcie trwałej stabilności cen w dłuższym okresie jest mało prawdopodobne w warunkach ekspansywnej polityki budżetowej, zakładającej nieodpowiedzialne zwiększanie długu. Jak wcześniej zaznaczyliśmy, taka sytuacja może wystąpić w warunkach dominacji polityki fiskalnej nad polityką pieniężną, co z kolei ma miejsce wówczas, gdy bank centralny nie jest instytucją niezależną bądź gdy dany kraj zmuszony jest ponosić bardzo duże wydatki z budżetu (na przykład w czasie wojny). Praca Sargenta i Wallace'a dostarczyła ekonomistom i politykom silnych argumentów na rzecz wprowadzenia zakazu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

finansowania deficytu budżetowego kreacją pieniądza (tzw. monetyzacja deficytu) i wzmocnienia niezależności banków centralnych[91].

5.12. Polityka pieniężna w Polsce

5.12.1. Lata 1990–1998: reguła celu kursowego i strategia „eklektyczna"

W pierwszych latach po transformacji systemowej polityka pieniężna w Polsce była podporządkowana walce z inflacją i ustabilizowaniu wartości złotego. Zasady formułowania tych celów oraz sposoby ich realizacji i kryteria oceny podlegały jednak wyraźnej ewolucji. Ponadto w niektórych latach polityka pieniężna była nakierowana jednocześnie na inne cele, niekoniecznie spójne z celem głównym.

W 1990 rok polska gospodarka weszła z ponad 70-procentową inflacją w skali miesiąca, z powszechnymi niedoborami na rynkach produktów i wysokim zadłużeniem. Polityka pieniężna była podporządkowana założeniom programu stabilizacyjnego (zwanego popularnie terapią szokową), w którym głównym priorytetem była stabilizacja makroekonomiczna i finansowa, w tym poskromienie inflacji. W ramach tego programu Narodowy Bank Polski zobowiązał się do: utrzymywania stóp procentowych na realnie dodatnim poziomie, radykalnego ograniczenia zakresu i skali preferencyjnych kredytów dla przedsiębiorstw państwowych oraz zaniechania bezpośredniego finansowania deficytu budżetowego emisją pieniądza. Działania te miały utrzymać wzrost podaży pieniądza w ramach określonego limitu, który był celem pośrednim polityki pieniężnej. Kluczowy element programu stabilizacyjnego stanowiła głęboka dewaluacja oficjalnego kursu złotego (o ponad 40%) i ustalenie sztywnego kursu walutowego na poziomie 9500 PLN/USD. Stały kurs miał pełnić w programie funkcję „kotwicy nominalnej" poprzez swój stabilizujący wpływ na ceny i oczekiwania inflacyjne. Wobec bardzo skromnego stanu oficjalnych rezerw walutowych wiarygodność tego rozwiązania miały zapewnić wysokie krajowe nominalne stopy procentowe oraz utworzony przez kraje OECD i udostępniony Polsce fundusz stabilizacyjny w wysokości 1 mld USD.

[91] Niektórzy ekonomiści wskazują, że niezależność banku centralnego może być przeszkodą do ściślejszej koordynacji między polityką fiskalna i polityka pieniężną, ponieważ bank centralny może stać się „więźniem" własnej niezależności (Cargill 2001, Marszałek 2003). Badania w zasadzie nie potwierdzają jednak tych obaw.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Instytucjonalne ramy prowadzenia polityki pieniężnej wyznaczała ustawa o Narodowym Banku Polskim z 1989 roku. Zgodnie z jej postanowieniami, celem działalności NBP miało być w szczególności „umacnianie wartości pieniądza polskiego". W zakresie swojej działalności NBP powinien uczestniczyć w „kształtowaniu realizacji polityki gospodarczej państwa oraz w jej realizacji, kierując się ustawami i uchwałami Sejmu", spełniać założenia polityki pieniężno-kredytowej uchwalane przez Sejm i współpracować z Ministrem Finansów. Prezes NBP miał być powoływany i odwoływany przez Sejm na wniosek Prezesa Rady Ministrów[92]. Przepisy te w żadnej mierze nie zapewniały polskiemu bankowi centralnemu niezależności politycznej, funkcjonalnej, operacyjnej i finansowej, co zresztą było typowe dla anglo-francuskiego modelu bankowości centralnej, obowiązującego w owych latach w większości państw europejskich.

Mimo że podstawowym celem polityki pieniężnej było obniżenie inflacji, sposób formułowania tego celu w *Założeniach polityki pieniężnej…* przyjmowanych na kolejne lata wykazywał spore różnice[93]. W pierwszym okresie towarzyszył mu drugi równorzędny cel, związany z osiągnięciem stabilnej pozycji złotego wobec walut obcych. Poczynając od 1994 roku, zaczęto w *Założeniach…* powoływać się także na ustawowy cel NBP, jakim było umacnianie wartości pieniądza polskiego. Pojawiło się też po raz pierwszy stwierdzenie, że polityka pieniężna powinna sprzyjać wzrostowi gospodarczemu.

W latach 1990–1991 okazjonalny konflikt, jaki występował między dwoma celami polityki pieniężnej – ograniczeniem inflacji i obroną stałego kursu – został rozwiązany w praktyce na korzyść tego drugiego. Funkcja „kotwicy nominalnej" była traktowana jako kluczowa w celu wyhamowania inflacji. Początkowy silny impuls dewaluacyjny („przestrzelenie" dewaluacji – zob. rozdział 6) spowodował jednak, że inflacja utrzymywała się na wysokim poziomie (w samym 1990 roku ceny wzrosły o 245%). Skutkiem była szybka realna aprecjacja złotego, co z kolei obniżało konkurencyjność polskiego eksportu i mogło zagrozić stabilności kursu w średnim okresie. Z tego powodu w październiku 1991 roku złagodzono cel kursowy, wprowadzając system pełzającego kursu (*crawling peg*), a następnie system pełzającego pasma (*crawling band*) o ustalonej miesięcznej stopie dewaluacji. Starano się w ten

[92] Zob. ustawa z 31 stycznia 1989 roku o Narodowym Banku Polskim.

[93] Na przykład w 1991 roku celem było „wygaszenie tendencji inflacyjnych w celu stabilizacji siły nabywczej złotego", z kolei w 1993 roku – „umacnianie siły nabywczej złotego na rynku wewnętrznym". Dopiero w *Założeniach polityki pieniężnej na 1996 rok* pojawia się po raz pierwszy rozróżnienie celu finalnego, pośredniego i operacyjnego w polityce pieniężnej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

sposób zarówno kontrolować kurs, jak i spowalniać inflację. W dokumentach roboczych Międzynarodowego Funduszu Walutowego ten sposób prowadzenia polityki pieniężnej w Polsce po 1991 roku był nazywany podejściem „eklektycznym".

Jak pokazaliśmy jednak wyżej, zgodnie z zasadą „trójkąta niemożności" (punkt 5.10.2), w gospodarce otwartej na transgraniczne przepływy kapitału nie można jednocześnie ustalać stopy inflacji i kursu walutowego. Powstaje w związku z tym pytanie, jak można było realistycznie oczekiwać, że polityka pieniężna nakierowana na realizację dwóch niezależnych celów przy pomocy jednego instrumentu (krótkookresowa stopa procentowa) może przynieść oczekiwane rezultaty? Odpowiedź jest prosta – liczono na skuteczność polityki „eklektycznej", ponieważ transgraniczne przepływy kapitału w Polsce w pierwszym okresie lat 90. były bardzo ograniczone. Przyczyną była nieuregulowana sprawa obsługi polskiego zadłużenia zagranicznego. Polska w 1980 roku za sprawą jednostronnej decyzji przestała obsługiwać oficjalne zadłużenie wobec rządów państw zachodnich zrzeszonych w tak zwanym Klubie Paryskim. Oznaczało to, że kraj znalazł się w sytuacji technicznego bankructwa (*sovereign default*)[94]. W rezultacie napływ kapitału do gospodarki został wstrzymany. Ograniczenia w przepływach kapitału pozwalały na stosowanie strategii „eklektycznej".

Taki stan rzeczy utrzymywał się aż do końca 1994 roku, kiedy polski rząd – po długich negocjacjach – podpisał porozumienie z Klubem Paryskim o restrukturyzacji i częściowym umorzeniu polskiego zadłużenia[95]. Umożliwiło to Polsce powrót na międzynarodowe rynki finansowe i otworzyło zagranicznym inwestorom możliwość nabywania polskich papierów wartościowych. Obligacje rządowe pozostawały w tym czasie bardzo atrakcyjne, ponieważ ich stopy rentowności przekraczały 20%, a ryzyko kursowe było praktycznie zerowe na skutek administracyjnego ustalania kursu w ramach systemu pełzającej dewaluacji. W efekcie mieliśmy do czynienia z bardzo znacznym naruszeniem parytetu stóp procentowych. Skutkiem tego był masowy napływ kapitału – tylko w samym 1995 roku było to ponad 13 mld USD.

Napływ kapitału na tak dużą skalę zmuszał NBP do dokonywania kosztownych operacji sterylizacyjnych i doprowadził ostatecznie do dwukrotnej rewaluacji kursu złotego – łącznie o 11% w 1995 roku – oraz do poszerzenia pasma wahań

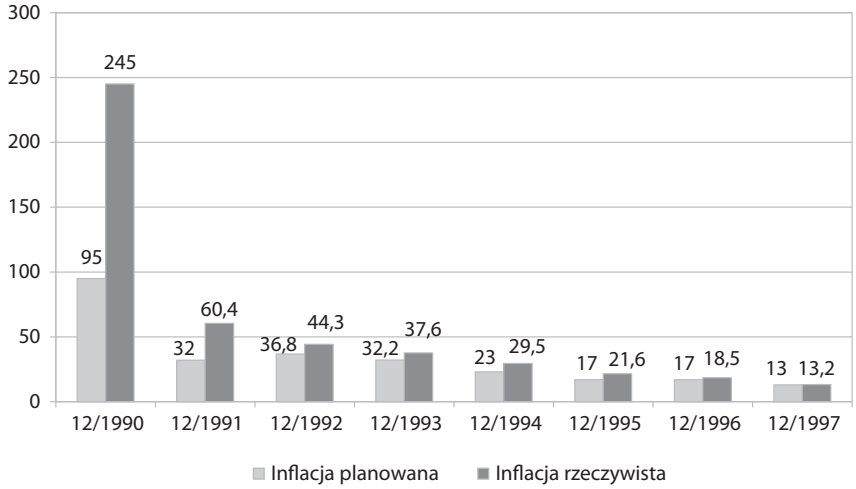
[94] Należy zaznaczyć, że Polska jednocześnie terminowo płaciła odsetki od zadłużenie prywatne wobec banków zagranicznych, zrzeszonych w tzw. Klubie Londyńskim.

[95] Równolegle Polska podpisała porozumienie z Klubem Londyńskim o uregulowaniu pozostałych płatności z tytułu zadłużenia prywatnego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kursu z +/–1% w 1994 roku do +/–7% w 1995 roku oraz do +/–10% na początku 1996 roku. Stałego kursu nie udało się więc obronić. Co więcej, wzrost podaży pieniądza i pełzająca dewaluacja (wówczas w tempie około 18% rocznie) podtrzymywały oczekiwania inflacyjne, wobec czego inflacja systematycznie utrzymywała się powyżej zakładanego poziomu. W 1997 roku NBP zmienił sposób prowadzenia polityki pieniężnej, zastępując cel pośredni w postaci tempa wzrostu podaży pieniądza celem pośrednim w postaci krótkookresowej stopy procentowej. Uzasadnieniem było obniżenie inflacji i systematyczny wzrost popytu na pieniądz[96]. Nie wyeliminowało to jednak podstawowej sprzeczności w realizowanej strategii pieniężnej (nadal dwa niezależne cele i tylko jeden instrument), a także nie zahamowało napływu kapitału zagranicznego. Spowodował on, że w kolejnych latach próby jednoczesnego kontrolowania kursu walutowego i inflacji w ramach strategii „eklektycznej" okazywały się coraz mniej skuteczne. Wykres 5.18 pokazuje, że cele inflacyjne zakładane w kolejnych rocznych *Założeniach polityki pieniężnej* NBP były systematycznie przekraczane. Ostateczny rozbrat ze strategią „eklektyczną" nastąpił w 1999 roku, kiedy zastąpiono ją strategią bezpośredniego celu inflacyjnego.

Wykres 5.18. Inflacja zakładana i inflacja rzeczywista w latach 1990–1997 (dane w %, r/r)



Uwaga: inflacja zakładana w corocznych *Założeniach polityki pieniężnej* NBP.

Źródło: opracowanie własne na podstawie danych NBP.

[96] W warunkach umiarkowanej inflacji i stabilnego popytu na pieniądz bardziej uzasadniona jest cenowa regulacja ilości pieniądza w obiegu, czyli przy pomocy stopy procentowej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Podsumowując, politykę pieniężną w Polsce w latach 1990–1998 można podzielić na trzy etapy:

- etap I (lata 1990–1991): strategia celu kursowego, czyli osiąganie celu inflacyjnego przy pomocy utrzymywania stałego kursu waluty;
- etap II (lata 1992–1996): strategia „eklektyczna I", czyli próba realizowania jednocześnie celu inflacyjnego (przy pomocy celu pośredniego M2) i celu kursowego (pełzające pasmo – *crawling band*);
- etap III (lata 1997–1998): strategia „eklektyczna II", czyli próba realizowania jednocześnie celu inflacyjnego (przy pomocy celu pośredniego w postaci krótkookresowej stopy procentowej NBP) i celu kursowego (pełzające pasmo – *crawling band*).

5.12.2. Lata 1999–2015: reguła BCI i stabilizacja inflacji

W 1997 roku w Polsce przyjęto nową Konstytucję, a w jej ramach art. 227 regulujący funkcjonowanie do Narodowego Banku Polskiego i polityki pieniężnej. Zapisy zawarte w tym artykule gwarantują niezależność NBP od władz politycznych, dają mu wyłączne prawo emisji pieniądza i prowadzenia polityki pieniężnej, nakładają obowiązek dbania o wartość polskiego pieniądza, a także powołują do życia nowy organ NBP – Radę Polityki Pieniężnej – odpowiedzialną za kształtowanie polityki pieniężnej w kraju (ramka 5.20). Na podstawie art. 227 została w 1997 roku przyjęta nowa ustawa o Narodowym Banku Polskim, która definiuje główny cel polityki pieniężnej jako „utrzymanie stabilności cen" (art. 3, ust. 1) i określa szczegółowo sposób powoływania i zasady funkcjonowania organów NBP, w tym Rady Polityki Pieniężnej.

Ramka 5.20. Art. 227 Konstytucji RP

1. Centralnym bankiem państwa jest Narodowy Bank Polski. Przysługuje mu wyłączne prawo emisji pieniądza oraz ustalania i realizowania polityki pieniężnej. Narodowy Bank Polski odpowiada za wartość polskiego pieniądza.
2. Organami Narodowego Banku Polskiego są: Prezes Narodowego Banku Polskiego, Rada Polityki Pieniężnej oraz Zarząd Narodowego Banku Polskiego.
3. Prezes Narodowego Banku Polskiego jest powoływany przez Sejm na wniosek Prezydenta Rzeczypospolitej, na 6 lat.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

4. Prezes Narodowego Banku Polskiego nie może należeć do partii politycznej, związku zawodowego ani prowadzić działalności publicznej nie dającej się pogodzić z godnością jego urzędu.
5. W skład Rady Polityki Pieniężnej wchodzą Prezes Narodowego Banku Polskiego jako przewodniczący oraz osoby wyróżniające się wiedzą z zakresu finansów, powoływane na 6 lat w równej liczbie przez Prezydenta Rzeczypospolitej, Sejm i Senat.
6. Rada Polityki Pieniężnej ustala corocznie założenia polityki pieniężnej i przedkłada je do wiadomości Sejmowi równocześnie z przedłożeniem przez Radę Ministrów projektu ustawy budżetowej. Rada Polityki Pieniężnej, w ciągu 5 miesięcy od zakończenia roku budżetowego, składa Sejmowi sprawozdanie z wykonania założeń polityki pieniężnej.
7. Organizację i zasady działania Narodowego Banku Polskiego oraz szczegółowe zasady powoływania i odwoływania jego organów określa ustawa.

Rada Polityki Pieniężnej pierwszej kadencji została powołana w lutym 1998 roku i już na jednym z pierwszych posiedzeń podjęła decyzję o przyjęciu strategii BCI i rozpoczęciu jej realizacji, poczynając od 1999 roku (NBP 1998). Ustalono, że głównym miernikiem inflacji ma być roczny wskaźnik CPI, choć równolegle dla potrzeb analitycznych wykorzystywano wskaźniki pomocnicze, takie jak PPI, wskaźnik cen budownictwa czy wskaźniki inflacji bazowej. RPP postanowiła również, że w horyzoncie 5 lat (czyli do 2003 roku) inflacja zostanie obniżona do 4% w skali roku (z ponad 13% w 1997 roku). Ten średniookresowy cel miał być realizowany stopniowo, za pośrednictwem systematycznie obniżanych celów rocznych, (zapisywanych w postaci dopuszczalnych przedziałów) ustalanych w *Założeniach polityki pieniężnej* na kolejne lata. Zdecydowano także, że formalnym kryterium wykonania celów rocznych będzie porównanie ustalonego krótkookresowego celu inflacyjnego z wielkością rocznego wskaźnika CPI w grudniu każdego roku. Rada podjęła jednocześnie decyzję o systematycznym opracowywaniu i publikowaniu powtarzalnych (trzy razy w roku) średniookresowych prognoz inflacji (w ramach okresowych raportów o inflacji).

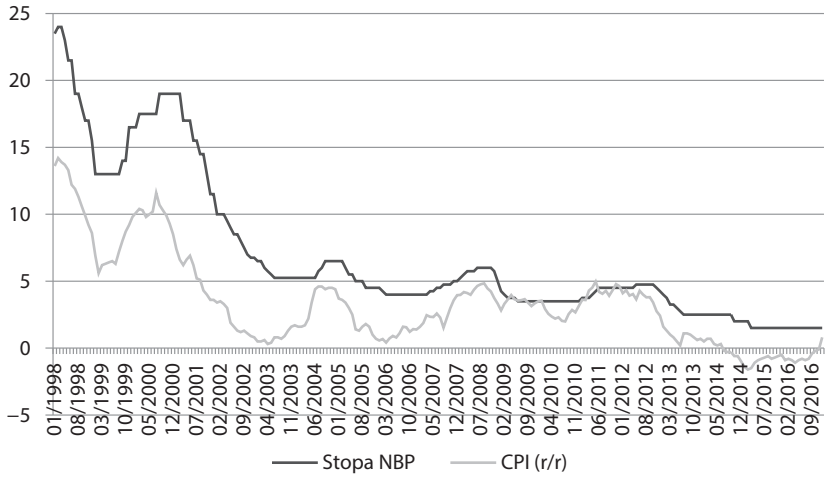
Po obniżeniu inflacji do 1,8%, w końcu 2003 roku Rada przyjęła średniookresowy cel inflacyjny w wysokości 2,5% rocznie (z przedziałem dopuszczalnych odchyleń +/–1 punkt procentowy). Tak określona strategia BCI była następnie realizowana przez kolejne kadencje RPP.

Czy z punktu widzenia rozwoju sytuacji makroekonomicznej i przebiegu zjawisk inflacyjnych w Polsce stosowanie strategii BCI można uznać za sukces? Wykres 5.19 przedstawia miesięczne obserwacje poziomu stopy referencyjnej NBP

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i wskaźnika CPI w latach 1998–2015. Zgodnie z przyjętą przez RPP strategią działania, można w tym czasie rozróżnić dwa okresy – okres silnej dezinflacji w latach 1998–2002 i okres umiarkowanej, względnie stabilnej inflacji w latach 2003–2015. Pierwszy z nich charakteryzują bardzo wysokie realne stopy procentowe (rzędu 6–12%). Ta wyjątkowo restrykcyjna polityka pieniężna była reakcją na przyspieszenie inflacji w 1999 roku. Skutkiem tego zaostrzenia stał się silny spadek inflacji w latach 2000–2002, co jednak było okupione gwałtownym obniżeniem tempa wzrostu gospodarczego (z ponad 6% w latach 1998–1999 do niewiele ponad 1% w 2001 roku). W kolejnych latach realne stopy procentowe pozostawały nadal dodatnie (z wyjątkiem kilku miesięcy w 2011 roku), przy czym ich poziom był wyraźnie niższy niż w pierwszych latach XXI wieku. Ogólnie można stwierdzić, że po okresie silnej dezinflacji inflacja w latach 2004–2013 pozostawała pod kontrolą, choć podlegała dość sporym wahaniom.

Wykres 5.19. Stopa referencyjna NBP i wskaźnik CPI w latach 1998–2016 (dane w %)



Źródło: opracowanie własne na podstawie danych NBP i GUS.

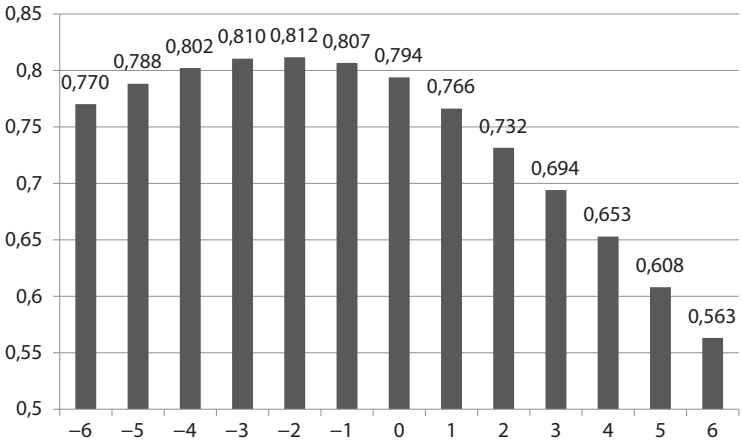
W latach 2014–2015 polityka pieniężna stanęła w obliczu nowego wyzwania, jakim była deflacja. Tendencje deflacyjne pojawiły się we wszystkich krajach rozwiniętych w odpowiedzi na spadające ceny surowców, paliw i żywności oraz w następstwie kryzysu zadłużenia. W reakcji na deflację RPP prowadziła dość konserwatywną politykę pieniężną, ostrożnie obniżając stopy procentowe. Można przypuszczać, że ta powściągliwość była spowodowana tym, że deflacja została

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wywołana głównie czynnikami leżącymi poza kontrolą polityki pieniężnej, a spadającym cenom towarzyszył dość stabilny wzrost gospodarczy i systematyczna poprawa na rynku pracy. W rezultacie prowadzenia tej ostrożnej polityki realne stopy procentowe wzrosły w latach 2013–2015 do poziomu 2–3%, czyli zbliżonego do poziomu obserwowanego w latach 2007–2008, kiedy tempo wzrostu PKB było niemal dwukrotnie wyższe.

Wykres 5.19 wskazuje na jeszcze jedną interesującą prawidłowość. Można mianowicie zauważyć, że momenty początku cyklu obniżek lub podwyżek stóp procentowych były z reguły opóźnione w stosunku do momentów występowania punktów zwrotnych w kształtowaniu się inflacji. Taka sytuacja miała miejsce, na przykład, przy rozpoczęciu cyklu podwyżek stóp w połowie 1999 roku i na początku 2007 roku oraz przy rozpoczęciu cyklu obniżek w 2001 roku i w połowie 2012 roku. Mogłoby to wskazywać na spóźnione reakcje polityki pieniężnej w stosunku do sygnałów cenowych. Gdyby tak rzeczywiście było, świadczyłoby to o tym, że polityka pieniężna była w praktyce bardziej zorientowana na reagowanie na *przeszłe* wartości inflacji (*backward-looking*) zamiast dążyć do oddziaływania na *przyszłe* tendencje cenowe (*forward-looking*).

Wykres 5.20. Współczynniki korelacji między zmianami stóp procentowych NBP i zmianami wskaźnika CPI w przypadku różnych miesięcznych opóźnień i wyprzedzeń (od –6 miesięcy do +6 miesięcy) w latach 1998–2015



Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Można w prosty sposób zweryfikować tę hipotezę, porównując korelację zmian stóp procentowych NBP i zmian cen odnośnie do różnych opóźnień i wyprzedzeń. Wykres 5.20 przedstawia wyniki tego eksperymentu. Jak widać, decyzje polityki pieniężnej były silniej skorelowane z wyprzedzającymi wartościami inflacji niż z wartościami opóźnionymi. Najsilniejsza korelacja występowała między decyzjami o zmianach stóp procentowych z okresu (t) i zmianami inflacji z okresu wcześniejszego o dwa miesiące ($t - 2$). Ogólnie rzecz biorąc, korelacja decyzji polityki pieniężnej z opóźnionymi (późniejszymi) zmianami wskaźnika CPI była wyraźnie znacznie słabsza niż korelacja z wyprzedzającymi (wcześniejszymi) zmianami wskaźnika CPI. Oznacza to, że polityka pieniężna w większym stopniu kierowała się obserwacjami z przeszłości niż przewidywaniami przyszłych zmian cen. Wynika z tego, że, realizując strategię BCI, Rada Polityki Pieniężnej nie przestrzegała konsekwentnie zasady „patrzenia do przodu" (*forward-looking*).

Na ile skuteczna w utrzymaniu stabilności cen okazała się realizowana w Polsce strategia BCI? W latach 1998–2003 dokonano silnej dezinflacji – dzięki restrykcyjnej polityce pieniężnej udało się obniżyć średnioroczną inflację z ponad 14% na początku 1998 roku do 1,8% w końcu 2003 roku, czyli znacznie poniżej celu wyznaczonego przez RPP (wynoszącego wówczas 4%)[97]. Zarazem z punktu widzenia formalnego kryterium realizacji celu inflacyjnego w ciągu całego okresu 1998–2003 (a także po 2012 roku) ani razu nie udało się jednak Radzie osiągnąć założonego celu inflacyjnego. Jak pokazuje wykres 5.21, faktyczna inflacja w tych latach mierzona wskaźnikiem CPI w końcu roku znajdowała się zawsze albo powyżej, albo poniżej celu inflacyjnego, ustalonego jako przedział dopuszczalnych wartości. Wśród przyczyn tego stanu rzeczy wskazać trzeba na przyjęcie obserwacji z jednego tylko miesiąca (grudzień) jako kryterium realizacji celu inflacyjnego i osadzanie go w zbyt wąskim przedziale wartości.

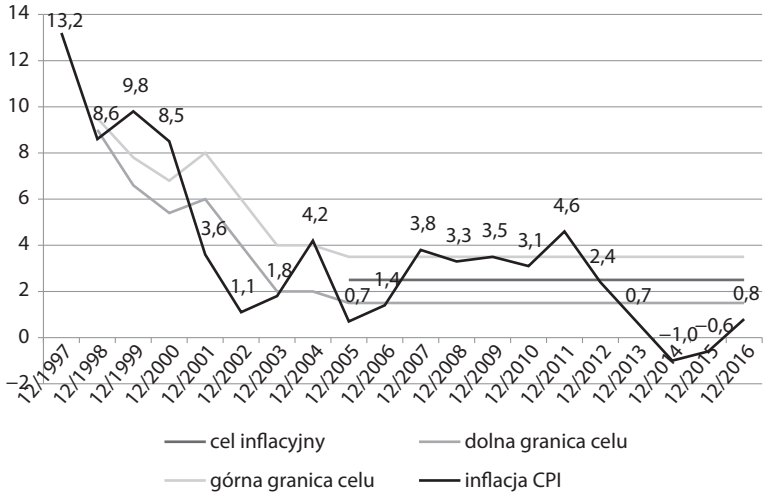
Poczynając od 2004 roku, Rada nie dokonywała już oceny realizacji celu na podstawie grudniowych obserwacji CPI, ponieważ przyjęty został stały cel „kroczący", ustalony na poziomie 2,5%. Gdyby jednak dalej stosować kryterium grudniowej inflacji, okazałoby się, że w większości lat inflacja wykraczała poza dopuszczalny przedział wahań. Pozwala to zauważyć, że inflacja w Polsce cechuje się stosunkowo

[97] Prowadzona w latach 2000–2002 polityka szybkiej dezinflacji nie tylko doprowadziła inflację do poziomu wyraźnie poniżej ustalonego celu (1,8% w porównaniu z 4,0%), lecz także silnie wyhamowała wzrost gospodarczy i przyczyniła się do wzrostu bezrobocia (do 18%). Z perspektywy twierdzenia o optymalnym konserwatyzmie banków centralnych polityka pieniężna w tym okresie była więc zdecydowanie zbyt restrykcyjna i wiązała się z dużym kosztem makroekonomicznym.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

dużymi wahaniami, co może mieć związek z wpływem czynników niezależnych od polityki pieniężnej, takich jak szoki zewnętrzne w postaci zmian cen surowców, paliw i żywności na rynkach światowych.

Wykres 5.21. Cel inflacyjny w Polsce i jego realizacja w 1998–2016 (CPI, dane w %, r/r)



Źródło: opracowanie własne.

Podsumowując, politykę pieniężną prowadzoną w Polsce w latach 1999–2015 można także podzielić na trzy główne etapy:

- etap I (lata 1999–2003): okres ostrej dezinflacji, w trakcie którego inflacja została obniżona do 1,8%;
- etap II (lata 2004–2013): okres stabilizacji inflacji wokół celu inflacyjnego zakładającego 2,5% +/–1 punkt procentowy;
- etap III (lata 2014–2016): okres bardzo niskiej inflacji i deflacji, kształtującej się wyraźnie poniżej celu inflacyjnego.

5.13. Polityka pieniężna a kryzys finansowy

5.13.1. Wpływ globalizacji

W okresie ostatnich dwóch dziesięcioleci XX wieku i w pierwszej dekadzie XXI wieku gospodarka światowa podlegała daleko idącym przeobrażeniom.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Głównym motorem tych zmian był proces globalizacji, powodujący rosnącą współzależność poszczególnych gospodarek narodowych. Zjawisko globalizacji jest wynikiem działania wielu czynników, w tym zwłaszcza rosnącej ekonomicznej i instytucjonalnej otwartości poszczególnych gospodarek na międzynarodowe przepływy towarów, usług i kapitału (a w wielu przypadkach także i siły roboczej) oraz szybkiego postępu technologicznego, zwłaszcza w informatyce i telekomunikacji. Dzięki liberalizacji handlu międzynarodowego dokonanej w ramach kolejnych rund wielostronnych negocjacji handlowych w ramach GATT i WTO, a także w ramach regionalnych ugrupowań integracyjnych, tempo wzrostu eksportu oraz importu przyspieszyło, a udział handlu zagranicznego w światowym PKB systematycznie wzrastał. Dzięki wyeliminowaniu ograniczeń w wymienialności walut narodowych i znoszeniu barier w przepływie kapitałów w skali międzynarodowej zwiększyły się strumienie bezpośrednich inwestycji zagranicznych oraz inwestycji portfelowych, a także przepływy kapitału finansowego (pożyczki, kredyty). Integracji rynków finansowych sprzyjał rozwój nowych technologii informatycznych i telekomunikacyjnych, dzięki którym obniżyły się koszty i skrócił się czas dokonywania transakcji obejmujących transgraniczne przepływy kapitału. Narodowe rynki finansowe w poszczególnych krajach podlegały spontanicznemu procesowi umiędzynarodowienia i intensywnej integracji. W wyniku rozwoju tych wzajemnych powiązań, doszło do wykształcenia się względnie jednolitego rynku finansowego na skalę światową.

Konsekwencją globalizacji jest zaś zwiększona współzależność poszczególnych gospodarek. Ma to dobre i złe strony. Rozwój wymiany międzynarodowej pozwala na osiągnięcie korzyści ze specjalizacji produkcji, a rosnąca skala przepływów kapitału umożliwia lepsze wykorzystanie zasobów produkcyjnych w skali transnarodowej. Wszystko to umożliwia szybszy wzrost gospodarczy. Jednocześnie globalizacja wywołuje spadek inflacji dzięki możliwościom arbitrażu cenowego, jakie powstają w wyniku integracji rynków towarowych i usługowych. Zwiększona otwartość na handel, a także na przepływy kapitałowe, spowodowała wzrost odporności gospodarek narodowych na wewnętrzne szoki inflacyjne, które można było neutralizować dzięki importowi. W rezultacie zwiększonej otwartości udało się skutecznie spowolnić napięcia inflacyjne dzięki taniemu importowi. Oznacza to, że nastąpiło spłaszczenie krzywej Phillipsa w poszczególnych krajach, a rola krajowej luki popytowej w generowaniu inflacji uległa ograniczeniu na rzecz światowej luki popytowej.

Z drugiej strony globalizacja zwiększyła jednak uzależnienie poszczególnych gospodarek od otoczenia, w tym zwłaszcza ich podatność na negatywne szoki

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zewnętrzne (*negative spillovers*). Sytuacja w danym kraju stawała się w rosnącym stopniu funkcją sytuacji w innych krajach. Przede wszystkim poszczególne kraje zaczęły coraz silniej odczuwać konsekwencje polityki gospodarczej prowadzonej w innych państwach, w tym zwłaszcza w krajach największych, takich jak Stany Zjednoczone czy Chiny, a także w Unii Europejskiej.

Jednym z najważniejszych skutków tych zmian było ograniczenie stopnia kontroli banku centralnego nad podażą pieniądza i nad inflacją. W szczególności inflacja w coraz większym stopniu stawała się procesem zależnym od czynników egzogenicznych, znajdujących się poza zakresem kontroli banków centralnych. Postawiło to przed polityką pieniężną nowe wyzwanie: czy i jak bank centralny, zwłaszcza w małej otwartej gospodarce, powinien reagować na zjawiska znajdujące się zasadniczo poza jego kontrolą, takie jak zmiany cen na rynkach światowych, masowe przepływy kapitału czy rozwój nowych instrumentów finansowych (zob. Rosati 1999). Dodatkowe wyzwania pojawiły się w latach 2008–2011, kiedy na długofalowe tendencje związane z globalizacją nałożyły się skutki najgłębszego kryzysu finansowego i gospodarczego, jaki dotknął gospodarkę krajów rozwiniętych od czasów Wielkiego Kryzysu z lat 1929–1933.

5.13.2. Kryzys finansowy w latach 2008–2009

Kryzys finansowy zapoczątkowało załamanie na rynku kredytów hipotecznych w Stanach Zjednoczonych. Seria bankructw dużych instytucji finansowych, do jakich doszło w 2008 roku, wywołała spadek zaufania na rynku międzybankowym. W warunkach niepełnej informacji o ryzyku obciążającym poszczególne kategorie aktywów finansowych wzrosło znacznie tzw. ryzyko partnera (lub kontrahenta – *counterparty risk*), a w konsekwencji rynek międzybankowy uległ zablokowaniu. Spowodowało to najpierw ostry kryzys płynności, który zmusił liczne banki do awaryjnej sprzedaży aktywów w celu uzupełnienia niedoborów płynności. To z kolei doprowadziło do załamania rynków akcji i obligacji, gwałtownego obniżenia ich cen i ogromnych strat kapitałowych. W rezultacie wiele instytucji finansowych znalazło się w stanie upadłości (zob. ramka 5.21).

Ostry deficyt płynności w sektorze finansowym spowodował ograniczenie akcji kredytowej (zjawisko tzw. *credit crunch*) ze strony banków i odcięcie wielu przedsiębiorstw oraz gospodarstw domowych od finansowania kredytowego. Kryzys w sektorze finansowym rozprzestrzenił się na sektor realny, a w rezultacie gospodarki krajów rozwiniętych weszły w głęboką recesję.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 5.21. Mechanizm transmisji ryzyka: bankructwo banku Lehman Brothers

Dobrym przykładem działania mechanizmu transmisji ryzyka, ilustrującym skalę i siłę powiązań między bankami, jest upadek banku Lehman Brothers we wrześniu 2008 roku. Jako czwarty co do wielkości bank inwestycyjny na świecie (ponad 600 mld USD aktywów) Lehman był powiązany z wieloma innymi bankami siecią wzajemnych pożyczek i inwestycji. Jego bankructwo spowodowało straty na ekspozycjach, jakie miały wobec niego inne banki, co zmusiło je do tworzenia odpisów na zagrożone aktywa, i wywołało niedobór płynności. Uzupełnienie deficytu płynności na rynku międzybankowym okazało się jednak niemożliwe, ponieważ banki wstrzymywały się z udzielaniem sobie pożyczek na rynku międzybankowym zarówno ze względu na własne zwiększone potrzeby płynnościowe, jak i z uwagi na podwyższone ryzyko niewypłacalności potencjalnych pożyczkobiorców – tzw. ryzyko partnera (*counterparty risk*) – w warunkach asymetrycznej informacji o ich kondycji finansowej. Skutkiem tego było zamrożenie całego rynku pożyczek międzybankowych w USA. To z kolei zmusiło banki mające kłopoty z utrzymaniem płynności do wyprzedaży aktywów w trybie awaryjnym, niekiedy panicznym, i to bez względu na cenę (*fire sale*), co doprowadziło do silnego spadku cen aktywów i poważnych strat bilansowych. Upadek Lehman Brothers spowodował więc nie tylko straty bezpośrednie tego banku oraz jego wierzycieli, lecz także straty instytucji, które Lehman bezpośrednio finansował.

Pośrednim, ale niesłychanie ważnym skutkiem tego bankructwa był gwałtowny wzrost kosztów finansowania na rynku. Banki, chcące uzupełnić swoje kapitały uszczuplone tworzeniem rezerw na pokrycie zagrożonych należności, napotykały na duże trudności z pozyskaniem środków. Zmuszało je to do dalszego ograniczenia akcji kredytowej, tak aby zmniejszyć sumę aktywów w dostosowaniu do kurczącej się bazy kapitałowej.

Banki centralne starały się przeciwdziałać negatywnym skutkom kryzysu, obniżając stopy procentowe niemal do zera i dostarczając płynność do sektora bankowego na masową skalę. Skuteczność tych działań okazała się jednak ograniczona. Duże zadłużenie przedsiębiorstw i gospodarstw domowych w sytuacji ogólnego brak zaufania i przy kiepskich perspektywach wzrostu powstrzymywało banki przed wzrostem akcji kredytowej. Dominowały nadal obawy o wypłacalność dłużników.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Na skutek przewlekłej recesji w wielu krajach, zwłaszcza w strefie euro, pojawiły się symptomy deflacji.

Dla banków centralnych była to pod wieloma względami nowa sytuacja. Po pierwsze, w sposób naturalny pojawiły się pytania o to, czy polityka pieniężna mogła zapobiec kryzysowi, w tym, w szczególności, czy mogła nie dopuścić do narastania baniek spekulacyjnych na rynkach aktywów, między innymi na rynkach nieruchomości. Gdyby okazało się, że banki centralne mogły zapobiec spekulacyjnemu wzrostowi cen aktywów, nie doszłoby do nadmiernej ekspansji kredytów hipotecznych, i można byłoby prawdopodobnie uniknąć kryzysu. Po drugie, co mogły zrobić banki centralne, aby pomóc gospodarce wyjść z recesji, w sytuacji, gdy stopy procentowe zostały już obniżone praktycznie do zera? Czy oznacza to, że polityka pieniężna osiągnęła kres swoich możliwości i jest w istocie bezsilna wobec zjawisk kryzysowych, czy możliwe jest jednak sięgnięcie do innych, niekonwencjonalnych metod, wykraczających poza dotychczasowe ramy polityki pieniężnej? W dalszej części niniejszego podrozdziału spróbujemy udzielić odpowiedzi na te pytania.

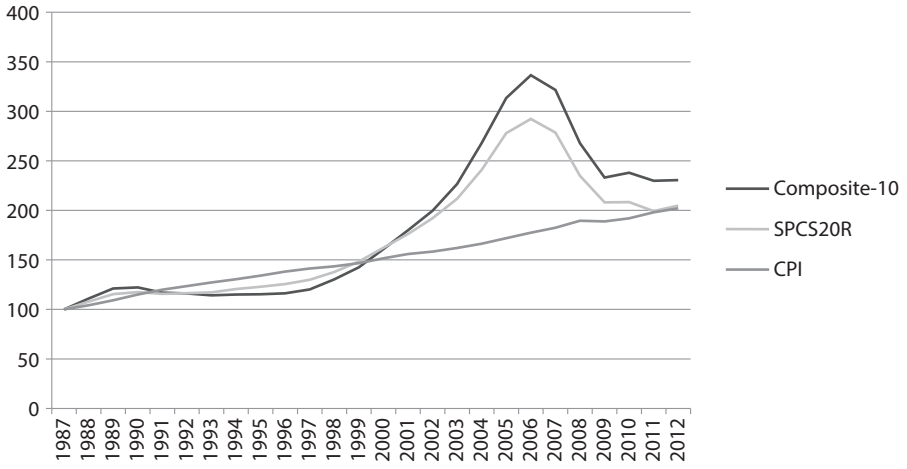
5.13.3. Polityka pieniężna a bańki spekulacyjne na rynku aktywów

Kryzys z lat 2008–2009 przypomniał o tym, że na rynkach finansowych możliwe są długotrwałe wzrosty cen niektórych aktywów, nieuzasadnione ich rzeczywistą wartością. Te anomalie cenowe prowadzą do powstawania baniek spekulacyjnych (*speculative bubbles*), które w pewnym momencie pękają, powodując gwałtowny spadek cen aktywów. Pęknięcie bańki spekulacyjnej może wywołać falę bankructw, a w dalszej konsekwencji recesję i nawet deflację.

Przypomnijmy, że cena danego aktywu równa się – w warunkach równowagi – sumie zdyskontowanych oczekiwanych dochodów z tego aktywu w przyszłości, wliczając w to cenę odsprzedaży lub wykupu. Są to czynniki fundamentalne, decydujące w długim okresie o rzeczywistej wartości aktywów (*fundamentals*). Odchylanie się bieżących cen rynkowych aktywów od tych wartości fundamentalnych oznacza, że aktywa są przeszacowane lub niedoszacowane. W okresie przedkryzysowym doszło do wyraźnego odchylenia się cen aktywów w górę od poziomu równowagi, co było spowodowane nadmiernie optymistycznymi prognozami zmiany cen i wynikającym z tych prognoz wysokim popytem na aktywa, finansowanym kredytami bankowymi. Wykres 5.22 pokazuje, że na rynku nieruchomości mieszkalnych w USA doszło w latach 2000–2006 do silnego wzrostu cen, w porównaniu ze wskaźnikiem CPI, co wyraźnie świadczyło o narastaniu bańki spekulacyjnej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 5.22. Zmiany cen domów w USA w latach 1987–2012 (1987 = 100)[a]



a) Composite-10 CSXR – wskaźnik cen domów w 10 największych obszarach metropolitalnych w USA (composite index of the home price index for 10 major Metropolitan Statistical Areas in the United States); SPCS20R – wskaźnik cen domów w 20 największych miastach USA (Standard&Poor/Case-Shiller 20-City Composite Home Price Index).

Źródło: opracowanie własne na podstawie danych Standard and Poor's.

Wzrost cen domów zachęcał do spekulacyjnych zakupów. Inwestorzy zaciągali kredyty, aby inwestować w nieruchomości w nadziei, że wzrost ich cen pozwoli spłacić zadłużenie i solidnie zarobić. Ten spekulacyjny popyt dodatkowo przyspieszał wzrost cen. W tym procesie wzajemnego napędzania się cen i popytu kluczową rolę odgrywał mechanizm akceleratora finansowego, który wyjaśnia, dlaczego zachowania uczestników rynku przyczyniają się do wzrostu ryzyka (Minsky 1992; Kindleberger 2005; Bernanke i Gertler 1999; Bernanke, Gertler i Gilchrist 1999). Zgodnie z tym mechanizmem nieuzasadniony, nadmierny optymizm (*irrational exuberance*) co do przyszłych zysków prowadzi do nadmiernych inwestycji, wzrostu cen aktywów, zwiększenia rynkowej wartości zabezpieczeń (wzrost względnych cen aktywów) i napływu kapitału (pożyczkowego) na rynek. W ten sposób akcelerator finansowy skłania firmy i gospodarstwa domowe do zwiększania zadłużenia i wydatków. Gdy jednak – prędzej czy później – rzeczywistość okazuje się gorsza od oczekiwań, ceny aktywów i wartości zabezpieczeń spadają. Akcelerator finansowy ulega odwróceniu – kapitał odpływa, ceny aktywów maleją, wartość zabezpieczeń spada. Wywołuje to relatywny wzrost zadłużenia (relacja długu do aktywów rośnie), co ogranicza zdolność uczestników rynku do obsługi długu i pozyskania nowych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kredytów. Akcelerator jest mechanizmem procyklicznym – powoduje wzmocnienie i przedłużenie boomu oraz pogłębienie i przedłużenie recesji.

Czy polityka pieniężna ponosi odpowiedzialność za narastanie baniek spekulacyjnych na rynkach aktywów? Część ekonomistów stawia zarzut bankom centralnym, że, utrzymując zbyt niskie stopy procentowe przez zbyt długi czas, doprowadziły do kryzysu finansowego. Między innymi John Taylor wskazuje na to, że stopy procentowe Fed w latach 2002–2006 były niższe o około 2–3 punkty procentowe od stóp hipotetycznych, spełniających regułę Taylora (Taylor 2009). Trudno jest jednak znaleźć empiryczne potwierdzenie na istnienie bezpośredniego związku pomiędzy niskimi stopami Fed a spekulacyjnym wzrostem cen aktywów, zwłaszcza nieruchomości (zob. Rosati 2009). Podobnie sceptycznie wobec roli niskich stóp procentowych w doprowadzaniu do kryzysu wypowiadają się Reinhart i Reinhart, (2011), którzy wskazują raczej na udział rosnącej globalnej nadpłynności (*global savings glut*), czyli napływu do USA oszczędności z krajów „nadwyżkowych", takich jak Chiny, Japonia, kraje naftowe czy Rosja. Kapitały te finansowały rosnący popyt na aktywa na rynkach finansowych w Stanach Zjednoczonych.

Bardziej przekonująca wydaje się teza o pośrednim wpływie luźnej polityki pieniężnej na skłonność do podejmowania ryzyka i stosowania dźwigni finansowej w sektorze bankowym (De Nicolo i in. 2010, Borio 2011). Ten kanał oddziaływania (*risk-taking channel*) zachęca banki do podejmowania ryzykownych inwestycji w sytuacji, gdy stopy procentowe znajdują się przez dłuższy czas na niskim poziomie i gdy uczestnicy rynku oczekują dalszego luzowania polityki pieniężnej w przyszłości.

Inne ważne pytanie, jakie pojawiło się w kontekście niedawnego kryzysu, dotyczy tego, czy i jak banki centralne powinny przeciwdziałać narastaniu baniek spekulacyjnych. Teoretycznie bank centralny może zareagować na pojawiające się napięcia na rynku aktywów na kilka różnych sposobów. Na przykład jest w stanie włączyć zmiany cen aktywów do głównego wskaźnika inflacji (CPI, HICP) i reagować na nadmierny wzrost tego wskaźnika poprzez podniesienie stóp procentowych, gdy tak zdefiniowana „inflacja" przekroczy pewien poziom. Jest to działanie wyprzedzające (*ex ante*) i sprowadza się do „przebicia" bańki spekulacyjnej jeszcze w fazie jej narastania[98]. Mankamentem tej metody jest jednak to, że naraża ona

[98] Wiele wskazuje na to, że tak właśnie powinien był zachować się bank centralny Japonii w latach 90., kiedy na skutek ekspansywnej polityki pieniężnej doszło do spekulacyjnego wzrostu cen aktywów, finansowanego kredytem bankowym. Pęknięcie bańki spekulacyjnej zagroziło stabilności sektora bankowego,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

politykę pieniężną na względnie dużą niestabilność. Gdy ceny aktywów szybko rosną, a ich udział we wskaźniku inflacji jest relatywnie wysoki, działanie *ex ante* wymaga dużych wzrostów stóp procentowych, co może istotnie utrudniać działalność sektora realnego i hamować wzrost gospodarczy. Gdy z kolei ceny aktywów rosną wolniej lub nawet spadają, wskaźnik CPI będzie zaniżał rzeczywistą inflację, a nadmiernie luźna polityka pieniężna może wówczas zagrozić stabilności cen. Co więcej, bank centralny często nie jest w stanie ocenić, czy i kiedy wzrost cen aktywów, w tym na przykład domów, jest na tyle duży, że uzasadnia zaostrzenie polityki pieniężnej.

Zamiast podejmować ryzyko działania wyprzedzającego bank centralny może utrzymywać umiarkowane stopy procentowe i czekać aż bańka spekulacyjna sama pęknie, a następnie przeciwdziałać tendencjom recesyjnym, agresywnie luzując politykę pieniężną *ex post*. Takie podejście ekonomiści zaczęli nazywać doktryną Greenspana, ponieważ to właśnie szef amerykańskiego Fed przekonywał do słuszności stosowania takiej polityki (Greenspan 1999, Bernanke i Gertler 2001). Zgodnie z tym poglądem, polityka pieniężna powinna uwzględniać zmiany cen aktywów tylko w takim stopniu, w jakim wpływają one na prognozę przyszłej inflacji. Gdy prognoza przyszłej inflacji uwzględnia bieżące zmiany cen aktywów, dodatkowa reakcja polityki pieniężnej na te zmiany nie jest już potrzebna. Tak właśnie postępowała większość banków centralnych w okresie poprzedzającym kryzys. Okazało się jednak, że nie było to w stanie zapobiec wejściu gospodarki USA w głęboką recesję w 2009 roku mimo bardzo agresywnego luzowania polityki pieniężnej i szybkiego obniżenia stóp procentowych praktycznie do zera. Ponadto w dłuższym okresie niepożądanym skutkiem takiego działania może być trwałe odchylenie nastawienia polityki pieniężnej w stronę nieuzasadnionej ekspansji, ponieważ poluzowanie polityki w okresie pokryzysowym nie jest równoważone odpowiednim zaostrzeniem polityki w okresie ekspansji gospodarczej. Ta asymetria zwiększa prawdopodobieństwo wystąpienia kolejnych baniek spekulacyjnych w przyszłości.

Jeszcze innym sposobem postępowania jest wykorzystanie strategii „opierania się wiatrowi" (*leaning against the wind*). Bank centralny stara się wówczas uwzględniać zagrożenia płynące z nadmiernego wzrostu cen aktywów w odniesieniu do stabilności cen w średnim horyzoncie czasu i reaguje, stopniowo zaostrzając politykę pieniężną. Celem nie jest w tym przypadku przebijanie bańki spekulacyjnej,

---

który musiał silnie ograniczyć kredyty dla sektora realnego, a to wpędziło japońska gospodarkę w wieloletnią stagnację.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

a raczej systematyczne ograniczanie dostępu do finansowania, tak aby ją „zagłodzić". Zaletą tej strategii jest działanie w sposób wyprzedzający i symetryczny. Lecz metoda „opierania się wiatrowi" wymaga bardzo dobrego rozeznania w sytuacji na rynku aktywów i umiejętności trafnego przewidywania jego dalszej dynamiki. Gdy tego rozeznania brakuje, strategia stopniowego zaostrzania polityki pieniężnej może nie być skuteczna, a według niektórych autorów okaże się nawet przeciwskuteczna, gdy wzrost stóp procentowych pobudzi dalszy wzrost cen aktywów (ramka 5.22).

Ramka 5.22. Ceny aktywów a stopy procentowe

Ekonomista Jordi Galì (2013) twierdzi, że cena aktywu na rynku finansowym może być rozpatrywana jako suma dwóch składników – składnika fundamentalnego (*fundamental component*) i składnika spekulacyjnego (*bubble component*). Przy założeniu, że uczestnicy rynku zachowują się racjonalnie, składnik fundamentalny ceny aktywu jest równy sumie zdyskontowanych przyszłych przychodów z danego aktywu, z kolei składnik spekulacyjny powinien (w równowadze) być równy krótkookresowej stopie procentowej, ponieważ stanowi ona w danym momencie koszt alternatywny inwestycji w dane aktywa. Podniesienie stóp procentowych przez bank centralny w celu zahamowania narastania bańki spekulacyjnej wywołuje więc wzrost kosztu alternatywnego, i w konsekwencji, składnika spekulacyjnego, co przyczynia się do dalszego zwiększania ceny aktywu i narastania spekulacji. W takich warunkach strategia „opierania się wiatrowi" może okazać się przeciwskuteczna, ponieważ pogłębia nierównowagę. Zamiast tego, bank centralny powinien stosować taką politykę pieniężną, która dąży do zachowania równowagi między stabilizowaniem bańki spekulacyjnej oraz stabilizowaniem zagregowanego popytu.

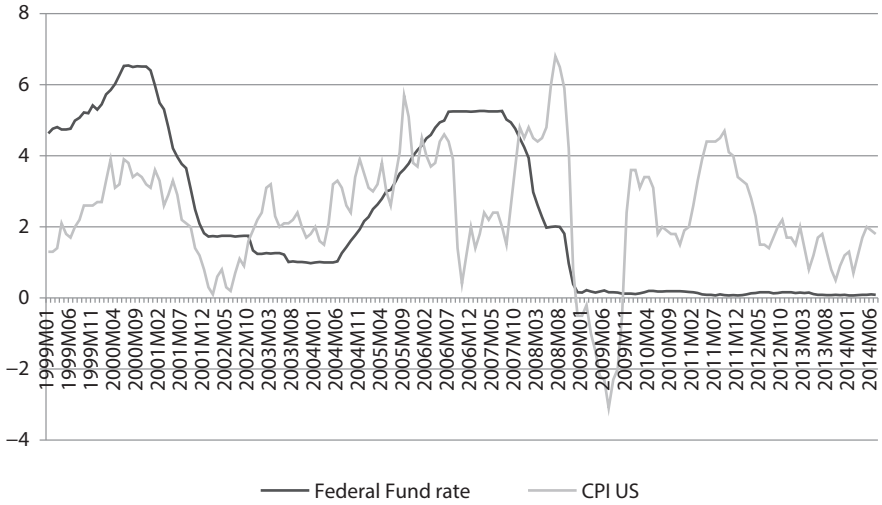
5.13.4. Czy realizacja strategii celu inflacyjnego może zapobiegać kryzysom finansowym?

Kryzys finansowy pokazał, że polityka pieniężna, która skupia się na realizacji jednego celu, jakim jest utrzymanie stabilności cen, nie jest w stanie zapobiec narastaniu spekulacyjnej nierównowagi na rynku aktywów i nie chroni przed

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

recesyjnymi skutkami pęknięcia bańki spekulacyjnej. W Stanach Zjednoczonych po okresie luzowania polityki pieniężnej w latach 2001–2003, Fed w kolejnych latach stopniowo podnosił stopy procentowe, reagując na wzrost inflacji w latach 2004–2005. Od końca 2005 roku do końca 2007 roku stopa funduszy federalnych wyraźnie przekraczała 4% i była wyższa od stopy inflacji (wykres 5.23), i Fed nie widział powodu, aby jeszcze bardziej zaostrzać politykę pieniężną. Z punktu widzenia kontroli inflacji polityka pieniężna Fed był w ogólnym zarysie prawidłowa; niemniej, nie zdołała zapobiec nadmiernemu wzrostowi cen aktywów i narastaniu bańki spekulacyjnej, co w ostatecznym rezultacie doprowadziło do kryzysu.

Wykres 5.23. Inflacja CPI i stopa funduszy federalnych (*federal funds rate*) w USA w latach 1999–2014 (dane w %)



Źródło: opracowanie własne na podstawie danych Eurostatu (CPI) i Systemu Rezerwy Federalnej (*federal funds rate*).

Ponieważ polityka pieniężna nakierowana na realizację celu inflacyjnego nie zapobiega kryzysom finansowym, rodzi się pytanie, czy i w jaki sposób bank centralny może zapewnić stabilność na rynkach finansowych? Trzeba przede wszystkim przypomnieć, że kondycja sektora finansowego, zwłaszcza banków, ma bardzo istotne znaczenie w przypadku realizacji polityki pieniężnej, ponieważ to banki są najważniejszym ogniwem mechanizmu transmisji monetarnej. Gdy angażują się w spekulacyjne zakupy aktywów, rośnie ryzyko kryzysu płynności oraz kryzysu wypłacalności w sektorze bankowym, co może zakłócić transmisję impulsów polityki pieniężnej. Bank centralny, pełniąc swoją funkcję „pożyczkodawcy

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ostatniej szansy", przychodzi z reguły z pomocą tym bankom, które są zasadniczo wypłacalne, ale zostały dotknięte przejściowym brakiem płynności, ponieważ upadek banku blokuje kanały transmisji impulsów polityki pieniężnej i pociąga za sobą duże koszty zewnętrzne dla gospodarki (*negative externality*). Bank centralny pożycza wówczas płynne środki bankom zagrożonym kryzysem, zwiększając skalę regularnych operacji otwartego rynku lub poprzez udzielanie awaryjnych pożyczek krótkoterminowych bezpośrednio poszczególnym bankom (na przykład w ramach *marginal lending facility* w strefie euro lub w formie kredytu lombardowego w Polsce).

Tego typu działanie należy do standardowych funkcji banku centralnego, lecz stanowi jedynie reakcję na powstałe zagrożenie, w żaden sposób mu nie zapobiegając. Doświadczenia ostatniego kryzysu pokazały, że zmiany stóp procentowych i podaży pieniądza nie wystarczają do zapewnienia stabilności na rynkach finansowych (Borio 2011). Nic w tym dziwnego, ponieważ zapewnienie stabilności finansowej jest czymś innym niż zapewnianie stabilności cen. Powierzenie bankowi centralnemu zadania utrzymania stabilności finansowej oznacza więc nałożenie nań dodatkowego celu – poza troską o stabilność cen, a to wymaga – zgodnie z zasadą Tinbergena – wyposażenia go w dodatkowy instrument. Mogą być nim regulacje ostrożnościowe (*prudential regulations*) oraz nadzór na funkcjonowaniem banków (*banking supervision*). Korzystając z tych narzędzi, bank centralny może hamować nadmierny wzrost kredytów, ograniczać skłonność banków do podejmowania przesadnie ryzykownych inwestycji, by w ten sposób zapobiegać narastaniu nadmiernych nierównowag.

Uwzględnienie w ramach mandatu banku centralnego celu w postaci utrzymania stabilności finansowej nie musi więc być sprzeczne z realizacją celu inflacyjnego. Bank centralny realizuje wówczas dwa odrębne cele przy pomocy dwóch różnych instrumentów. Nie chodzi zatem o „włączenie" celu stabilności finansowej do celu inflacyjnego poprzez, na przykład, objęcie wskaźnikiem inflacji zmian cen aktywów. Oba cele są traktowane odrębnie, a bank centralny – przy pomocy dwóch instrumentów – stosuje równolegle dwie równoległe polityki – politykę pieniężną i politykę nadzoru ostrożnościowego[99].

[99] Oczywiście nie oznacza to, że stan sektora finansowego nie ma wpływu na inflację i odwrotnie. Lecz na poziomie operacyjnym założenie o niezależności obu celów jest dopuszczalne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.14. Nowe tendencje w polityce pieniężnej

5.14.1. Problem „zerowej dolnej granicy" i potrzeba zastosowania niekonwencjonalnych instrumentów polityki pieniężnej

Upadek amerykańskiego banku inwestycyjnego Lehman Brothers we wrześniu 2008 roku spowodował zablokowanie rynku międzybankowego w USA. Brak płynności i straty na aktywach zmusiły banki do ograniczenia kredytów udzielanych dla sektora realnego i w rezultacie pojawiło się zjawisko tzw. zapaści kredytowej (*credit crunch*). Reagując na gwałtowny kryzys płynności systemu bankowego, banki centralne agresywnie luzowały politykę pieniężną, obniżając szybko stopy procentowe i zwiększając podaż pieniądza. Wykres 5.24 pozwala zauważyć, że między sierpniem 2008 roku a marcem 2009 roku podstawowe stopy procentowe w głównych krajach rozwiniętych zostały drastycznie obniżone i od tego czasu utrzymują się na historycznie niskim poziomie w przedziale od 0 do 0,5% w skali rocznej. Niektóre banki centralne (strefa euro, Szwajcaria, Dania) wprowadziły też ujemne stopy od depozytów utrzymywanych przez banki komercyjne na rachunkach w banku centralnym. Poza obniżaniem ceny pieniądza, banki centralne w krajach rozwiniętych podjęły też – na różną skalę – bezwarunkowy skup aktywów na rynku finansowym (tzw. transakcje *outright*), aby bezpośrednio zwiększać zasoby płynności znajdujące się w posiadaniu banków komercyjnych. Działania te miały na celu skłonienie banków komercyjnych do zwiększenia kredytu dla sektora realnego.

Mimo obniżenia stóp procentowych niemal do zera, gospodarki krajów rozwiniętych nie powróciły jednak na ścieżkę szybkiego wzrostu, a stopy bezrobocia pozostawały nadal na wysokim poziomie[100]. Dotyczy to zwłaszcza państw strefy euro, gdzie w latach 2009–2014 mieliśmy do czynienia praktycznie ze stagnacją gospodarczą. Co więcej, mimo bardzo niskiej ceny pieniądza i „pompowania" płynności do gospodarki, inflacja utrzymuje się wyraźnie poniżej celu inflacyjnego, a w niektórych krajach pojawiły się tendencje deflacyjne[101]. Jako jedną z głównych przyczyn tego stanu rzeczy wymienia się wysoki poziom zadłużenia i dążenie

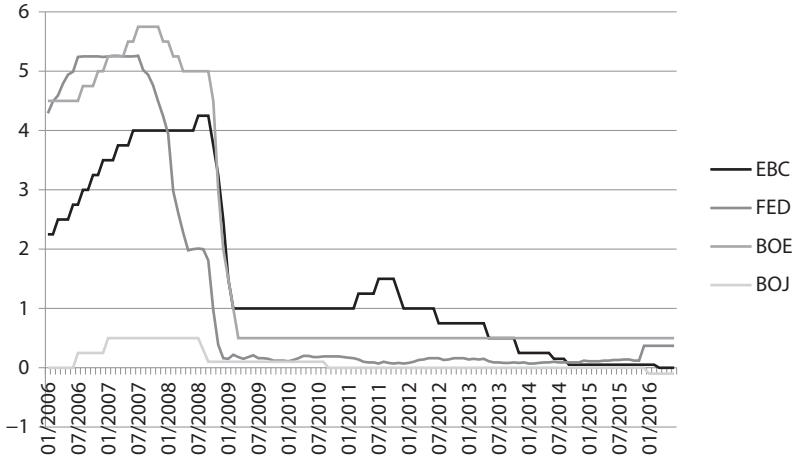
[100] Na ryzyka związane z polityką bardzo niskich stóp procentowych wskazuje Rzońca (2014).

[101] Niektórzy autorzy zalecają, aby w obliczu ryzyka „zerowego limitu" konwencjonalna polityka pieniężna została podporządkowana zasadzie „niżej na dłużej" (*lower for longer*), która nakazuje agresywnie obniżać stopy i utrzymywać je na niskim poziomie przez dłuższy czas, niż oczekują tego rynki (Williams 2014).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przedsiębiorstw oraz gospodarstw domowych do obniżenia długu, czyli tzw. delewarowanie (*delevereging* – zob. Eggertson i Krugman 2010, Gros 2012).

Wykres 5.24. Podstawowe stopy procentowe banków centralnych w USA, Wielkiej Brytanii, Japonii oraz w strefie euro w latach 2006–2016 (dane w % – stan na koniec miesiąca)



Źródło: opracowanie własne na podstawie danych pochodzących z banków centralnych.

Gdy nominalne stopy procentowe są bliskie zeru, bank centralny traci możliwość dalszego luzowania polityki pieniężnej tradycyjnymi metodami, dlatego że – zgodnie z konwencjonalną teorią ekonomii – nominalna cena pieniądza nie może być ujemna[102]. W tych warunkach jakikolwiek dalszy spadek inflacji wywołuje wzrost realnych stóp procentowych, a to z kolei przekłada się na kolejny spadek popytu i wydatków oraz dalszy spadek inflacji. Prędzej czy później pojawia się deflacja, która jeszcze bardziej podnosi realne stopy procentowe, utrudniając gospodarce wyjście z recesji.

Wyzwanie, jakie staje przed polityką pieniężną, gdy nominalne stopy zbliżają się do zera (*zero lower bound*), można przedstawić za pomocą rysunku 5.10. Górna część rysunku pokazuje kształtowanie się nominalnej stopy procentowej w zależności od poziomu inflacji, przy założeniu, że bank centralny stosuje regułę celu

[102] Ten konwencjonalny pogląd został w zasadniczy sposób podważony przez rozwój wydarzeń, który miały miejsce w krajach rozwiniętych w latach 2014–2015. W istocie, w pewnych bardzo szczególnych warunkach, nominalne stopy procentowe banku centralnego mogą być ujemne, podobnie zresztą jak poziom rentowności papierów wartościowych (zob. ramka 5.19).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

inflacyjnego i zmienia stopę procentową zgodnie z zasadą Taylora (zob. ramka 5.16). Prosta nachylona pod kątem 45° wyznacza linię zerowej realnej stopy procentowej, ponieważ w każdym jej punkcie nominalna stopa procentowa pozostaje równa stopie inflacji, a linia *ACD* jest linią reakcji banku centralnego na zmiany inflacji. Bank centralny wybiera pewien (dodatni) cel inflacyjny, na przykład na poziomie π, oraz spójną z tym celem nominalną stopę procentową i_A, zapewniającą realną stopę procentową równą odcinkowi *AB*. Gdy poziom inflacji przekracza cel inflacyjny ($p > \pi$), bank centralny zaostrza politykę pieniężną, podnosząc nominalne stopy procentowe powyżej przyrostu inflacji, tak aby wzrosły także *realne* stopy procentowe. Gdy inflacja spada natomiast poniżej celu inflacyjnego ($p < \pi$), bank łagodzi politykę pieniężną, obniżając nominalne stopy procentowe poniżej spadku inflacji, tak aby *realne* stopy procentowe także uległy obniżeniu.

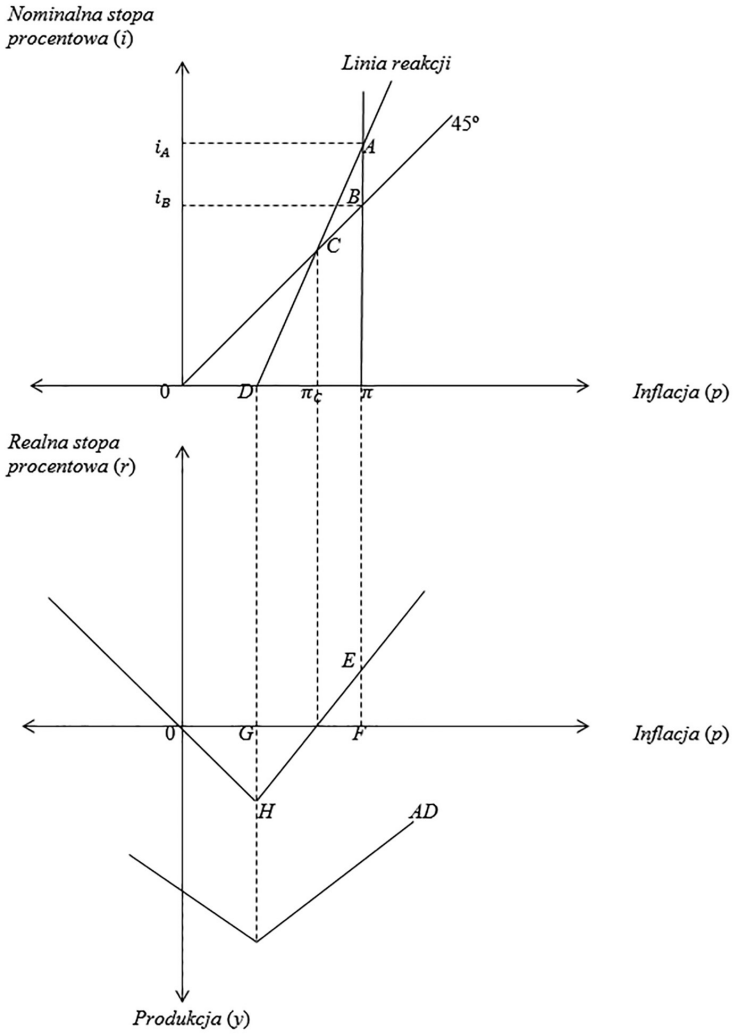
Gdy inflacja spada poniżej poziomu π_C, nominalne stopy procentowe zbliżają się do zera, a stopy realne stają się ujemne. Kształtowanie realnej stopy procentowej ukazane zostało w dolnej części rysunku 5.10. Począwszy od punktu *D*, dalsze obniżanie poziomu stóp nominalnych zasadniczo nie jest już możliwe. Wówczas dalszy spadek inflacji – poniżej poziomu π – wywołuje bowiem wzrost realnych stóp procentowych. Linia realnej stopy procentowej zmienia nachylenie z opadającego na rosnące. W punkcie *H* realna stopa procentowa jest maksymalnie ujemna, równa odcinkowi *GH*, lecz następnie, w miarę postępującego obniżania inflacji, zaczyna rosnąć w tempie równym spadkowi inflacji (nachylenie 45°).

Zmiany w zakresie realnej stopy procentowej przekładają się na zmiany produkcji poprzez wpływ na zagregowany popyt. Linia zagregowanego popytu *AD* jest w normalnych warunkach krzywą opadającą, ponieważ popyt rośnie wraz ze spadkiem inflacji i realnych stóp procentowych[103]. W punkcie *H* następuje jednak zmiana tej zależności i stopy realne zaczynają rosnąć pod wpływem spadku inflacji przy zerowej stopie nominalnej, a w konsekwencji następuje spadek produkcji. Jak pokazuje dolna część rysunku 5.10, krzywa *AD* zmienia nachylenie z ujemnego na dodatnie. Gdy nominalne stopy procentowe zbliżą się do zera, dalsze ich obniżanie stanie się niemożliwe (Taylor 2000). Osiągnięcie „zerowej granicy" (*zero lower bound*) sprawia, że konwencjonalna polityka pieniężna przestaje być skuteczna (Woodford 2012).

[103] Model AS-AD został omówiony w rozdziale 2.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 5.10. Wpływ niskiej inflacji na realne stopy procentowe, zagregowany popyt i produkcję



Źródło: opracowanie własne.

W tej sytuacji powstaje pytanie, co w tej sytuacji mogą jeszcze zrobić banki centralne, aby pobudzić wzrost gospodarczy i przybliżyć poziom inflacji do celu inflacyjnego. W normalnych warunkach bank centralny stosuje operacje otwartego rynku w celu utrzymania pożądanego poziomu rynkowej stopy procentowej, wielkość rezerw znajdujących się w posiadaniu banków komercyjnych ma zaś mniejsze

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

znaczenie. Gdy nominalne stopy procentowe są wyraźnie wyższe od zera, bank centralny może zawsze je obniżyć, jeśli chce poluzować politykę pieniężną. Jednakże kiedy stopy procentowe są już bliskie lub równe zeru, te konwencjonalne możliwości dalszego luzowania polityki wyczerpują się. Chociaż doświadczenia wyniesione z ostatniego kryzysu pokazały, że banki centralne mogą obniżyć stopy do poziomu ujemnego (zob. ramka 5.23), pole do takiego działania pozostaje w tym przypadku ograniczone.

Ramka 5.23. Ujemne nominalne stopy procentowe

Zawirowania na rynkach finansowych w okresie po kryzysie lat 2008–2009 postawiły pod znakiem zapytania niektóre z najbardziej oczywistych dogmatów finansowych. Dotyczy to między innymi możliwości wystąpienia ujemnych nominalnych stóp procentowych. Zgodnie z konwencjonalną teorią polityki pieniężnej, nominalne stopy procentowe na depozytach nie mogą być ujemne, ponieważ wówczas opłaca się gromadzić gotówkę zamiast depozytów bankowych. Podobnie nominalne stopy procentowe od pożyczek udzielanych przez bank centralny także nie mogą być ujemne, bo oznaczałoby to konieczność dopłacania do pożyczek udzielanych bankom komercyjnym. Oba twierdzenia wymagają gruntownej rewizji w świetle wydarzeń ostatnich lat.

Jak wiadomo, poczynając od 2013 roku, banki centralne w niektórych krajach wprowadziły ujemne stopy procentowe od jednodniowych depozytów. Na przykład 31 grudnia 2014 roku stopa oprocentowania krótkoterminowych depozytów trzymanych przez banki w banku centralnym wynosiła: w Danii –0,2%, w Szwajcarii –0,25%, w Szwecji –0,75%, a w strefie euro –0,2%. Wyjaśnienie tego zjawiska jest akurat proste. Alternatywa w postaci zamiany depozytów na gotówkę jest realistyczna tylko w przypadku detalicznych klientów, posiadających relatywnie niewielkie depozyty. Z kolei w odniesieniu do banków, które zarządzają często płynnymi środkami o równowartości setek milionów euro lub więcej, zamiana depozytów na pieniądz gotówkowy po prostu nie wchodzi w rachubę ze względu na ogromne koszty związane z transportem, przechowywaniem i zabezpieczeniem tak dużych zasobów gotówki. Również rezygnacja z korzystania z rachunków bankowych i przejście na gotówkowe rozliczenia codziennych transakcji pożyczkowych na rynku międzybankowym wiązałyby się z wysokimi kosztami. W takiej sytuacji banki są gotowe dopłacać do depozytów utrzymywanych w banku centralnym, aby

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

tych kosztów uniknąć. Ujemna stopa depozytowa staje się więc czymś w rodzaju opłaty za wygodę i możliwość bezpiecznego przechowywania środków płynnych. Ponadto, ujemne stopy depozytowe są konieczne w sytuacji, gdy podstawowa stopa refinansowa zostaje zredukowana do zera. Jak pamiętamy, stopa depozytowa pełni funkcję dolnego limitu przedziału wahań stóp procentowych na rynku międzybankowym. Aby zapewnić odpowiednią rozpiętość tego przedziału, stopa depozytowa zostaje więc obniżona do poziomu poniżej zera.

Na przełomie 2014 i 2015 roku niektóre banki centralne obniżyły również podstawowe stopy procentowe poniżej zera. W październiku 2014 roku bank centralny Szwecji jako pierwszy obniżył stopę procentową od operacji *repo* do poziomu –0,10%[104]. W jego ślady poszedł Szwajcarski Bank Narodowy, obniżając na początku stycznia 2015 roku podstawową stopę procentową (stopę *repo*) najpierw do przedziału zawierającego się między –0,75 a –0,25%, a następnie – od połowy miesiąca – do przedziału między –1,25 i –0,25%. Oznacza to, że bank centralny jest gotów *dopłacać* do krótkoterminowych pożyczek udzielanych bankom komercyjnym. Na przykład 5 lutego 2014 roku Szwajcarski Bank Narodowy kwotował stopę jednodniowych operacji *repo* (SARON) na poziomie –0,72%. Główną przyczyną tej nietypowej sytuacji była oczywiście przewlekła stagnacja w wielu krajach europejskich i próba zachęcenia banków komercyjnych do zwiększenia skali akcji kredytowej dla gospodarki poprzez ułatwienie im refinansowania się w banku centralnym. W obu sytuacjach – zarówno w przypadku depozytów, jak i operacji *repo* – ujemne nominalne stopy procentowe były składową wyjątkowo ekspansywnej polityki pieniężnej prowadzonej przez większość banków centralnych w latach 2013–2015.

Ujemne stopy procentowe pojawiły się także na rynku papierów wartościowych. Na przykład 5 lutego 2015 roku rynkowe rentowności 2-letnich rządowych obligacji niemieckich wynosiły –0,19%, 5-letnich obligacji niemieckich –0,05%, a 10-letnich rządowych obligacji szwajcarskich –0,12%. W ciągu 2015 roku ujemne rentowności na obligacjach o terminach zapadalności od jednego roku do pięciu lat notowane były także w wielu innych krajach europejskich[105]. Powstaje zatem pytanie, dlaczego inwestorzy są gotowi w pełni świadomie kupować papiery wartościowe o ujemnej dochodowości.

[104] 18 marca 2015 roku Rijksbank jeszcze bardziej obniżył podstawową stopę procentową, do poziomu –0,25%, starając się zapobiec umocnieniu korony szwedzkiej wobec euro w następstwie rozpoczęcia przez EBC operacji skupu obligacji rządowych na masowa skalę.

[105] M.in. w Holandii, Finlandii, Austrii, Belgii, Danii, Szwecji i Francji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Można wskazać co najmniej kilka przyczyn takiego postępowania. Jedną z nich jest brak bezpiecznej alternatywy do inwestowania ogromnych kwot, jakimi dysponują inwestorzy, zwłaszcza instytucjonalni (wielkie fundusze emerytalne, ubezpieczeniowe i inwestycyjne). Jak już zauważyliśmy, gromadzenie dużych zasobów gotówki okazuje się niepraktyczne i kosztowne, a z kolei kupowanie obligacji rządowych innych krajów jest w sytuacji przewlekłego kryzysu może być traktowane jako zbyt ryzykowne. Inny jeszcze powód stanowi deflacja, która na przełomie 2014 i 2015 roku zagościła w wielu krajach europejskich, w tym w strefie euro. Gdy ceny spadają, ujemne rentowności wciąż mogą zapewnić niewielką dodatnią stopę przychodu w kategoriach realnych. Jeszcze inną przyczyną takiego stanu rzeczy mogą być przewidywania aprecjacji waluty, w jakiej denominowane są obligacje. Jeśli inwestorzy spodziewają się umocnienia kursu euro lub franka szwajcarskiego, są gotowi akceptować ujemne rentowności, licząc na zyski z aprecjacji tych walut. Wreszcie też zapowiedź rozpoczęcia przez Europejski Bank Centralny z początkiem 2015 roku skupu aktywów finansowych na szeroką skalę w ramach tzw. luzowania ilościowego (QE) wywołała oczekiwanie wzrostu cen tych aktywów, w tym obligacji skarbowych. Antycypując te zakupy, inwestorzy nabywali więc obligacje, licząc na to, że będą mogli je korzystnie odsprzedać EBC w nieodległej przyszłości. Spowodowało to wzrost cen obligacji i dalszy spadek ich rentowności. Wszystkie wymienione przyczyny mają oczywiście charakter przejściowy – w perspektywie kilku lat nominalne stopy procentowe z pewnością powrócą do dodatnich poziomów.

Ogólnie rzecz biorąc, wydarzenia lat 2014–2015 w zasadniczy sposób podważyły ugruntowane od lat przekonanie o istnieniu „zerowej dolnej granicy" (*zero lower bound*) stóp procentowych. Okazało się, że nominalne stopy procentowe – i to zarówno stopy rynkowe, jak i stopy banku centralnego – mogą przyjmować wartości ujemne, i to przez dłuższy okres czasu (oczywiście, tym bardziej dotyczy to realnych stóp procentowych). Dolna granica prawdopodobnie nadal istnieje, lecz jest wyraźnie niższa od zera – być może nawet o 2–3 punkty procentowe[106].

Wyjaśnieniem tego paradoksu może być malejąca użyteczność gotówki w miarę wzrostu skali jej wykorzystania. Jak wspomniano, utrzymywanie płynności w gotówce

[106] Już pod koniec XIX wieku niemiecki ekonomista Silvio Gesell zauważył, że, gdy gospodarka znajduje się w tzw. pułapce płynności (*liquidity trap*), rząd może nałożyć „podatek" na zasoby płynnych środków utrzymywanych przez uczestników rynku, w postaci na przykład konieczności odpłatnego ostemplowania banknotów. Gesell chciał w ten sposób zniechęcić uczestników rynku do gromadzenia gotówki. W bardzo podobny sposób funkcjonują ujemne stopy procentowe nałożone przez bank centralny na płynne rezerwy banków komercyjnych. Zob. Buiter i Panigirtzoglou (2003).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przez uczestników rynku pociąga za sobą koszty związane z przechowywaniem, zabezpieczeniem i transportem jej zasobów. Jeśli koszty te wynoszą, przykładowo, 1% wartości posiadanych środków, to bezpieczną „dolną granicą" dla nominalnych stóp procentowych jest –1%, a nie zero. W praktyce trudno oszacować te koszty, pewne przybliżenie mogą stanowić jednak opłaty za stosowanie kart kredytowych, które umożliwiają dokonywanie transakcji bezgotówkowych. Opłaty te (uiszczane zarówno przez sprzedawcę, jak i przez konsumenta) przez wiele lat zawierały się w przedziale między 1% a 3% wartości transakcji[107]. Pokazuje to, na jakie koszty gotowi są godzić się konsumenci, aby uniknąć posługiwania się gotówką. Może to również oznaczać, że rzeczywista „dolna granica" nominalnych stóp procentowych mieści się właśnie między –1% a –3%. Dodatkowo warto zauważyć, że użyteczność gotówki maleje na skutek rozwoju elektronicznych form płatności i handlu internetowego.

5.14.2. Polityka luzowania ilościowego

Jednym z najbardziej znanych niekonwencjonalnych instrumentów polityki pieniężnej, po który sięgnęły niektóre banki centralne w okresie kryzysu finansowego, jest tzw. luzowanie ilościowe (*quantitative easing* – QE). Polega ono na bezwarunkowym nabyciu przez bank centralny papierów wartościowych od banków komercyjnych na określoną kwotę (są to operacje typu *outright*). Celem tej operacji jest zwiększenie zasobów pieniądza rezerwowego znajdującego się w posiadaniu banków i doprowadzenie, poprzez wykorzystanie mechanizmu mnożnika pieniężnego, do zwiększenia całej podaży pieniądza w gospodarce. Stosując luzowanie ilościowe, bank centralny nie dąży więc do obniżenia krótkoterminowych stóp procentowych, ponieważ z reguły są one już bliskie zeru. Polityka luzowania ilościowego wpływa natomiast na obniżenie długoterminowych stóp procentowych, ponieważ nabycie papierów wartościowych przez bank centralny wywołuje także wzrost rynkowych cen papierów wartościowych, a zatem także spadek ich rentowności. Ubocznym skutkiem stosowania polityki luzowania ilościowego jest również osłabienie kursu waluty narodowej, co poprzez przesunięcie popytu z dóbr zagranicznych na dobra krajowe dodatkowo stymuluje gospodarkę. Polityka luzowania

[107] Po 2012 roku wysokość tych opłat (tzw. *interchange fee*) w wielu krajach uległa obniżeniu wobec nałożenia ograniczeń administracyjnych przez władze publiczne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ilościowego polega więc na powiększaniu ogólnej sumy aktywów w bilansie banku centralnego i równoległym powiększaniu sumy pasywów poprzez nowe zobowiązania związane ze wzrostem sumy pieniądza rezerwowego o określoną wielkość (Marciniak 2015). Taką właśnie formę luzowania ilościowego zastosowano po raz pierwszy na początku XXI wieku w Japonii[108].

Luzowanie ilościowe odróżnia się czasem od luzowania jakościowego oraz od luzowania kredytowego. Luzowanie jakościowe (*qualitative easing*) polega na tym, że bank centralny zastępuje w swoim bilansie aktywa płynne, mniej ryzykowne, aktywami mniej płynnymi i bardziej ryzykownymi, lecz bez powiększania ogólnej sumy aktywów (Buiter 2008). W ten sposób bank centralny zmienia strukturę jakościową aktywnej strony swojego bilansu, ułatwiając bankom proces dostosowania ich struktury aktywów do bieżącej sytuacji. Z kolei celem luzowania kredytowego (*credit easing*) nie jest samo zwiększanie ilości pieniądza rezerwowego znajdującego się w posiadaniu banków, jak w przypadku luzowania ilościowego, a raczej powiększanie strony aktywnej bilansu banku centralnego poprzez zakup takich aktywów (papiery wartościowe, pakiety kredytów), które pozwolą uzyskać wzrost akcji kredytowej w całym sektorze bankowym. Tę właśnie formę luzowania ilościowego zastosowano w USA (zob. ramka 5.24)[109, 110]. Politykę luzowania ilościowego w strefie euro zdecydował się wprowadzić w 2015 roku także Europejski Bank Centralny (zob. ramka 5.25).

[108] Bank Japonii jako pierwszy bank centralny na świecie zastosował politykę luzowania ilościowego w marcu 2001 roku. Gdy krótkoterminowe stopy procentowe utrzymywały się od wielu miesięcy na poziomie bliskim zeru, a mimo to gospodarka pozostawała w stanie stagnacji, BoJ rozpoczął na wielką skalę skup obligacji rządowych i papierów wartościowych emitowanych pod zastaw innych aktywów, czyli tzw. ABS (*asset-backed securities*), i „zalewając" sektor bankowy tanią płynnością, liczył na to, że ułatwi to bankom wznowienie akcji kredytowej. W okresie 2001–2004 bank centralny zasilił banki komercyjne pieniądzem rezerwowym na łączną kwotę 300 mld USD. Polityka ta nie przyniosła jednak oczekiwanych rezultatów. Wzrost podaży pieniądza rezerwowego, będący rezultatem zastosowania polityki QE, nie przełożył się w dostatecznym stopniu na wzrost kredytów dla gospodarki oraz wzrost masy pieniądza ogółem m.in. z powodu spadku skłonności banków do podejmowania ryzyka na skutek strat kapitałowych poniesionych w poprzednich latach, braku opłacalnych projektów inwestycyjnych oraz słabego popytu na kredyty ze strony mocno zadłużonego sektora przedsiębiorstw (zob. Hiroshi, Okina i Shiratsuka 2001, s. 89–130).

[109] Zob. Bernanke (2009), http://www.federalreserve.gov/newsevents/speech/bernanke20090113a.htm (16.03.2010).

[110] Politykę QE zastosował również Bank Anglii, skupując w latach 2009–2012 aktywa od banków komercyjnych na łączną kwotę 375 mld GBP.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 5.24. Polityka QE w Stanach Zjednoczonych

Kryzys finansowy lat 2008–2009 wywołał wzrost zainteresowania niekonwencjonalnymi instrumentami polityki pieniężnej. Najszybciej zareagował amerykański Fed, który już w listopadzie 2008 roku rozpoczął zakupy papierów wartościowych, w tym papierów zabezpieczonych kredytami hipotecznymi (*mortgage-backed securities*) na kwotę 600 mld USD. Do marca 2009 roku suma aktywów Fed wzrosła z 900 USD dolarów do 1,75 biliona USD, a w czerwcu 2010 roku osiągnęła wartość 2,1 biliona USD. W kolejnych miesiącach kwota papierów wartościowych znajdujących się w posiadaniu Fed uległa obniżeniu, lecz, gdy po przejściowym ożywieniu gospodarka amerykańska ponownie zwolniła w 2010 roku, Fed wznowił operacje skupu aktywów. W ramach drugiej rundy luzowania ilościowego (QE2) bank centralny zgromadził obligacje rządu USA na łączną kwotę 600 mld USD w okresie od listopada 2010 do czerwca 2011. Celem było utrzymanie sumy aktywów na stabilnym poziomie 2,054 biliona USD[111].

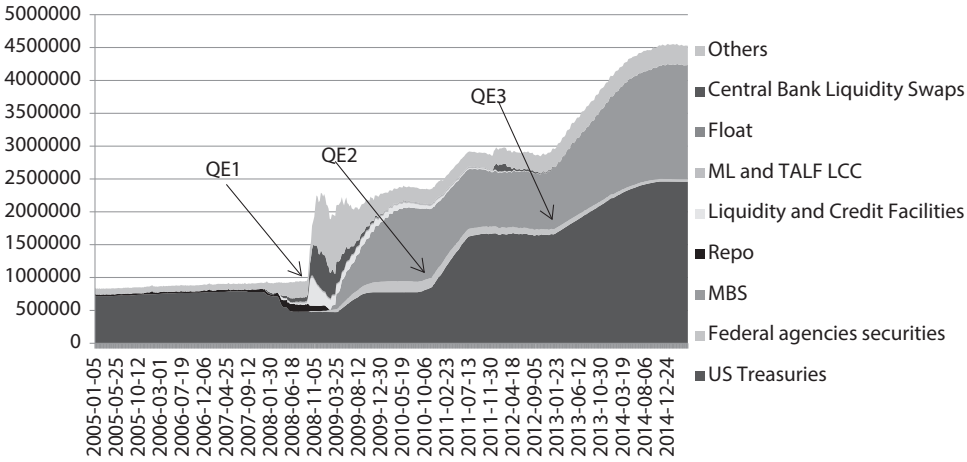
Trzecia runda (QE3) została uruchomiona we wrześniu 2012 roku. Fed zobowiązał się wówczas do kupowania aktywów w wysokości 40 mld USD miesięcznie, przy czym skup miał dotyczyć przede wszystkim obligacji i instrumentów dłużnych agencji rządowych finansujących rynek nieruchomości. W grudniu 2012 roku kwota miesięcznych zakupów została zwiększona do 85 mld USD. Jednocześnie Fed nie wyznaczył żadnych ram czasowych odnośnie do programu skupu, zobowiązując się jednak do utrzymania superniskich stóp procentowych „co najmniej do 2015 roku". W rezultacie suma aktywów w bilansie Fed przekroczyła na koniec 2013 roku 4 bln USD.

W czerwcu 2013 roku Prezes Rezerwy Federalnej Ben Bernanke zapowiedział stopniowe ograniczanie skali skupu aktywów (*tapering*) aż do całkowitego zaprzestania tych operacji w połowie 2014 roku. Rynki finansowe zareagowały jednak na te deklaracje z wielkim niepokojem, co zmusiło Fed do wycofania się z tych zapowiedzi i wydłużenia okresu stopniowego wygaszania luzowania ilościowego. Ostateczne zaprzestanie skupu aktywów nastąpiło dopiero w październiku 2014 roku.

[111] Poza bezwarunkowym skupem aktywów Fed realizował także w tym okresie tzw. operację TWIST, czyli zakup długoterminowych papierów i równoczesną sprzedaż krótkookresowych papierów w celu obniżenia długoterminowych stóp procentowych (na łączną kwotę 667 mld USD w okresie od września 2011 roku do grudnia 2012 roku). Dzięki operacji TWIST udało się zwiększyć zasoby płynnych aktywów w systemie bankowym i obniżyć ryzyko płynności (bez wzrostu podaży pieniądza).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 5.25. Zmiany w aktywach FED z okresu od stycznia 2005 roku do kwietnia 2015 roku[112] (dane w mln USD)



Objaśnienia: MBS (mortgage backed securities) oznaczają papiery zabezpieczone kredytami hipotecznymi, Repo to saldo operacji repo, Liquidity and Credit facilities, a także ML and TALF (*Term Asset-Backed Loan Facility*) oznaczają udzielone przez Fed pożyczki dla banków pod zastaw różnych instrumentów, a Float to środki pieniężne w obrocie.

Źródło: opracowanie własne na podstawie danych Fed.

Wykres 5.25 przedstawia zmiany w zakresie stanu aktywów Fed w latach 2005–2015. Jak widać, w wyniku trzech rund luzowania ilościowego w USA, między wrześniem 2008 roku a grudniem 2013 roku, suma aktywów w bilansie banku centralnego wzrosła prawie 4,5 razy. O ile w pierwszej fazie kryzysu dominowały operacje związane z zakupem papierów MBS, a także transakcje swapowe i bezpośrednie pożyczki dla banków, o tyle w drugiej i trzeciej rundzie Fed skupował przede wszystkim papiery skarbowe (*US Treasuries*).

Czy rzeczywiście luzowanie ilościowe przyniosło oczekiwaną poprawę sytuacji gospodarczej w USA? Badania przeprowadzone przez Bank Rezerwy Federalnej w San Francisco wskazują, że QE1 i QE2 spowodowały łączny spadek długoterminowych stóp procentowych o 15–20 punktów bazowych, obniżyły stopę bezrobocia w USA o 1,5 punktu procentowego i zwiększyły poziom PKB o 3%, przy czym zdecydowanie większa część tych zmian była skutkiem pierwszej rundy luzowania

[112] *Float* oznacza „pieniądz w ruchu", który liczony jest w systemie bankowym podwójnie ze względu na opóźnienie w realizacji czeków (pomiędzy momentem zaksięgowania kwoty czeku na rachunku odbiorcy a momentem obciążenia kwotą czeku rachunku wystawcy).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(Chen, Curdia i Ferrero 2012)[113]. Wraz z upływem czasu skuteczność polityki QE w USA wyraźnie malała. Na przykład QE1 obniżył stopę oprocentowania kredytu dla korporacji o 1 punkt procentowy i zwiększył poziom PKB o 2,8%, a QE2 – odpowiednio tylko o 13 punktów bazowych i 0,13% PKB. Jednocześnie badanie wykazało, że QE działa skuteczniej, gdy towarzyszy mu *forward guidance*, ponieważ wtedy rynki nie obawiają się odwrócenia kierunku polityki[114].

Przypomnijmy, że luzowanie ilościowe oddziałuje na gospodarkę realną kilkoma kanałami. Po pierwsze, skup obligacji wywołuje wzrost ich cen i obniża poziom rentowności. To z kolei prowadzi do spadku wartości długoterminowych stóp procentowych i zmniejsza koszt pozyskiwania kapitału na rynku przez firmy, banki, a także przez sektor publiczny[115]. Po drugie, występuje efekt równoważenia portfela aktywów przez banki (*portfolio rebalancing*), ponieważ, gdy papiery wartościowe zostaną odsprzedane bankowi centralnemu, banki komercyjne inwestują pozyskane środki w inne aktywa, w tym obligacje, kredyty i instrumenty udziałowe, co dodatkowo ułatwia dostęp do kapitału sektorowi realnemu. Co więcej, spadek stóp procentowych powoduje odpływ kapitału za granicę i osłabia kurs waluty krajowej.

Polityka luzowania ilościowego pociąga jednak za sobą także określone ryzyko. Skup aktywów zwiększa nadpłynność banków i czyni je bardziej niezależnymi od banku centralnego. Może to w przyszłości utrudnić zaostrzenie polityki pieniężnej i konieczną walkę z inflacją (Marciniak 2015). Wielu ekonomistów uważa też, że polityka QE zmniejsza presję na przeprowadzenie koniecznych reform strukturalnych i konsolidację finansów publicznych, ponieważ ułatwia rządowi zaciąganie kolejnych pożyczek na rynku finansowym. Wywołuje to wzrost długu publicznego, co może pociągać za sobą wyższe stopy procentowe i wolniejszy wzrost gospodarczy, a w przyszłości jest także w stanie zmusić władze do podniesienia podatków.

[113] Polityka QE została również zastosowana w Wielkiej Brytanii. Wyniki badań przeprowadzonych w Banku Anglii wskazują, że skup aktywów na kwotę 200 mld GBP zwiększył brytyjski PKB o 2% oraz inflację o 1,5 punktu procentowego. Odpowiada to obniżce głównych stóp procentowych o 3 punkty procentowe.

[114] Również badania przeprowadzone przez Międzynarodowy Fundusz Walutowy potwierdzają, że polityka QE pozwoliła ograniczyć ryzyko załamania i ustabilizować sytuację w systemach finansowych wśród krajów rozwiniętych (zob. Klyuev, de Imus i Srinivasan 2009).

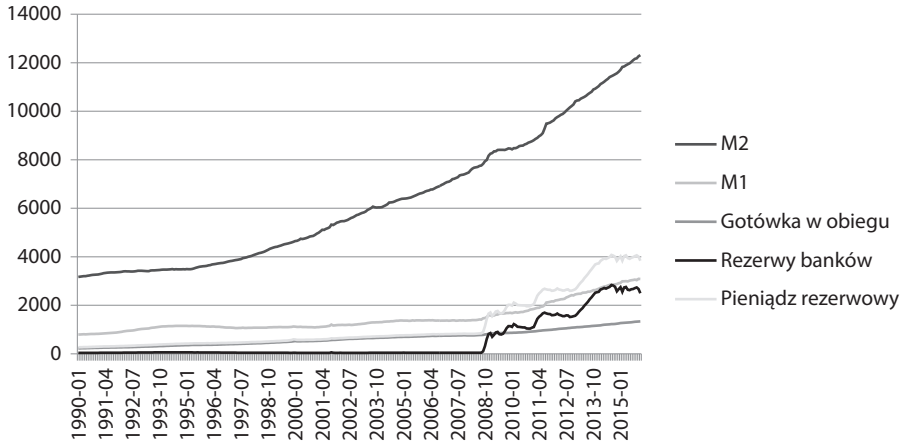
[115] Odmianą polityki QE jest także nabywanie papierów wartościowych w dysfunkcjonalnych segmentach rynku finansowego (np. na rynku długu państw zadłużonych), mające doprowadzić do pożądanego poluzowania polityki pieniężnej w tych segmentach. Zakup taki może być sterylizowany lub nie – w zależności od konkretnych zamiarów banku centralnego co do zmian całkowitej podaży pieniądza.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ogólnie rzecz biorąc, przy niskich stopach procentowych i obfitości finansowania dłużnicy są mniej skłonni do oddłużania się i restrukturyzacji, co zdaniem niektórych ekonomistów może stanowić istotny hamulec wzrostu, gdy koniunktura ulegnie poprawie (Rzońca 2014).

Polityka QE może też zniekształcać bodźce rynkowe i zachęcać do spekulacji, ponieważ na skutek masowego wzrostu płynności rynkowe stopy procentowe znajdują się poniżej poziomu trwałej równowagi. Sztucznie zaniżone stopy procentowe i łatwa dostępność kredytu skłaniają wówczas do inwestowania w aktywa nieprodukcyjne, takie jak na przykład nieruchomości, lub wywołują nadmierny wzrost cen papierów wartościowych na giełdach. Wskazuje się także, że polityka QE realizowana w danym momencie wymusza działania odwrotne w (nieodległej) przyszłości, co oznacza zacieśnienie polityki pieniężnej na jakiś czas. Jeśli QE jest traktowane przez uczestników rynku jako polityka tymczasowa, przejściowa, to perspektywa jej odwrócenia i zaostrzenia osłabia pożądane efekty luzowania. Racjonalnie zachowujący się uczestnicy rynku powstrzymają się bowiem w takiej sytuacji od zwiększania wydatków, w tym zwłaszcza w zakresie inwestycji.

Wykres 5.26. Zmiany podaży agregatów pieniężnych w USA w latach 1990–2015 (dane w mld USD)



Źródło: opracowanie własne na podstawie danych Fed.

Szczególnie kontrowersyjnym zagadnieniem jest potencjalny wpływ luzowania ilościowego na inflację. Rzeczywiście, definitywny skup papierów wartościowych od banków komercyjnych przez bank centralny powoduje wzrost podaży pieniądza

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

rezerwowego. Lecz nie musi to samo z siebie wywoływać zwiększonej inflacji, ponieważ pozostaje ona zależna od *całkowitej* podaży pieniądza. Tak długo zatem, jak wzrost podaży pieniądza rezerwowego nie przekłada się na wzrost podaży pieniądza ogółem, inflacja nie będzie rosła. Tymczasem wzrost szerszych agregatów pieniężnych zależy nie tylko od wzrostu podaży pieniądza rezerwowego, lecz także od tego, czy banki komercyjne zechcą zwiększyć akcję kredytową. Tak długo, jak nie następuje wzrost akcji kredytowej, nie dojdzie również do wzrostu wydatków, a zatem nie może być również mowy o wzroście inflacji. Jest ona bowiem możliwa tylko w warunkach wzrostu faktycznych wydatków finansowanych kredytem. W normalnych warunkach wzrost podaży pieniądza rezerwowego przekłada się na zwiększenie kredytu, podczas gdy w warunkach kryzysu i utrzymującej się „zapaści kredytowej" (*credit crunch*) mnożniki pieniężne maleją.

Potwierdzają to dane dotyczące zmian poziomu agregatów pieniężnych w USA. Kolosalny wzrost podaży pieniądza rezerwowego nie przełożył się proporcjonalnie na wzrost szerszych agregatów pieniężnych. Wykres 5.26 odzwierciedla zmiany agregatów pieniężnych w latach 1990–2015. Po pierwsze, widać wyraźnie, że w kształtowaniu podaży pieniądza rezerwowego zaszła fundamentalna zmiana po wybuchu kryzysu finansowego we wrześniu 2008 roku. Długookresowy trend uległ całkowitemu zakłóceniu na skutek stosowania przez Fed polityki agresywnego zwiększania płynności w systemie bankowym. Dobrze widoczne na wykresie są trzy tury luzowania ilościowego w przebiegu zmian podaży pieniądza rezerwowego i rezerw banków po 2008 roku. Po drugie, wzrost podaży pieniądza rezerwowego był prawie w całości spowodowany wzrostem stanu rezerw na rachunkach banków komercyjnych w systemie rezerwy federalnej, natomiast tempo wzrostu podaży pieniądza gotówkowego praktycznie nie uległo zmianie w porównaniu z poprzednimi latami. Wreszcie, po trzecie, podaż najszerszego agregatu, M2, praktycznie nie zareagowała na skokowy wzrost podaży pieniądza rezerwowego[116]. Oznacza to, że – przynajmniej w okresie pierwszych kilku lat – polityka luzowania ilościowego nie wywołała nadmiernego wzrostu całkowitego zasobu pieniądza w obiegu. Leniwie reagowała również podaż agregatu M1. Po 2008 roku doszło nawet do zupełnie nietypowej sytuacji, kiedy podaż pieniądza rezerwowego przewyższała podaż agregatu M1. Oznacza to, że płynne rezerwy banków komercyjnych były większe niż wszystkie depozyty bieżące zgromadzone w całym sektorze

[116] Tempo wzrostu podaży pieniądza rezerwowego wzrosło z 6,4% w latach 1990–2007 do 28,5% w latach 2008–2013, z kolei tempo wzrostu agregatu M2 wynosiło wówczas odpowiednio 4,9% oraz 6,5%.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

bankowym. W końcu 2015 roku wielkość pieniądza rezerwowego przewyższała wielkość M1 o ponad 700 mld USD.

Brak bezpośredniego zagrożenia wzrostem inflacji w wyniku polityki QE nie oznacza jednak, że luzowanie ilościowe jest z tego punktu widzenia całkowicie obojętne. Duże zasoby płynnych środków zgromadzone w bankach mogą zostać po pewnym czasie wykorzystane w celu zwiększenia akcji kredytowej, gdy polepszą się perspektywy wzrostu gospodarczego i poprawie ulegną nastroje rynkowe. Nadwyżkowa płynność zmniejsza także potrzeby banków w zakresie zasilania się w płynność w banku centralnym. Wówczas bank centralny może stanąć w obliczu konieczności silnego podnoszenia stóp procentowych – ponad pożądany poziom równowagi, aby hamować nadmierną emisję kredytu. Może to w przyszłości utrudnić zaostrzenie polityki pieniężnej i konieczną walkę z inflacją.

Wykres 5.26 pozwala także odnieść się do często pojawiającego się w mediach przekazu, mówiącego o tym, że polityka QE oznacza drukowanie pieniądza[117]. Należy przypomnieć, że bank centralny, skupując papiery wartościowe od banków komercyjnych, zapisuje należne im kwoty na ich rachunkach w banku centralnym. Wywołuje to wzrost sumy aktywów i pasywów banku centralnego, lecz nie oznacza jeszcze automatycznego druku pieniądza, ponieważ ilość gotówki w obiegu nie ulega zmianie. Precyzując, bank centralny tworzy w ten sposób pieniądz rezerwowy w postaci zapisów na rachunkach banków, lecz nie zwiększa podaży gotówki w fizycznej postaci banknotów i monet. Nie następuje więc „druk" w dosłownym znaczeniu tego słowa[118]. W okresie między sierpniem 2008 roku a grudniem 2013 roku stan rezerw bankowych w USA wzrósł z 47 mld USD do 2541 mld USD (czyli ponad 50-krotnie), z kolei zasób gotówki znajdującej się w obiegu wzrósł w tym samym czasie z 777 mld USD do 1160 mld USD, a więc tylko o 50%.

Dzięki utrzymywaniu większych rezerw banki komercyjne mogą, ale nie muszą, poszerzać akcję kredytową. Lecz nawet jeśli się tak stanie, nie oznacza to jeszcze automatycznego druku pieniędzy – dopiero, gdy udzielony kredyt pociąga za sobą zwiększone wypłaty gotówkowe z banków, rozpoczyna się dodatkowy druk pieniądza. Tak więc stosowanie polityki QE wymaga dodruku pieniądza dopiero wówczas,

[117] Na przykład 29 października 2014 roku „Gazeta Wyborcza", informując o stopniowym wycofywaniu się centralnego banku USA z polityki luzowania ilościowego, donosiła w tytule, że „Fed zaprzestaje dodruku dolarów".

[118] Zarazem trzeba pamiętać o tym, że rezerwy bankowe mogą być w każdym momencie wykorzystane przez banki do dokonania wypłat gotówkowych dla klientów. W tym sensie rezerwy pozostają równie płynne jak gotówka.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gdy banki zwiększą akcję kredytową oraz kiedy w wyniku udzielenia tych dodatkowych kredytów wzrasta popyt na gotówkę.

Ramka 5.25. Polityka QE w strefie euro

22 stycznia 2015 roku Europejski Bank Centralny zapowiedział rozpoczęcie programu mającego na celu skup aktywów od banków komercyjnych na masową skalę. Uzupełniając wcześniejsze programy skupu papierów wartościowych zabezpieczonych aktywami (*Asset-Backed Securities Purchase Programme* – ABSPP) oraz skupu zabezpieczonych obligacji bankowych (*Covered Bonds Purchase Programme* – CBPP), EBC poinformował o planowanym uruchomieniu programu mającego na celu skup rządowych papierów wartościowych (*Public Sector Purchase Programme* – PSPP). W ramach PSPP Eurosystem (obejmujący EBC oraz banki centralne państw członkowskich strefy euro) ma nabywać papiery rządowe państw członkowskich oraz papiery emitowane przez instytucje europejskie i agencje rządowe o łącznej wartości 60 mld EUR miesięcznie. EBC dołączył w ten sposób do grupy banków centralnych USA, Wielkiej Brytanii i Japonii, które stosowały politykę QE w związku z kryzysem finansowym i gospodarczym lat 2008–2012. EBC ogłosił również, że zakupy prowadzone w ramach PPSP będą kontynuowane co najmniej do września 2016 roku, lecz mogą zostać przedłużone, jeśli inflacja będzie utrzymywać się nadal poniżej celu średniookresowego. Oznacza to, że w istocie program ten ma charakter „otwarty" (*open-ended*)[119]. EBC rozpoczął operacje skupu w marcu 2015 roku.

W ramach miesięcznej kwoty 60 mld EUR EBC przeznacza 10 mld EUR na zakup zabezpieczonych obligacji bankowych (*covered bonds*), kontynuując program CBPP uruchomiony w październiku 2014 roku. Pozostałe 50 mld EUR obejmuje zakupy w ramach programu PSPP, z czego 6 mld EUR (12%) przeznaczone jest na zakup obligacji emitowanych przez ponadnarodowe instytucje europejskie mające siedzibę w strefie euro[120], a 44 mld EUR (88%) pozwala na zakup obligacji rządowych wyemitowanych przez rządy państw strefy euro oraz wybrane

[119] Draghi (2015).

[120] European Financial Stability Facility, European Investment Bank, European Stability Mechanism, European Union, European Atomic Energy Community, Council of Europe Development Bank oraz Nordic Investment Bank.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

agencje rządowe[121], przy czym EBC przypada do rozdysponowania 4 mld EUR, a narodowym bankom centralnym 40 mld EUR (Claeys, Leandro i Mandra 2015).

5.14.3. Pieniądze z „helikoptera" (*helicopter money*)

„Money for nothing and cheques for free" – śpiewała grupa rockowa Dire Straits w latach 80. Czy to, co stanowiło wówczas tylko wyraz niemożliwych do spełnienia marzeń brytyjskich artystów, może istotnie stać się rzeczywistością? Czy pieniądze mogą być rozdawane za darmo? Dziś okazuje się, że nie jest to wykluczone, a to za sprawą tak zwanych pieniędzy z „helikoptera".

Termin *helicopter money* spopularyzował Milton Friedman w swoim głośnym eseju na temat optymalnej ilości pieniądza w obiegu (Friedman 1969). Jest to metaforyczny opis szczególnej odmiany polityki pieniężnej, która polega po prostu na rozdawaniu uczestnikom rynku pieniądza gotówkowego przez bank centralny. Autor rozpatruje przypadek hipotetycznego społeczeństwa, którego nominalny produkt roczny wynosi 10 000 USD, z czego jedna dziesiąta – czyli 1 000 USD – utrzymywana jest w formie pieniądza (gotówkowego), a następnie analizuje efekt jednorazowego zwiększenia podaży pieniądza o 100%. Stosowny fragment tekstu Friedmana brzmi następująco: „Let us suppose now that one day a helicopter flies over this community and drops an additional $ 1,000 in bills from the sky, which is, of course, hastily collected by members of the community. Let us suppose further that everyone is convinced that this is a unique event which will never be repeated... (s. 4–5)".

Przyjmując szereg upraszczających założeń, autor eseju pokazuje, że w opisanej sytuacji nie istnieją żadne szczególne powody, dla których popyt na pieniądz gotówkowy miałby wzrosnąć. Dlatego należy przyjąć, że członkowie danego społeczeństwa wydaliby natychmiast dodatkowe pieniądze na przejściowe zwiększenie konsumpcji. Jedynym trwałym skutkiem jednorazowego wzrostu podaży pieniądza okazałby się zatem wzrost cen. Ponieważ dzięki „helikopterowi" podaż pieniądza zwiększa się dwukrotnie, a popyt na pieniądz pozostaje stały, ceny muszą wzrosnąć także dwukrotnie, w związku z czym całkowity produkt roczny społeczeństwa w cenach nominalnych wzrośnie do 20 000 USD.

[121] Do agencji tych zaliczono: Instituto de Credito Oficial (Spain), KfW (Germany), Landeskreditbank Baden-Württemberg Foerderbank (Germany), Landwirtschaftliche Rentenbank (Germany), NRW.Bank (Germany), CADES (France) oraz UNEDIC (France).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Przypadek omówiony przez Friedmana dowodzi, że wzrost nominalnej podaży pieniądza wywołuje – w ramach przyjętych założeń – wzrost bieżących wydatków, ponieważ ludzie starają się utrzymać rezerwy kasowe na dotychczasowym poziomie, czyli w tej samej proporcji do stałego dochodu. Ta właśnie zależność dostarcza uzasadnienia w przypadku zastosowania polityki luzowania ilościowego. Stosując QE, bank centralny zwiększa rezerwy płynne znajdujące się w posiadaniu banków komercyjnych, licząc na to, że banki odpowiednio dostosują strukturę swojego portfela aktywów, a więc zwiększą zakupy papierów wartościowych i – przede wszystkim – emisję kredytu do gospodarki. To z kolei ma pozwolić przedsiębiorstwom i gospodarstwom domowym na podniesienie wydatków.

Istnieją jednak dwie ważne różnice między polityką *helicopter money* a luzowaniem ilościowym. Pierwsza z nich dotyczy wpływu na bilans banku centralnego. W przypadku QE wartość bilansu banku wzrasta po stronie aktywów o kwotę zakupionych papierów wartościowych, będącą dokładnie równą kwocie przyrostu podaży pieniądza rezerwowego po stronie pasywów. Z kolei w ramach funkcjonowania *helicopter money* zwiększenie zobowiązań po stronie pasywów banku centralnego w postaci przyrostu masy pieniądza gotówkowego w obiegu jest dokładnie kompensowane jednoczesnym spadkiem wielkości pasywów na skutek obniżenia wyniku finansowego banku centralnego. Oznacza to, że „rozdawanie" pieniędzy jest transferem bogactwa, który musi zostać w całości pokryty z wyniku finansowego banku centralnego, czyli z bieżącego zysku (lub z rezerw, jeśli bank centralny wykazuje stratę). Jeśli tych środków jest za mało, różnicę musi pokryć właściciel banku centralnego, czyli skarb państwa, reprezentowany przez rząd. Lecz w takim razie deficyt powstały w bilansie banku centralnego staje się częścią deficytu sektora publicznego, który w ostatecznym rachunku będzie musiał być sfinansowany przez przyszłe podatki. W takiej sytuacji pieniądze z „helikoptera" pozwolą na wzrost rzeczywistych wydatków tylko w takim stopniu, w jakim nie występuje w gospodarce „równoważność rikardiańska"[122].

Druga różnica pozostaje równie istotna. Polityka QE wpływa na udział pieniądza w bilansach uczestników rynku nie bezpośrednio, ale za pośrednictwem banków. Skala wzrostu ilości pieniądza znajdującego się w posiadaniu uczestników rynku zależy – zgodnie z mechanizmem mnożnika – od skłonności banków

[122] Jak wskazaliśmy w rozdziale 3 (punkt 3.9.6), efekt „rikardiański" występuje wśród państw europejskich jedynie w ograniczonym zakresie. Polityka *helicopter money* byłaby więc pod tym względem skuteczna.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

do zwiększania kredytu. W warunkach dużej niepewności i bardzo niskich stóp procentowych, charakterystycznych dla okresów kryzysowych, banki zmieniają strukturę swojego portfela na rzecz wzrostu udziału środków płynnych. Oznacza to, że banki zwiększają popyt na pieniądz rezerwowy i absorbują zwiększoną podaż – na tym polega wspomniana wcześniej „pułapka płynności" (*liquidity trap* – zob. punkt 5.14.1). W takiej sytuacji do pozostałych uczestników rynku dociera tylko niewielka część dodatkowej płynności, w związku z czym pośrednictwo systemu bankowego w przekazywaniu płynnych środków do uczestników rynku zawodzi. W rezultacie polityka QE staje się mało skuteczna w pobudzaniu wydatków.

Pieniądze „zrzucane z helikoptera" i trafiające bezpośrednio do uczestników rynku, mogą w takiej sytuacji okazać się bardziej skutecznym sposobem stymulowania popytu i poprawy koniunktury. Dzieje się tak w sytuacji, gdy na skutek zwiększonej niepewności co do rozwoju sytuacji gospodarczej w przyszłości, banki niechętnie udzielają kredytów. Zasilanie banków w płynność przez bank centralny nie przekłada się w takich warunkach na wzrost akcji kredytowej – na rynku występuje bowiem zacieśnienie kredytowe (*credit crunch*). Wówczas zamiast zasilać banki w dodatkową płynność, można przekazać ją bezpośrednio do uczestników rynku. W szczególności gospodarstwa domowe cechują się stosunkowo wysoką skłonnością do ponoszenia wydatków z bieżącego dochodu. Co więcej, stojąc w obliczu ostrzejszych ograniczeń budżetowych w okresie kryzysu i zarazem nie obawiając się, że nastąpi wzrost podatków (efekt rikardiański), będą bardziej skłonne przeznaczyć dodatkowe środki na bieżącą konsumpcję (zgodnie z zasadą wygładzania konsumpcji w czasie). Może to spowodować – zgodnie z sugestią Friedmana – przejściowy wzrost popytu i wydatków, zapewne na znacznie większą skalę niż polityka QE (zob. Buiter 2014).

Jakie warunki muszą zostać spełnione, aby pieniądze z „helikoptera" stały się skuteczniejszym instrumentem pobudzania popytu niż luzowanie ilościowe? Jeśli mnożnik pieniężny jest wysoki, to system bankowy pozostaje bardzo efektywnym mechanizmem zwiększania podaży „szerokiego" pieniądza. Na przykład, gdy mnożnik M3 wynosi 10, zwiększenie podaży bazy monetarnej o jednostkę wywołuje wzrost podaży M3 o 10 jednostek. Lecz jeśli mnożnik jest niski, wzrost kredytów, a w konsekwencji wzrost agregatu M3, okaże się niewielki. Tak też dzieje się, gdy banki przechowują większość płynnych środków na rachunkach w banku centralnym.

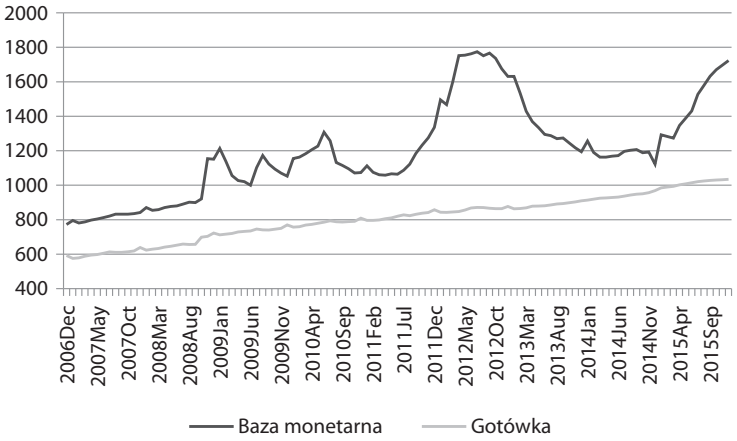
Może się wydawać, że z taka właśnie sytuacja miała miejsce w strefie euro podczas kryzysu zadłużenia. Wykres 5.7 (punkt 5.2.2) pokazuje, że wartość mnożnika pieniądza M3 spadła w latach 2008–2012 niemal o połowę, tj. z około 10,0–10,2 do

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

około 5,5. Wzrost wartości mnożnika od połowy 2012 roku do końca 2014 roku do poziomu około 8,0–9,0 jednostek był głównie skutkiem ustabilizowania sytuacji na rynku długu suwerennego w związku z ogłoszeniem programu OMT, a także rozpoczęcia przez EBC skupu aktywów w ramach programu ABSPP i CBPP (zob. ramka 5.25). W 2015 roku wartość mnożnika M3 zaczęła jednak ponownie spadać i pod koniec roku wynosiła około 6,2, co świadczy o wciąż niewielkiej skuteczności polityki luzowania ilościowego rozpoczętej w marcu 2015 roku.

Z kolei wykres 5.27 ilustruje zmiany poziomu bazy monetarnej na tle zmian zasobu gotówki znajdującej się w obiegu. Przebieg zmian bazy monetarnej jest lustrzanym odbiciem zmian wartości mnożnika. Wynika z tego, że posiadane nadwyżki płynności banki lokowały na rachunkach w EBC – stan tych rachunków (różnica między wielkością bazy monetarnej a wielkością gotówki w obiegu) zwiększył się z około 200 mld EUR w latach poprzedzających kryzys do niemal 1000 mld EUR w połowie 2012 roku, by po przejściowym spadku w latach 2013–2014 znowu wzrosnąć do około 600 mld EUR. Oznacza to, że wysiłki EBC, mające na celu zwiększenie zasilania banków strefy euro w płynność, podejmowane w latach 2009–2015 nie miały praktycznie wpływu na wzrost zasobu gotówki znajdującej się w obiegu, co pośrednio dowodzi, że kredyty i wydatki rosły w tym okresie bardzo wolno.

Wykres 5.27. Baza monetarna i gotówka w obiegu w strefie euro w okresie od grudnia 2006 roku do grudnia 2015 roku (dane w mld EUR)



Uwaga: ze względu na brak niektórych danych miesięcznych w przypadku 2015 roku (luty, maj, sierpień, listopad) brakujące obserwacje oszacowano metodą interpolacji.

Źródło: opracowanie własne na podstawie danych EBC.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Odmianą *helicopter money* jest monetyzacja deficytu budżetowego, czyli finansowanie deficytu bezpośrednio dodrukiem pieniądza przez bank centralny. W tym przypadku bank centralny „rozdaje" pieniądze rządowi, który następnie przeznacza je na dodatkowe wydatki lub na obniżanie podatków. Jak wiadomo, w większości krajów prawo zakazuje stosowania tego typu praktyk, które w przeszłości wielokrotnie doprowadzały do galopującej inflacji. Niemniej, monetyzacja może mieć ekonomiczny sens w pewnych szczególnych sytuacjach. Szef amerykańskiego Urzędu Rezerwy Federalnej Ben Bernanke nawiązał do tej formy *helicopter money* w kontekście dyskusji o sposobach przełamania deflacji w Japonii na początku pierwszej dekady XXI wieku[123]. Hiszpański ekonomista Jordi Galí wykazał, że finansowanie deficytu budżetowego bezpośrednim dodrukiem pieniądza może być bardzo skuteczną metodą stymulowania gospodarki przy niewielkim ryzyku inflacji (Galí 2014).

Kłopoty ze skutecznym pobudzaniem wzrostu gospodarczego w strefie euro i podniesieniem inflacji z bardzo niskich poziomów w okolice celu inflacyjnego spowodowały, że Europejski Bank Centralny zainteresował się w 2016 roku ideą „*helicopter money*", mimo że w żadnym kraju na świecie polityka ta nie była dotychczas stosowana[124]. Przewlekła stagnacja gospodarcza w strefie euro, znacznie dłuższa niż w USA, była między innymi skutkiem niedostatecznie szybkiej reakcji ze strony EBC w podejmowaniu niekonwencjonalnych działań (takich jak QE), co z kolei wynikało z politycznych ograniczeń, w ramach których musi funkcjonować ta instytucja. Zastosowanie *helicopter money* wymagałoby prawdopodobnie zmiany istniejących traktatów europejskich, co politycznie może okazać się niewykonalne (na polityczne aspekty zwraca uwagę Turner 2015).

Idea *helicopter money* może być – co zrozumiałe – bardzo atrakcyjna dla niektórych polityków i ekonomistów o bardziej lewicowych poglądach. Uzasadnieniem w przypadku zastosowania tej niekonwencjonalnej formy polityki jest przekonanie, że „rozdawanie" pieniędzy bezpośrednio między członków społeczeństwa może znajdować uzasadnienie na gruncie sprawiedliwości społecznej, gdyby postrzegać

[123] W przemówieniu z 21 listopada 2002 roku Ben Bernanke nawiązał do eseju Friedmana i jego pomysłu dostarczania „pieniędzy z helikoptera", sugerując, że rząd japoński zawsze może dowolnie zwiększyć ilość pieniądza w obiegu, ponieważ ma do swojej dyspozycji maszyny drukarskie, i w ten sposób może podnieść inflację. Po tej wypowiedzi krytycy nadali mu przydomek „helikopterowy Ben".

[124] Prezes EBC Mario Draghi nie wykluczył możliwości wykorzystania *helicopter money* w strefie euro, określając tę koncepcję podczas konferencji prasowej z 10 marca 2016 roku jako „bardzo interesującą". Draghi poinformował zarazem, że w EBC prowadzone są prace analityczne nad możliwością zastosowania tej polityki w praktyce.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

je jako podział przyszłego dochodu z emisji (seniorat) w formie swoistej „obywatelskiej dywidendy”[125].

5.14.4. Polityka „wytycznych na przyszłość” (*forward guidance*)

Polityka *forward guidance*, która w języku polskim może być tłumaczona jako polityka „wytycznych na przyszłość” (lub polityka „wskazywania kierunku”), jest nowym instrumentem polityki pieniężnej, polegającym na komunikowaniu uczestnikom rynku zamiarów oraz intencji banku centralnego w odniesieniu do planowanych zmian w polityce pieniężnej[126]. *Forward guidance* jest stosowany przez banki centralne w celu ograniczenia niepewności co do rozwoju sytuacji gospodarczej w nadchodzących latach i zwiększenia przejrzystości polityki pieniężnej oraz kształtowania pożądanych oczekiwań co do przyszłego poziomu stóp procentowych. Ma to szczególne znaczenie w okresach zawirowań i kryzysów na rynkach finansowych, kiedy spada zaufanie i rośnie ryzyko zaburzeń w funkcjonowaniu mechanizmów dostarczania płynności do systemu bankowego. W warunkach dużej luki podażowej oraz niskiej i stabilnej inflacji bank centralny obniża silnie stopy procentowe. Lecz przy nominalnych stopach procentowych bliskich zeru konwencjonalna polityka pieniężna przestaje być skuteczna. *Forward guidance*, utrwalając oczekiwania co do utrzymania zerowych stop procentowych w przyszłości w warunkach rosnącej inflacji, może w tych warunkach stać się instrumentem dalszego łagodzenia polityki pieniężnej.

Stosowanie *forward guidance* oznacza znaczące odstępstwo od tradycyjnej postawy banku centralnego, cechującej się skłonnością do unikania wiążących zobowiązań co do przyszłego postępowania. Ze względu na niepewność przyszłości bank centralny w zasadzie woli nie wiązać sobie rąk i pozostawić możliwie szeroki zakres swobody decyzji. Zdarzają się jednak okresy, kiedy niepewność co do

[125] W wywiadzie udzielonym dla włoskiego dziennika członek zarządu EBC Peter Praet, zapytany o możliwość zastosowania *helicopter money*, oświadczył: „Yes, all central banks can do it. You can issue currency and you distribute it to people. That's helicopter money. Helicopter money is giving to the people part of the net present value of your future seigniorage, the profit you make on the future banknotes. The question is, if and when is it opportune to make recourse to that sort of instrument which is really an extreme sort of instrument”. “La Repubblica”, 18 March 2016, http://www.ecb.europa.eu/press/inter/date/2016/html/sp160318.en.html (12.04.2016). Więcej na ten temat na portalu kampanii na rzecz *Quantitative Easing for People* (www.qe4people.eu).

[126] Termin *forward guidance* nie ma jeszcze ustalonego odpowiednika w języku polskim. W niniejszej książce tłumaczymy go jako „wytyczne na przyszłość”, ale możliwe jest też tłumaczenie oznaczające „wskazywanie kierunku”. W dalszym tekście publikacji będziemy posługiwać się zamiennie terminem angielskim i jego polskim tłumaczeniem.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przyszłych przedsięwzięć banku centralnego silnie ogranicza inwestycje i hamuje działalność gospodarczą. Pokazał to kryzys finansowy i gospodarczy w latach 2009–2013, kiedy nawet bardzo swobodna polityka pieniężna nie była w stanie przełamać stagnacji gospodarczej z powodu utrzymującego się braku zaufania i braku pespektyw wzrostu.

Stosując *forward guidance*, bank centralny zobowiązuje się publicznie do utrzymywania zerowych (lub bliskich zeru) stóp procentowych przez okres dłuższy niż byłoby to w normalnych warunkach uzasadnione. Jeśli taka deklaracja jest wiarygodna, wpływa ona na obniżenie oczekiwań co do przyszłego poziomu realnych stóp procentowych, a to zwiększa skłonność uczestników rynku do dokonywania wydatków (Krugman 1998, Eggertsson i Woodford 2003, Werning 2012). Polityka *forward guidance* może występować w dwóch odmianach. Ta pierwsza polega na bardzo silnym zobowiązaniu banku centralnego do utrzymywania stóp procentowych na zadeklarowanym poziomie (odmiana „odysejska"), druga z nich ma zaś charakter bardziej prognozowania sytuacji i sugestii co do prawdopodobnych reakcji polityki pieniężnej (odmiana „delficka")[127].

Politykę *forward guidance* stosowała najdłużej i w najpełniejszej formie Rezerwa Federalna w Stanach Zjednoczonych (zob. ramka 5.26). Instrument ten wykorzystały także banki centralne Kanady, Wielkiej Brytanii, Japonii i strefy euro, a nawet, choć w bardzo ograniczonym zakresie, także Narodowy Bank Polski[128]. Oceniając skuteczność polityki *forward guidance*, trzeba podkreślić, że usprawnia ona pobudzanie gospodarki tylko wówczas, gdy rynki wierzą, że bank centralny będzie tolerował wyższą inflację przynajmniej przez jakiś czas.

Ramka 5.26. Polityka *forward guidance* w USA

Już w grudniu 2008 roku Federalny Komitet do spraw Operacji Otwartego Rynku (FOMC) zakomunikował, iż "Komitet przewiduje, że słabe warunki ekonomiczne

[127] Pierwsza z tych „antycznych" nazw nawiązuje do postaci Odyseusza, który kazał przywiązać się do masztu własnego okrętu, aby nie ulec pokusie syrenich śpiewów i uniknąć wpadnięcia na skały. Druga odwołuje się zaś do wyroczni delfickiej, której przepowiednie były bardziej ambiwalentne i zagadkowe, zapewniając tym samym większą swobodę interpretacji.

[128] Od września 2013 roku RPP zapowiadała w swoich komunikatach utrzymanie obniżonych stóp procentowych: najpierw „do końca" 2013 roku, od stycznia 2014 roku „co najmniej do końca pierwszego półrocza" 2014 roku, a od marca 2014 roku „w dłuższym okresie, tj. co najmniej do końca trzeciego kwartału" 2014 roku. Od lipca 2014 roku RPP zaprzestała stosowania *forward guidance*.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

uzasadniają utrzymanie wyjątkowo niskich poziomów stopy oprocentowania funduszy federalnych przez *pewien czas*" (*for some time*). W marcu 2009 roku, po zakończeniu pierwszej rundy nabywania skarbowych papierów wartościowych na dużą skalę, w komunikacie FOMC słowa „pewien czas" zostały zastąpione sformułowaniem „dłuższy okres czasu" (*for an extended period*). W sierpniu 2011 roku FOMC poszedł jeszcze o krok dalej, ogłaszając, że „Komitet przewiduje, iż stopa funduszy federalnych pozostanie wyjątkowo niska co najmniej do połowy 2013 roku", a więc uściślił, że horyzont jego własnego zobowiązania do utrzymania niskich stóp procentowych został wydłużony do dwóch lat. W komunikacie wystosowanym po posiedzeniu w styczniu 2012 roku FOMC zapowiedział wydłużenie okresu obowiązywania niskich stóp procentowych do końcowych miesięcy 2014 roku („*to late 2014*" – zob. Cambell i in. 2012), czyli na okres niemal trzech lat. Tego rodzaju publiczne komunikaty ogłaszane przez bank centralny nazywamy *forward guidance.* Każdy z nich wywoływał spadek długookresowych stóp procentowych (Williams 2014).

Na początku 2013 roku Fed ogłosił, że „stopy procentowe nie będą podnoszone, aż stopa bezrobocia spadnie do 6,5%, o ile zarazem inflacja nie będzie wyższa niż 2,5%". Tego typu powiązanie stopy procentowej z publicznie dostępnymi danymi statystycznymi niezależnymi od banku centralnego sprawia, że polityka *forward guidance*[129] staje się jeszcze bardziej wiarygodna. W październiku 2014 roku FOMC podtrzymał stanowisko dotyczące tego, że w świetle aktualnych danych wyjątkowo niski poziom stopy funduszy federalnych (mieszczący się w przedziale od 0% do 0,25%) jest uzasadniony i powinien zostać utrzymany przez „znaczny okres czasu (*a considerable time*) po zakończeniu bieżącego programu skupu aktywów, zwłaszcza jeśli inflacja utrzyma się poniżej długookresowego celu 2%, a długookresowe oczekiwania inflacji pozostaną dobrze zakotwiczone (*well anchored*)"[130].

[129] Mimo że już w kwietniu 2014 roku stopa bezrobocia w USA spadła poniżej 6,5% (do 6,3%), Fed nie zdecydował się na zaostrzenie polityki pieniężnej, a jedynie ogłosił, że stopniowo będzie ograniczał skalę skupu aktywów. Pierwsza podwyżka stóp procentowych (o 25 punktów bazowych) miała miejsce dopiero w grudniu 2015 roku.

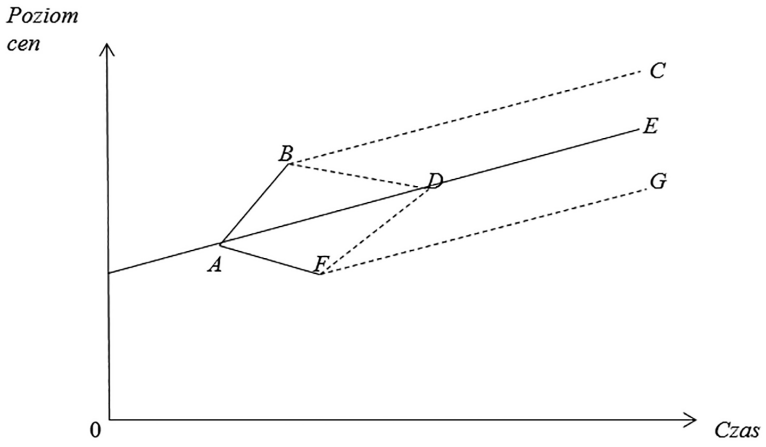
[130] Jednym z elementów polityki *forward guidance* w USA jest słynny wykres „kropkowy" (*dot chart*), publikowany przez Fed po każdym posiedzeniu FOMC, przedstawiający to, jaki poziom stopy oprocentowania funduszy federalnych członkowie FOMC uważają za właściwy w perspektywie średniookresowej. Na przykład na posiedzeniu FOMC z 16 marca 2016 roku mediana oczekiwań członków Komitetu wynosiła 0,875% na koniec 2016 roku, 1,937% na koniec 2017 roku, 3,0% na koniec 2018 roku i 3,25% w dłuższym okresie (*Economic projections of Federal Reserve Board members and Federal Reserve Bank presidents under their individual assessments of projected appropriate monetary policy,* Federal Reserve Board, Washington, 16 March, 2016).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.14.5. Zmiany w strategii BCI

Niezależnie od poszukiwań niekonwencjonalnych instrumentów polityki pieniężnej podejmowane są próby udoskonalenia i korygowania strategii bezpośredniego celu inflacyjnego. Prace te zmierzają w kilku kierunkach. Jednym z najbardziej popularnych pomysłów jest odejście od strategii BCI na rzecz strategii osiągania określonego poziomu cen, czyli *price-level targeting* – PLT (zob. Eggertsson i Woodford 2003; Kahn 2009; Carney 2009; Bean, Paustian, Penalver i Taylor 2010; Wojtyna 2012). O ile w strategii BCI chodzi o powrót do określonej *stopy* inflacji, o tyle w strategii PLT odchylenia od ścieżki cen są korygowane zmianami stóp procentowych w taki sposób, aby powrócić do ustalonego *poziomu* cen. Ilustruje to rysunek 5.11.

Rysunek 5.11. Strategia BCI a strategia PLT



Źródło: opracowanie własne.

Gdy inflacja rośnie i poziom cen podnosi się z punktu A do B, strategia BCI wymaga ustabilizowania inflacji na ścieżce BC, podczas gdy strategia PLT wymaga powrotu na ścieżkę *BDE*. Z kolei gdy inflacja spada i poziom cen obniża się z punktu *A* do *F*, strategia BCI wymaga ustabilizowania inflacji na ścieżce *FG*, a tymczasem strategia PLT wymaga powrotu na ścieżkę *BDE*. Oznacza to, że strategia PLT wymaga generalnie silniejszego zaostrzenia polityki pieniężnej niż strategia PLT przy wzroście inflacji (ponieważ poziom cen ma powrócić na wyjściową

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ścieżkę), lecz także wymaga silniejszego luzowania polityki pieniężnej wobec spadku inflacji niż strategia BCI. Różnica ta jest szczególnie ważna przy zagrożeniu deflacją – wobec spadku cen z punktu *A* do *F* strategia PLT umożliwia bardziej ekspansywną politykę pieniężną. Gdy stopy procentowe znajdują się już na dolnej zerowej granicy, strategia PLT ma przewagę nad strategią BCI, ponieważ wywiera presję na spadek realnych stóp i stabilizuje oczekiwania inflacyjne na wyższym poziomie, co przy danym poziomie nominalnych stóp powinno sprzyjać szybszemu wzrostowi gospodarczemu.

Problemem wymagającym rozwiązania w przypadku zastosowania strategii PLT jest sposób komunikowania się z rynkami. Wymaga ona bowiem porzucenia jasnego i zrozumiałego przekazu w postaci określonego, stałego celu inflacyjnego i przejścia na nowy wskaźnik, mniej powszechnie zrozumiały. Uzasadnianie przez bank centralny ciągłych zmian poziomu pożądanej inflacji może grozić utratą wiarygodności banku. Ogólnie rzecz biorąc, strategia PLT wiąże się z większą zmiennością stóp procentowych w czasie, co może destabilizować działalność gospodarczą. Być może z tego powodu żaden bank centralny nie zdecydował się na jej konsekwentne wdrożenie.

Kolejną interesującą propozycję stanowi porzucenie celu inflacyjnego na rzecz celu w postaci określonej stopy wzrostu nominalnego PKB (*nominal GDP growth targeting* – NGT). Załóżmy, że ustala się cel wzrostu nominalnego PKB na poziomie 5%, zakładając, że w normalnych warunkach potencjalne tempo realnego wzrostu wynosi 3%, a inflacja pozostaje na poziomie 2%. Gdy realne tempo wzrostu spada poniżej tempa potencjalnego, trzeba luzować politykę pieniężną. Na przykład, gdy realne tempo spada do 0%, należy tak obniżyć stopy procentowe, aby inflacja wzrosła do 5%. Z kolei gdy realne tempo wzrostu przyspiesza do 5%, przychodzi zaostrzać politykę pieniężną, aby obniżyć inflację do zera. Reguła stopy wzrostu nominalnego PKB zakłada więc pełną wymienność pomiędzy preferowaną stopą inflacji a stopą realnego wzrostu PKB (Sumner 2011, 2014; McCallum 2011a; Sheedy 2014).

Również i ta strategia wiąże się z szeregiem problemów. Z jednej strony brakuje wiarygodnych modeli i dobrego rozpoznania zależności pomiędzy poziomem PKB a inflacją. Z drugiej zaś znacznie trudniej jest stabilizować oczekiwania inflacyjne w ramach tej strategii, ponieważ wskaźnik nominalnego PKB pozostaje mniej czytelny niż wskaźnik CPI. Rodzi to ryzyko utraty wiarygodności antyinflacyjnej przez bank centralny. Wreszcie też istotną trudność techniczną stanowi brak miesięcznych danych dotyczących PKB większości państw.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Trzecim możliwym sposobem korekty strategii BCI jest po prostu podwyższenie celu inflacyjnego w okresach przewlekłej recesji. Na przykład Blanchard, Dell'Ariccia i Mauro (2010) proponują, aby w czasie kryzysu podnieść cel inflacyjny z 2% do 4%, żeby ograniczyć ryzyko spadku stóp procentowych do dolnej „zerowej granicy" i ogłosić, że ten wyższy cel inflacyjny będzie utrzymany aż do momentu, w którym stopa bezrobocia spadnie do ustalonego wcześniej poziomu, na przykład 6,5%. Gdy nie udało się dalej obniżać stóp nominalnych poniżej zera, należy po prostu wykreować oczekiwania wyższej inflacji. Aby uwiarygodnić tę politykę, można podnieść cel inflacyjny albo obniżyć wymagany próg bezrobocia. Uzasadniając tę propozycję, autorzy wskazują na duże prawdopodobieństwo powtarzania się w przyszłości kryzysów finansowych i występowania problemu „zerowej granicy"[131].

Krytycy tej propozycji wskazują, że pociąga ona za sobą wyższe koszty netto inflacji – w zasadzie przy wyższym poziomie inflacji ma z reguły miejsce zwiększona dyspersja cen i następuje spadek popytu na realne zasoby pieniądza, co ujemnie wpływa na wzrost gospodarczy (zob. punkt 5.3.6). Podkreśla się także trudności komunikacyjne i ryzyko spadku wiarygodności banku centralnego, związane z akceptowaniem wyższego poziomu inflacji i ogólnie z większym zakresem dyskrecjonalności polityki pieniężnej (Bean, Paustian, Penalver i Taylor 2010; Issing 2011; McCallum 2011b).

Podsumowując, wydaje się, że strategia BCI wyszła w zasadzie zwycięsko z trudnej próby, jaką okazał się kryzys finansowy i gospodarczy lat 2008–2012, i wciąż jest traktowana jako najbezpieczniejsza i najmniej kontrowersyjna odmiana reguły polityki pieniężnej. Zarazem można jednak przypuszczać, że lekcje odebrane w czasie kryzysu posłużą bankom centralnym jako materiał do dokonywania kolejnych modyfikacji i usprawnień strategii BCI, przy ogólnym zachowaniu jej ram.

5.14.6. Funkcja „pożyczkodawcy ostatniej szansy" dla rządów i wielorakie równowagi na rynku długu

Kryzys zadłużenia w strefie euro ujawnił istotne niedoskonałości systemu dyscypliny budżetowej i finansowej obowiązującego w Unii Europejskiej. Uświadomił

[131] Również Schmitt-Grohé i Uribe (2012) proponują przejściowe podwyższenie celu inflacyjnego w horyzoncie średniookresowym, tak aby uzyskać niezbędny spadek płac realnych w zadłużonych krajach strefy euro w warunkach sztywności płac nominalnych. Autorzy szacują, że podniesienie inflacji do 4,2–4,3% w okresie 5 lat pozwoliłoby powrócić do poziomu pełnego zatrudnienia w zadłużonych (peryferyjnych) państwach strefy euro.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

też ograniczenia, jakim podlega Europejski Bank Centralny wobec zaangażowania się we wspólną politykę pieniężną. Ożywiona dyskusja rozwinęła się zwłaszcza wokół pytania o to, czy w sytuacji wystąpienia napięć na rynku długu suwerennego państw członkowskich strefy euro EBC powinien pełnić funkcję „pożyczkodawcy ostatniej szansy" dla rządów tych państw (zob. np. De Grauwe 2011).

Rządy pożyczają pieniądze, sprzedając papiery wartościowe na rynku, aby pokryć nadwyżkę wydatków publicznych nad dochodami. Inwestorzy chętnie kupują papiery skarbowe, traktując je jako bardzo bezpieczną lokatę, ponieważ za tym szczególnym dłużnikiem, jakim jest rząd, stoi władza polityczna dysponująca prawem i zdolnością do nakładania i ściągania podatków. Dlatego też obligacje rządowe z reguły traktuje się jako papiery pozbawione ryzyka (*risk-free*)[132]. Zdarzają się jednak sytuacje, w których rząd traci zaufanie inwestorów i napotyka na trudności w pozyskiwaniu środków. Dzieje się tak wówczas, gdy zadłużenie kraju przekracza poziom uznawany przez inwestorów za bezpieczny lub kiedy z innych powodów rynki finansowe dojdą do wniosku, że rząd może mieć kłopoty z obsługą długu. W większości krajów, gdy w wyniku utraty zaufania pojawia się ryzyko utraty płynności przez rząd danego państwa bądź nawet ryzyko niewypłacalności, bank centralny z reguły przychodzi z pomocą swojemu rządowi, skupując na rynku wtórnym obligacje skarbowe w zamian za pieniądz rezerwowy. Dzieje się tak dlatego, że bank centralny stanowi część władzy publicznej. Obligacje rządu państwa, które posiada własny bank centralny, są więc obietnicą spłaty jednego zobowiązania rządowego, jakim jest dług, za pomocą innego zobowiązania rządowego, jakie stanowi gotówka. Dlatego posiadacz obligacji wie, że zawsze może ją sprzedać na rynku, ponieważ w sytuacji braku popytu ze strony inwestorów prywatnych zawsze odkupi ją albo rząd, albo bank centralny. Należy zauważyć, że bank centralny jest zawsze w stanie emitować w dowolnej ilości krajowy pieniądz rezerwowy. Tak długo więc, jak gotów jest on emitować pieniądz na potrzeby skupu obligacji, rząd pozostaje – przynajmniej formalnie – wypłacalny[133]. Nie występuje zatem tzw. ryzyko kontrahenta (*counterparty risk*). Gdy kraj posiada więc własny bank centralny, który

[132] Banki chętnie kupują papiery skarbowe, ponieważ przypisuje się im zerową wagę ryzyka przy wyliczaniu wartości aktywów ważonych ryzykiem wobec potrzeby wyznaczania współczynnika adekwatności kapitałowej.

[133] Oczywiście w takiej sytuacji może także wystąpić kryzys płynności lub kryzys wypłacalności, jeśli zobowiązania zostały zaciągnięte w walutach obcych. Krajowy bank centralny może bowiem dostarczać dowolną ilość pieniądza krajowego, lecz posiada jednocześnie ograniczone rezerwy walut zagranicznych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przygotowany jest na to, aby w sytuacji kryzysu płynności wykupić dług danego rządu, rynki nigdy nie doprowadzą do niewypłacalności państwa[134].

Ten model zależności zachodzącej między rządem, długiem publicznym i bankiem centralnym zmienił się po utworzeniu unii walutowej w Europie. Banki centralne państw członkowskich utraciły prawo emitowania własnej waluty, a prawo emitowania wspólnej waluty (euro) pozostało wyłącznie w kompetencji Europejskiego Banku Centralnego. Na długu suwerennym wystąpiło ryzyko kontrahenta. Pojedynczy kraj członkowski strefy euro, pozbawiony wsparcia własnego banku centralnego, jest zmuszony rekompensować ten brak wyższymi premiami za ryzyko w stopach procentowych. Zwiększenie tych stawek podnosi jednak koszt obsługi długu, co sprawia, że mogą stać się one przyczyną kryzysu zadłużenia. Dojdzie do niego wówczas, gdy inwestorzy uznają, że koszt obsługi zadłużenia przekracza możliwości płatnicze rządu. Wówczas rozpoczyna się wyprzedaż długu. To z kolei przekłada się na spadek cen obligacji i wzrost ich rentowności, co oznacza dalszy wzrost kosztów obsługi długu i trudności w refinansowaniu zadłużenia. Uruchamiany jest mechanizm samospełniających się obaw i następuje kryzys płynności. Jeśli kraj nie uzyska pomocy z zewnątrz, kryzys płynności szybko przekształci się w kryzys niewypłacalności, a kraj stanie się teoretycznie bankrutem. Charakterystyczne w przypadku kryzysu zadłużenia jest to, że dany kraj może być fundamentalnie wypłacalny, a jednocześnie zbankrutować na skutek utraty zaufania inwestorów i braku płynnych środków potrzebnych na pokrycie zapadających na bieżąco zobowiązań.

Dzieje się tak dlatego, że na rynkach finansowych, w odróżnieniu od rynków dóbr i usług, mamy do czynienia z istnieniem tzw. wielorakich równowag (*multiple equilibria*)[135]. Zjawisko to jest często ilustrowane w sposób przedstawiony na rysunku 5.12, który ukazuje wielorakie równowagi na rynku finansowym w dwóch wariantach – z uwzględnieniem dyskrecjonalnych zmian popytu (lewa strona) oraz ciągłych zmian popytu (prawa strona). Punktem wyjścia jest założenie, że rynek aktywów finansowych różni się zasadniczo od rynku produktów. Różnica

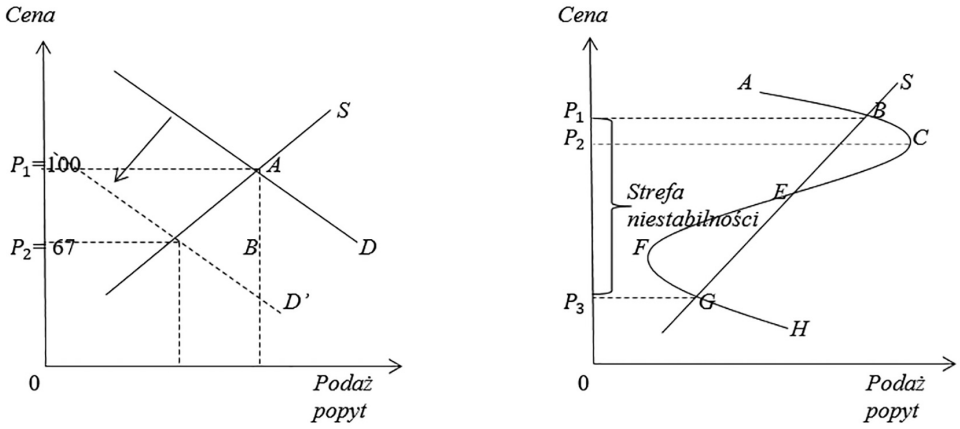
[134] Oczywiście, przy wysokim poziomie zadłużenia i dużej skali skupu obligacji przez bank centralny może wystąpić rosnąca inflacja. Powiększa ona nominalny PKB, co przy stałej nominalnej wartości długu wywołuje spadek relacji długu do PKB. Inflacja „zjada" więc wartość realną długu, co daje możliwość „inflacyjnego wyjścia" z nadmiernego zadłużenia krajom dysponującym własną walutą. Jest to ważny fiskalny amortyzator szoków – czasami tańszy z punktu widzenia kosztów makroekonomicznych niż wysoka premia za ryzyko lub ogłoszenie niewypłacalności (*outright default*).

[135] Zob. Eichengreen, Rose i Wyplosz (1995); Obstfeld i Rogoff (1997); Morris i Shin (2000); De Grauwe i Li (2012); Gros (2012).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

polega na tym, że o ile celem zakupu dóbr i usług jest ich konsumpcja (pomijamy zakupy pewnych dóbr w celach spekulacyjnych, ponieważ wówczas przestają one być dobrami konsumpcyjnymi, a stają się aktywami), o tyle celem zakupu aktywów finansowych jest z reguły uzyskiwanie dochodu, w tym dochodu z ich odsprzedaży po wyższej cenie. Ze względu na to, że uzyskiwanie dochodu z aktywów wymaga czasu, a przyszłe strumienie dochodów i przyszłe zmiany cen są zawsze obarczone niepewnością, istotnym parametrem branym pod uwagę przez inwestorów jest ryzyko związane z danym aktywem. Czynnik ten na ogół nie występuje przy zakupach dóbr i usług, które są przeznaczone do bieżącej konsumpcji.

Rysunek 5.12. Zjawisko wielorakich równowag na rynku długu suwerennego



Źródło: http://www.economist.com/blogs/freeexchange/2011/10/illustrated-euro-crisis (13.06.2012) oraz opracowanie własne.

Po lewej stronie rysunku 5.12 linia *S* wyznacza funkcję podaży obligacji skarbowych danego kraju, a linia *D* jest funkcją popytu. W normalnych warunkach i przy danym poziomie ryzyka popyt na obligacje rośnie wówczas, gdy ich cena spada. Punkt przecięcia funkcji podaży i popytu wyznacza poziom ceny obligacji, zapewniający równowagę na rynku. Cena ta zależy od wymaganego przez inwestorów poziomu rentowności obligacji. Na przykład, jeśli wymagana rentowność wynosi $r = 4\%$, to cena znajdzie się na poziomie $P_1 = 100$. Lecz cena obligacji stanowi funkcję nie tylko oczekiwanej stopy przychodu, ale także poziomu ryzyka. Gdy na rynek trafia informacja negatywna lub z innego powodu awersja do ryzyka wzrasta, rośnie również premia za ryzyko, co wywołuje wzrost wymaganej rentowności,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

na przykład do poziomu $r = 6\%$. Funkcja popytu przesuwa się z położenia *D* do położenia *D*', a cena spada do poziomu $P_2 = 67$.

Prawa strona rysunku 5.12 ukazuje ten sam mechanizm w warunkach ciągłych zmian popytu. W normalnych okolicznościach krzywa popytu ma kształt *ABC* (odpowiednik krzywej *D* w części lewej), a punkt równowagi rynkowej wyznacza *B*. W tym miejscu równowaga jest stabilna, ponieważ wzrost ceny przekraczający poziom P_1 wywołuje spadek popytu, co wywiera presję na obniżenie ceny do poprzedniego poziomu. Podobnie też spadek ceny poniżej poziomu P_1 przekłada się na wzrost popytu, co stwarza presję na podniesienie ceny do poprzedniego poziomu. W obu sytuacjach mamy więc do czynienia z automatycznym mechanizmem powrotu do punktu równowagi.

Lecz gdy spadek ceny jest silniejszy, mechanizm samoczynnego przywracania równowagi na rynku finansowym przestaje działać. Tak dzieje się wówczas, gdy na rynek trafia nieoczekiwana informacja negatywna lub silnie rośnie awersja do ryzyka. Wówczas spadek ceny obligacji poniżej poziomu P_2 powoduje zmianę nachylenia krzywej popytu z ujemnego na dodatnie (*CEF*). Oznacza to, że spadek cen obligacji wywołuje spadek popytu i dalszą wyprzedaż obligacji aż do punktu *F*, po czym krzywa popytu powraca do normalnego nachylenia *FGH* (odpowiednik krzywej *D*' w części lewej), a nową ceną równowagi jest P_3, odpowiadający punktowi równowagi *G*. Zauważmy, że w punkcie *E* równowaga ma charakter niestabilny, ponieważ spadek ceny wywołuje dalszy spadek popytu zamiast jego wzrostu, jak miało to miejsce w punkcie równowagi stabilnej *B*. W obszarze zawierającym się między punktami *C* i *F* mamy do czynienia z ostrą fazą kryzysu, kiedy rynki finansowe dostosowują się spontanicznie do nowej sytuacji. W tym przedziale ceny obligacji ulegają załamaniu ze względu na panikę rynkową i skłonność inwestorów do stadnych zachowań[136]. Taki właśnie był przebieg wydarzeń na rynku długu suwerennego we Włoszech i w Hiszpanii w lecie 2011 roku i w lecie 2012 roku (zob. ramka 5.27).

[136] Niestabilna równowaga może wystąpić nie tylko w okresie kryzysu, lecz także w czasie boomu spekulacyjnego: wówczas wzrost cen obligacji wywołuje wzrost popytu, a równowaga rynkowa przesuwa się z punktu *G* do punktu *B*.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 5.27. Napięcia na rynku długu suwerennego Włoch i Hiszpanii w latach 2011–2012 i rola EBC

Z sytuacją kryzysu płynności i mechanizmem samospełniających się obaw mieliśmy do czynienia w lecie 2011 roku na rynku długu suwerennego Włoch i Hiszpanii. Oba kraje, a zwłaszcza Włochy, były już wcześniej poważnie zadłużone, co nie zaniepokoiło jednak rynków finansowych, ponieważ oba państwa były uważane za fundamentalnie wypłacalne. Oprocentowanie długu suwerennego pozostawało na umiarkowanym poziomie (około 4%), dzięki czemu zarówno Włochy, jak i Hiszpania bez kłopotów finansowały się na rynku długu. Strefa euro znalazła się jednak już wtedy w niestabilnej sytuacji ze względu na kryzys zadłużenia suwerennego, jaki wystąpiły w Grecji, Irlandii i Portugalii. Gdy w czerwcu 2011 roku okazało się, że Grecja nie radzi sobie z programem oszczędnościowym i zmuszona będzie prosić o kolejną pomoc, poziom zaufania na rynkach silnie zmalał, a inwestorzy zaczęli pozbywać się obligacji włoskich i hiszpańskich. Był to klasyczny przykład efektu „zarażenia" (*contagion effect*) – inwestorzy doszli do wniosku, że jeśli Grecja mogła zbankrutować, to mogą zbankrutować także inne wysoko zadłużone kraje. W ciągu kilku miesięcy rentowność 10-letnich obligacji hiszpańskich wzrosła o 80 punktów bazowych (do 5,8%), włoskich zaś aż o 220 (do 5,5%). Wobec ryzyka rozprzestrzenienia się kryzysu zadłużenia na inne kraje Europejski Bank Centralny postanowił interweniować i w sierpniu 2011 roku rozpoczął skup obligacji włoskich i hiszpańskich na rynku wtórnym. Dzięki temu rentowność obligacji obu państw w ciągu kilku następnych miesięcy wróciła do poprzedniego poziomu.

Brak wyraźnej deklaracji ze strony EBC co do skali i trwałości dalszej interwencji spowodował jednak ponowny spadek zaufania na rynku długu. W połowie 2012 roku rentowność papierów włoskich i hiszpańskich znowu wzrosła do poziomu odpowiednio 6,0% i 6,8%. Dalszy ich wzrost groził utratą zdolności obsługi długu przez oba kraje. Wówczas EBC zdecydował się zainterweniować w sposób bardziej radykalny. 26 lipca 2012 roku prezes EBC Mario Draghi ogłosił, że EBC zrobi „wszystko co trzeba" (*whatever it takes*), aby chronić euro, dając do zrozumienia, że nie dopuści do załamania na rynku długu suwerennego. Spowodowało to momentalne uspokojenie sytuacji na rynku długu[137]. Kilka tygodni

[137] Te słynne słowa padły podczas Global Investor Conference w Londynie 26 lipca 2012 roku ("Within our mandate, the ECB is ready to do whatever it takes to preserve the euro. And believe me, it will be enough").

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

później EBC sprecyzował, że jest gotów skupować na rynku wtórnym nieograniczone ilości obligacji krajów strefy euro, w tym Włoch i Hiszpanii, w ramach specjalnego programu OMT (*outright monetary transactions*). W ciągu tygodnia rentowność włoskich obligacji 10-letnich spadła o 85 punktów bazowych, a hiszpańskich o 140. W połowie 2014 roku kształtowały się już na poziomie około 2,5%, znacznie poniżej poziomu notowanego przed kryzysem. Ryzyko kryzysu zadłużenia zostało zażegnane.

Trzeba podkreślić, że efekt stabilizujący osiągnięto nie dzięki interwencyjnemu skupowi obligacji, a za sprawą wiarygodnej deklaracji banku centralnego. Rynki finansowe uwierzyły, że EBC jest w stanie kupować nieograniczone ilości obligacji krajów zadłużonych, ponieważ jedynie on, jako wyłączny „producent" euro, może dokonać zakupów na tak wielką skalę. W rzeczywistości EBC nie wydał ani jednego euro na operacje OMT. Wystarczyła jedynie wiarygodna deklaracja o gotowości do skupu, aby inwestorzy przestali wyprzedawać papiery włoskie i hiszpańskie. W takiej sytuacji inwestorzy zyskali bowiem pewność, że nawet jeśli rządy Włoch i Hiszpanii miałyby trudności z wykupem papierów wartościowych, zawsze będą mogli sprzedać te papiery EBC.

Rysunek 5.12 pokazuje, dlaczego potrzebny jest „pożyczkodawca ostatniej szansy dla rządów". W czasie kryzysu ceny obligacji spadają drastycznie z powodu paniki rynkowej, co wywołuje skokowy wzrost rentowności. Bez wsparcia z zewnątrz może to grozić bankructwem. Wiarygodna interwencja banku centralnego lub chociażby sama jej zapowiedź zapewnia uspokojenie nastrojów i powrót zaufania. W rezultacie krzywa popytu na obligacje powraca do początkowego położenia i kryzys zostaje zażegnany.

EBC nie posiada funkcji „pożyczkodawcy ostatniej szansy dla rządów" zapisanej *expressis verbis* w swoim mandacie, lecz w praktyce może bez zasadniczych przeszkód wystąpić w tej roli. Statut EBC zezwala mu na dokonywanie transakcji zakupu i sprzedaży papierów wartościowych na rynku wtórnym, w tym skarbowych papierów wartościowych, w trybie warunkowym lub definitywnym (art. 18 Statutu EBC). Trzeba jednak pamiętać, że rolą EBC jest zapobieganie epizodom paniki na rynku długu i dbanie o stabilność wspólnej waluty, a nie przejmowanie długów państw członkowskich, które są fundamentalnie niewypłacalne. Tak więc operacje przeprowadzane w ramach funkcji „pożyczkodawcy ostatniej szansy dla rządów" muszą być podporządkowane celom polityki pieniężnej, w tym stabilizowaniu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

rynkowych stóp procentowych, a nie celom fiskalnym, takim jak na przykład ułatwianie finansowania deficytu budżetowego.

5.15. Przyszłość pieniądza

5.15.1. Jaka będzie przyszłość pieniądza gotówkowego?

Pieniądz w postaci znaków pieniężnych towarzyszy ludzkości od ponad trzech tysięcy lat, zapewniając płynność transakcji, możliwość przechowywania majątku i wygodę. Rozwój elektronicznych form pieniądza i systemów płatności w okresie ostatnich kilkudziesięciu lat znacząco ograniczył jednak zakres wykorzystywania pieniądza gotówkowego. Pierwsze karty kredytowe weszły do użycia na początku lat 50. ubiegłego wieku, lecz ich prawdziwa ekspansja, rozciągająca się na masową skalę, nastąpiła na przełomie roku 1989 i 1990 po wprowadzeniu elektronicznego systemu rozliczania transakcji konsumpcyjnych. Obecnie stosowane karty płatnicze (debetowe i kredytowe) wyparły już znaczną część gotówki z rynku. W USA udział transakcji gotówkowych w całkowitej liczbie transakcji związanych z zakupem dóbr i usług systematycznie maleje; co ciekawe, w ostatnich latach spada również liczba bankomatów w relacji do liczby ludności (Williams 2012), a to świadczy o malejącym popycie na gotówkę.

Obserwując wskazane tendencje, amerykański ekonomista Kenneth Rogoff doszedł do wniosku, że całkowita rezygnacja z pieniądza gotówkowego byłaby technicznie możliwa, a dodatkowo przyniosłaby szereg istotnych korzyści (Rogoff 2014). Przede wszystkim znikłyby koszty związane z drukiem banknotów i emisją monet. Przykładowo koszt emisji pieniądza gotówkowego w Polsce w 2013 roku wyniósł prawie 269 mln PLN (NBP 2014). Po drugie rezygnacja z gotówki pozwoliłaby wyeliminować zjawisko fałszowania banknotów. Straty z tego tytułu są szczególnie dotkliwe w krajach, których waluty pełnią funkcję międzynarodowych walut rezerwowych (USA, strefa euro, Wielka Brytania). Na przykład w samym tylko 2013 roku Bank Anglii zidentyfikował i wycofał z obiegu ponad 680 000 fałszywych banknotów na łączną kwotę 11,5 mln GBP. Po trzecie likwidacja gotówki ograniczyłaby wydatnie zakres szarej strefy w gospodarce, w tym zwłaszcza skalę nielegalnych transakcji o podłożu kryminalnym. Według szacunków OECD, transakcje takie, od nierejestrowanych usług pomocy domowych do nielegalnego handlu bronią i narkotykami, stanowią w rozwiniętych krajach równowartość około 20%

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

PKB. Część z nich musiałaby pojawić się w oficjalnym obiegu, co z kolei pozwoliłoby uzyskać dodatkowe dochody z podatków.

Co więcej, rezygnacja z obrotu gotówką stworzyłaby dodatkową przestrzeń dla jeszcze silniejszego luzowania polityki pieniężnej, niż ma to miejsce obecnie. Jest to szczególnie ważne w przypadku gospodarek borykających się z problemem „zerowej granicy" stóp procentowych (*zero lower bound*). W obecnej sytuacji próba obniżenia nominalnych stóp procentowych głębiej poniżej zera mogłaby skłonić indywidualnych posiadaczy depozytów do zamienienia ich na gotówkę. Gdyby takiej możliwości jednak nie było, bank centralny mógłby śmielej obcinać stopy procentowe, sprowadzając je do ujemnych wartości, by jeszcze bardziej obniżać w ten sposób stopy realne.

Lecz rezygnacja z pieniądza gotówkowego pociąga za sobą również pewne niedogodności. Przede wszystkim znika dochód z emisji, który w krajach o największym obiegu pieniężnym, takich jak USA czy strefa euro, może osiągać równowartość nawet 0,3–0,5% PKB (w zależności od aktualnego poziomu stóp procentowych). Ponadto, pieniądz gotówkowy jest wciąż ważny dla wielu ludzi, którzy z różnych powodów nie chcą lub nie potrafią korzystać z kart płatniczych. Co więcej, niskie lub zerowe stopy procentowe obniżają koszt alternatywny gromadzenia gotówki i czynią z niej, a zwłaszcza z banknotów o dużych nominałach, atrakcyjny środek przechowywania wartości. Jak podaje tygodnik *The Economist* w artykule z 3 października 2015 roku, w latach 2008–2014 wartość banknotów znajdujących się w obiegu w strefie euro wzrosła o ponad 30%, w Wielkiej Brytanii o ponad 35%, w USA o ponad 50%, a w Korei Południowej aż o ponad 130%[138]!

Wreszcie pojawia się kwestia bezpieczeństwa. Całkowita dematerializacja pieniądza może nieść ryzyko poważnych zaburzeń w sytuacji, gdy systemy elektronicznych rozliczeń ulegają awarii, czy to z powodu naturalnych usterek technicznych, czy ze względu na świadome akty sabotażu. Wobec wszystkich tych argumentów nie należy więc raczej spodziewać się rychłej rezygnacji z pieniądza gotówkowego. A zatem „Gotówka umarła! Niech żyje gotówka!" (*the cash is dead, long live the cash!*) – jak niedawno pisał John C. Williams, prezes regionalnego Banku Rezerwy Federalnej w San Francisco.

[138] Choć w Szwecji wartość pieniądza gotówkowego w obiegu spadła w tym samym czasie o 25%.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

5.15.2. Pieniądz wirtualny

Poza upowszechnianiem się kart płatniczych początek XXI wieku przyniósł także szybki rozwój alternatywnych form pieniądza, których konstrukcja opiera się na wykorzystaniu technologii informatycznych. Jedną z nich jest pieniądz wirtualny, zwany także kryptopieniądzem. Przyjmuje on postać zapisu cyfrowego umożliwiającego przechowywanie wartości i dokonywanie płatności między użytkownikami, którzy go akceptują. Najbardziej popularnym pieniądzem wirtualnym jest Bitcoin, wprowadzony do obiegu gospodarczego w 2009 roku przez osobę (lub grupę osób) o pseudonimie Satoshi Nakamoto. Lecz w obiegu funkcjonują także inne popularne odmiany kryptowalut, takie jak na przykład mile zbierane w ramach programów lojalnościowych firm lotniczych czy punkty uzyskiwane za zakupy w sieciach handlowych.

Bitcoin jest zarówno pieniądzem wirtualnym, jak i zdecentralizowanym systemem płatniczym, który pozwala swoim uczestnikom na wymienianie się „wartościami" w sieci bez pośrednictwa jakiegokolwiek banku lub instytucji rządowej. Przy pomocy Bitcoina można dokonywać różnych transakcji między użytkownikami łącza, a także nabywać realne dobra i usługi. Emisja Bitcoina jest ograniczona przez specyficzny mechanizm probabilistyczny, generujący wirtualne monety w taki sposób, aby ich obieg nigdy nie przekroczył 21 milionów. Sposób emisji i wymiany Bitcoinów zapewnia bezpieczeństwo transakcji poprzez stały dostęp do pełnej historii wykorzystania danej monety (w postaci tzw. bloków łańcuchowych – *blockchains*), a zarazem oferuje użytkownikom wysoki stopień anonimowości. W przeciwieństwie do tradycyjnych form pieniądza posługiwanie się Bitcoinem nie opiera się na zaufaniu do pieniądza, a raczej na zaufaniu do generującego go systemu informatycznego i technologii, które zapewniają przejrzystość i funkcjonowanie zgodne z określonymi regułami. Ze względu na oderwanie od instytucji publicznych Bitcoin bywa nazywany przez jego zwolenników najbardziej „demokratycznym" – lub prywatnym – pieniądzem na świecie.

Spośród zalet związanych z posługiwaniem się wirtualna walutą najczęściej wymienia się niskie koszty transakcyjne, przejrzystość i szybie tempo dokonywania transakcji oraz stały dostęp do pełnej historii wykorzystania każdej monety. Dzięki temu wiele firm posługuje się wirtualnymi pieniędzmi, uzupełniając w ten sposób podaż pieniądza tradycyjnego. Z kolei do wad takiego rozwiązania zalicza się dużą zmienność wartości jednostki Bitcoina, potencjalną podatność na usterki

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

techniczne i ataki hackerskie, a także brak jakiejkolwiek zewnętrznej regulacji prawnej odnoszącej się do obrotu środkami i ochrony użytkowników. Wskazuje się też na to, że anonimowość transakcji może zachęcać do wykorzystywania walut wirtualnych w działalności nielegalnej i kryminalnej[139].

Skala posługiwania się walutami wirtualnymi jest nadal znikoma w porównaniu z tradycyjnym obiegiem pieniężnym. Kryptopieniądze nie wywołują więc większych zakłóceń w polityce pieniężnej banku centralnego. Czy nie można jednak spodziewać się, że w przyszłości ta sytuacja może ulec zmianie, biorąc pod uwagę szybki rozwój cyfrowych form pieniądza zaobserwowany w ostatnich latach? Owszem, nie jest to wykluczone, lecz w rzeczywistości mało prawdopodobne. Nie należy bowiem sądzić, że różne odmiany pieniądza wirtualnego lub szerzej – prywatnego – mogą w dającej się przewidzieć przyszłości zdominować obieg pieniężny i skutecznie wyprzeć z rynku pieniądz publiczny, czyli oficjalny. Tak długo, jak rządy państw będą wymagały uiszczania podatków w pieniądzu oficjalnym, oraz tak długo, jak będą wypłacały wynagrodzenia oraz opłacały rachunki w pieniądzu publicznym, tak długo pieniądz wirtualny i inne odmiany kryptopieniądza będą jedynie dodatkiem pozostającym na marginesie oficjalnego obiegu pieniężnego.

Literatura cytowana

Akerlof G.A., Yellen J., *Efficiency Wage Models of the Labor Market*, Cambridge University Press, New York 1986.

Barnett W., *Understanding the New Divisia Monetary Aggregates*, w: *The Theory of Monetary Aggregation*, red. W. Barnett, A. Serletis, Elsevier, Amsterdam 2000.

Barro R.J., *Inflation and Economic Growth,* Bank of England Quarterly Bulletin, May 1995.

Barro R.J., Gordon D., *Rules, Discretion and Reputation in a Model of Monetary Policy*, "Journal of Monetary Economics" 1983, no. 12.

Batini N., Turnbull K., *Monetary Conditions Indices for the UK: A Survey*, Bank of England, "External MPC Unit Discussion Paper" 2000, no. 1.

Bean C., Paustian M., Penalver A., Taylor T., *Monetary Policy After the Fall*, w: *Macroeconomic Challenges: The Decade Ahead*, A Symposium Sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, July 2011.

Benassy-Quéré A., Qoeuré B., Jacquet P., Pisani-Ferry J., *Economic Policy. Theory and Practice*, Oxford University Press, Oxford–New York 2010.

[139] Więcej na temat walut wirtualnych zob. European Central Bank (2012, 2015).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Bernanke B.S., Gertler M., *Inside the Black Box: The Credit Channel of Monetary Policy Transmission*, "Journal of Economic Perspectives" 1995, vol. 9, no. 2.

Bernanke B.S., Gertler M., Gilchrist S., *The Financial Accelerator in a Quantitative Business Cycle Framework*, "NBER Working Paper" 1998, no. 6455.

Bernanke B.S., Gertler M., Watson M., *Systematic Monetary Policy and the Effects of Oil Price Shocks*, "Brookings Papers on Economic Activity" 1997, no. 1.

Bernanke B.S., Gertler M., *Should Central Banks Respond to Movements in Asset Prices*, "American Economic Review" 1999, vol. 91, no. 2.

Bernanke B.S., Laubach T., Mishkin F.S., Posen A.S., *Inflation targeting. Lessons from International Experience*, Princeton University Press, Princeton–New Jersey 1999.

Bernanke B.S., Mishkin F.S., *Inflation Targeting: A New Framework for Monetary Policy?*, "NBER Working Paper" 1997, no. 5893.

Bernanke B.S., *The Crisis and the Policy Response*, speech, 13 January, 2009, http://www.federalreserve.gov/newsevents/speech/bernanke20090113a.htm (13.01.2010).

Billio M., Getmansky M., Lo A.W., Pelizzon L., *Measuring Systemic Risk in the Finance and Insurance Sectors*, "MIT Sloan School Working Paper" 2010, no. 4774–103/17/2010.

Blanchard O.J., Kiyotaki N., *Monopolistic competition and the effects of aggregate demand*, "American Economic Review" 1987, no. 77(4).

Blanchard O.J., Galí J., *Real Wage Rigidities and the New Keynesian Model*, "NBER Working Paper" 2005, no. 11806.

Blanchard O.J., Galí J., *Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment*, "American Economic Journal: Macroeconomics" 2010, no. 2.

Blanchard O.J., *Macroeconomics,* Upper Saddle River, Prentice-Hall 1997.

Blanchard O.J., Diamond N., *The Beveridge Curve*, "Brookings Papers on Economic Activity" 1989, no. 1.

Blanchard O.J., Dell'Ariccia G., Mauro P., *Rethinking Macroeconomic Policy*, IMF Staff Position Note, SPN/10/03, 2010.

Blanchard O.J., Dell'Ariccia G., Mauro P., *Rethinking Macroeconomic Policy. Getting Granular*, IMF Staff Discussion Note, SDN/13/03, 2013.

Blinder A.S., *What Central Bankers Can Learn from Academics – and Vice Versa*, "Journal of Economic Perspectives" 1997, vol. 11, no. 2.

Borio C., *Central banking Post-Crisis. What Compass for Uncharted Waters?*, "BIS Working Papers" 2011, no. 353.

Boskin M.J. i in., *Toward a More Accurate Measure of the Cost of Living*, Final Report to the Senate Finance Committee, 4 December, 1996.

Brunnermeier M. i in., *The Fundamental Principles of Financial Regulation,* 11th Geneva Report on the World Economy, Geneva 2009.

Brzoza-Brzezina M., *Polska polityka pieniężna. Badania teoretyczne i empiryczne*, C.H. Beck, Warszawa 2011.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Buiter W., *The elusive welfare economics of price stability as a monetary policy objective: Why new Keynesian central bankers should validate core inflation*, "ECB Working Paper Series" 2006, no. 609.

Buiter W., *Quantitative easing and qualitative easing: a terminological and taxonomic proposal*, 9 December, 2008, http://blogs.ft.com/maverecon/2008/12/quantitative-easing--and-qualitative-easing-a-terminological-and-taxonomic-proposal/#axzz3Gn2czDrq (27.04.2010).

Buiter W.H., *The Simple Analytics of Helicopter Money: Why it Works Always*, "Economics" 2014, no. 8.

Buiter W.H., Panigirtzoglou N., *Ovecoming the Zero Bound on Nominal Interest Rates with negative Interest on Currency. Gesell's Solution*, "The Economic Journal" 2003, vol. 113(490).

Burda M., Wyplosz C., *Makroekonomia. Podręcznik europejski*, wydanie 2 zmienione, PWE, Warszawa 2000.

Cagan Ph., *The Monetary Dynamics of Hyperinflation*, w: *Studies in the Quantity Theory of Money*, red. M. Friedman, University of Chicago Press, Chicago 1956.

Calvo G.A., *Staggered prices in a utility-maximizing framework*, "Journal of Monetary Economics" 1983, no. 12(3).

Campbell J.R., Evans C., Fisher J.D.M., Justiniano A., *Macroeconomic Effects of Federal Reserve Forward Guidance*, "FRB of Chicago Working Paper" 2012, no. 3.

Cargill T.F., *Monetary policy, deflation and economic history: lessons for the Bank of Japan*, Bank of Japan, "Monetary and Economic Studies" 2001, February, Special Edition.

Carney M.J., *Commentary: Using Monetary Policy to Stabilize Economic Activity*, w: *Financial Stability and Economic Policy*, A Symposium Sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, August 2009.

Chen H., Curdia V., Ferrero A., *The Macroeconomic Effects of Large-Scale Asset Purchase Programs*, "FRB of San Francisco Working Paper" 2012, no. 22.

Cieśla N., *Konstrukcja pieniężnych agregatów Divisia w warunkach polskich*, Narodowy Bank Polski, „Materiały i Studia" 1999, nr 89.

Claeys G., Leandro A., Mandra A., *European Central Bank Quantitative Easing: A Detailed Manual*, „Bruegel Policy Contribution", issue 1, March 2015.

Clarida R., Galí J., Gertler M., *The Science of Monetary Policy: A New Keynesian Perspective*, "Journal of Economic Literature" 1999, vol. XXXVII.

Clement P., *The term 'macroprudential': origins and evolution*, "BIS Quarterly Review", March 2010.

Cochrane J.H., *Who's Afraid of a Little Deflation?*, "The Wall Street Journal", 17 November, 2014.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Crockett A., *Marrying the Micro- and Macro-prudential Dimensions of Financial Stability*, remarks before the 11th International Conference of Banking Supervisors, 20–21September, Basel 2000.

Currie D., Levine P., *Rules, Reputation and Macroeconomic Policy Coordination*, Cambridge University Press, Cambridge 1993.

Davig T., Leeper E.M., *Generalizing the Taylor Principle*, "American Economic Review" 2007, vol. 97, no. 3.

De Grauwe P., *The ECB as a lender of last resort*, 18 August 2011, www.voxeu.org

De Grauwe P., Ji Y., *Mispricing of Sovereign Risk and Multiple Equilibria in the Eurozone*, "CEPS Working Documents" 2012, no. 361.

De Larosière J., *Report of the High-Level Group on Financial Supervision in the EU*, chaired by Jacques de Larosière, Bruxelles, 25 February, 2009.

De Nicolo G., Dell'Ariccia G., Laeven L., Valencia F., *Monetary Policy and Bank Risk Taking*, IMF Staff Position Note, SPN/10/09, 27 July, 2010.

Demchuk O. i in., *Mechanizm transmisji polityki pieniężnej w Polsce. Co wiemy w 2011 roku?*, „Materiały i Studia" 2012, nr 270, Narodowy Bank Polski.

Dixit A., Lambertini L., *Interactions of Commitment and Discretion in Monetary and Fiscal Policies*, "American Economic Review" 2003, vol. 93.

Draghi M., *Introductory statement to the press conference (with Q&A)*, European Central Bank, 22 January 2015, http://www.ecb.europa.eu/press/pressconf/2015/html/is150122.en.html (27.04.2016).

Duisenberg F., *Monetary and fiscal policy in the euro area*, introduction at the International Monetary Conference, 3 June 2003, Berlin, http://www.ecb.int/press/key/date/2003/html/sp030603.en.html (15.04.2013).

Easterly W., Fischer S., *Inflation and the Poor*, The World Bank and International Monetary Fund, May 2000, Washington.

Eggertsson G.B., Woodford M., *The Zero Bound on Interest Rates and Optimal Monetary Policy*, "Brookings Papers on Economic Activity" 2003, no. 1.

Eichengreen B., Rose A.K., Wyplosz C., *Exchange Market Mayhem: The Antecedents and Aftermath of Speculative Attacks*, "Economic Policy" 1995, vol. 21.

Elliott D.J., Feldberg G., Lehnert A., *The History of Cyclical Macroprudential Policy in the United States*, "Divisions of Research and Statistics and Monetary Affairs" 2013, no. 29, FRB Finance and Economics Discussion Series.

Ericsson N.R., Eilev S.J., Neva A.K., Nymoen R., *Interpreting a Monetary Conditions Index in economic policy*, w: *Topics in monetary policy modelling*, "BIS Conference Papers" 1998, vol. 6.

European Central Bank, *Virtual currency schemes*, October 2012.

European Central Bank, *Virtual currency schemes – a further analysis*, February 2015.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Feige E.L., *New estimates of U.S. currency abroad, the domestic money supply and the unreported Economy*, "MPRA Paper" 2011, no. 34778, Munich Personal Repec Archive.

Fischer S., *Long-Term Contracts, Rational Expectations, and the Optimal Money Supply Rule*, "Journal of Political Economy" 1977, vol. 85, no. 1.

Fischer S., *The Role of Macroeconomic Factors in Growth*, "Journal of Monetary Economics" 1993, vol. 32, no. 2.

Fisher I., *Purchasing Power of Money*, Macmillan, New York 1911.

Fisher I., *The Theory of Interest,* Macmillan, New York 1930.

Friedman M., *A Program for Monetary Stability*, Fordham University Press, New York 1960.

Friedman M., *The Role of Monetary Policy*, "American Economic Review" 1968, vol. 58.

Friedman M., *The Optimum Quantity of Money*, Chicago University Press, Chicago 1969.

Friedman M., *The Counter-Revolution in Monetary Theory*, "IEA Occasional Paper" 1970, no. 33.

Friedman M., *Unemployment versus inflation?*, IEA, London 1975.

Friedman M., *Nobel Lecture: Inflation and Unemployment*, "Journal of Political Economy" 1977, no. 85.

Galí J., *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press, Princeton–New York 2008.

Galí J., *Monetary Policy and Rational Asset Price Bubbles*, "CEPR Discussion Paper" 2013, no. DP9355.

Galí J., *The Effects of a Money-Financed Fiscal Stimulus*, "CEPR Discussion Paper" 2014, no. DP10165.

Gerson M., *Dying Silently In Zimbabwe*, "Washington Post", 20 February, 2008.

Goldberg L.S., *Is the International Role of the Dollar Changing?*, "Current Issues in Economics and Finance" 2010, no. 16(1), Federal Reserve Bank Of New York.

Goodfriend M., King R.G., *The New Neoclassical Synthesis and the Role of Monetary Policy*, "NBER Macroeconomics Annual" 1997, vol. 12.

Goodhart C.A.E., *The Macro-Prudential Authority: Powers, Scope and Accountability*, "OECD Journal: Financial Market Trends" 2011, issue 2.

Gordon R.J., *Price inertia and policy ineffectiveness in the United States*: 1890–1980, "Journal of Political Economy" 1982, no. 90(6).

Greenspan A., *Testimony before the Committee on Banking and Financial Services*, US House of Representatives, 22 July, 1999.

Gros D., *A simple model of multiple equilibria and default*, "CEPS Working Document" 2012, no. 366.

Haldane A., *Central bank psychology*, speech at the Royal Society of Medicine, "The Economist", 22 November, 2014.

Hall R.E., Taylor J.B., *Makroekonomia*, wydanie 3 zmienione, Wydawnictwo Naukowe PWN, Warszawa 2000.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Hall R.E., *Separating the Business Cycle from Other Economic Fluctuations*, Jackson Hole Conference, August 2005.

Hanke S., Krus N., *World Hyperinflations*, "CATO Institute Working Paper", 15 August, 2012.

Hałka A., Leszczyńska A., *Wady i zalety wskaźnika cen towarów i usług konsumpcyjnych – szacunki obciążenia dla Polski*, „Gospodarka Narodowa" 2011, nr 9.

Hiroshi F., Okina K., Shiratsuka S., *Monetary Policy under Zero Interest Rate: Viewpoints of Central Bank Economists*, Bank of Japan, „Monetary and Economic Studies", February 2001.

IMF, *Macroprudential Policy: An Organizing Framework*, 14 March, 2011a, Washington.

IMF, *Exchange Arrangements and Exchange Restrictions*, Annual Report, 2012, 2014.

Issing O., *Lessons for Monetary Policy. What Should the Consensus Be?*, "IMF Working Paper" 2011, no. WP/11/97.

Jakóbik W., Reguły polityki makroekonomicznej, w: *Polityka gospodarcza. Teoria i realia*, red. J. Stacewicz, Oficyna Wydawnicza SGH, Warszawa 2008.

Kahn G.A., *Beyond Inflation Targeting: Should Central Banks target the Price Level?*, "FRB of Kansas City Economic Review" 2009, no. 3.

Kahneman D., Tversky A., *Choices, Values and Frames*, Cambridge University Press, Cambridge 2000.

Kapuściński M. i in., *Mechanizm transmisji polityki pieniężnej w Polsce. Co wiemy w 2013 roku?*, „Materiały i Studia" 2014, nr 306, Narodowy Bank Polski.

Kaźmierczak A., *Polityka pieniężna w gospodarce rynkowej*, Wydawnictwo Naukowe PWN, Warszawa 2003.

Keynes J.M., *Ogólna teoria zatrudnienia, procentu i pieniądza*, Londyn 1936.

Kindleberger C., *Manias, Panics, and Crashes: A History of Financial Crises*, wydanie 5, Palgrave Macmillan, Basingstoke 2005.

Kluza S., *Zastosowanie indeksów Divisia w Polsce*, nieopublikowana praca doktorska, SGH, Warszawa 2001.

Klyuev V., de Imus Ph., Srinivasan K., *Unconventional Choices for Unconventional Times: Credit and Quantitative Easing in Advanced Economies*, IMF Staff Position Note, SPN/09/27, 2009.

Kokoszczyński R., *Współczesna polityka pieniężna w Polsce*, PWE, Warszawa 2004.

Koo R., *The Holy Grail of Macroeconomics – Lessons from Japan's Great Recession*, Wiley, New York 2008.

Koo R., *The world in balance sheet recession: causes, cure, and politics*, "Real-World Economics Review" 2011, no. 58.

Kot A., *Metody kwantyfikacji restrykcyjności monetarnej, fiskalnej i policy mix w krajach akcesyjnych*, „Bank i Kredyt" 2003, nr 6.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Krugman P.R., *It's Baaack: Japan's Slump and the Return of the Liquidity Trap*, "Brookings Papers on Economic Activity" 1998, no. 2.

Krugman P.R., *Crisis in Prices?*, "New York Times", 31 December, 2002.

Kydland F.E., Prescott E.C., *Rules Rather Than Discretion: The Inconsistency of Optimal Plans*, "Journal of Political Economy" 1977, no. 85.

Lim C. i in., *Macroprudential Policy: What Instruments and How to Use Them? Lessons from Country Experiences*, "IMF Working Paper" 2011, no. WP/11/238.

Lucas R.E., Jr., *Econometric policy evaluation: A critique*, "Carnegie-Rochester Conference Series on Public Policy" 1976, no. 1.

Łyziak T., *Consumer inflation expectations in Poland*, "ECB Working Paper" 2003, no. 287.

Łyziak T., *Oczekiwania inflacyjne*, w: *Polityka pieniężna*, red. A. Sławiński, C.H. Beck, Warszawa 2011.

Łyziak T., *Oczekiwania inflacyjne w Polsce*, „Materiały i Studia" 2012, nr 271, Narodowy Bank Polski.

Mankiw G.N., *Real Business Cycles: A New Keynesian Perspective*, "Journal of Economic Perspectives" 1989, no. 3.

Mankiw G.N., *A Quick Refresher Course in Macroeconomics*, "Journal of Economic Literature" 1990, vol. XXVIII.

Mankiw G.N., *Comments on: "Separating the Business Cycle from Other Economic Fluctuations" by Robert E. Hall*, Jackson Hole Conference, August 2005.

Mankiw G.N., Romer D., *New Keynesian Macroeconomics*, vol. 1–2, MIT Press, Cambridge 1991.

Marciniak Z., *Niekonwencjonalna polityka pieniężna*, w: *Wybrane aspekty funkcjonowania Polski w Unii Europejskiej*, red. G. Wojtkowska-Łodej, H. Bąk, Oficyna Wydawnicza SGH, Warszawa 2014.

Marczewski K., *Normatywne podejście do polityki gospodarczej*, w: *Polityka gospodarcza. Teoria i realia*, red. J. Stacewicz, Oficyna Wydawnicza SGH, Warszawa 2008.

Marszałek P., *Coordination of monetary and fiscal policy*, "The Poznań University of Economics Review" 2003, vol. 3, no. 2.

Meltzer A., *The Phillips Curve and Labor Markets*, "Carnegie-Rochester Conference Series on Public Policy" 1976, vol. 1.

McCallum B.T., *Robustness Properties of a Rule for Monetary Policy*, "Carnegie-Rochester Conference Series on Public Policy" 1988, vol. 29.

McCallum B.T., *Crucial Issues Concerning Central Bank Independence*, "Journal of Monetary Economics" 1997a, no. 39.

McCallum B.T., *Issues in the Design of Monetary Policy Rules*, "NBER Working Paper" 1997b, no. 6016.

McCallum, B.T., *Nominal GDP Targeting*, Shadow Open Market Committee, 21 October, 2011a.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

McCallum B.T., *Should Central Banks Raise their Inflation Targets? Some Relevant Issues*, "Federal Reserve Bank of Richmond Economic Quarterly" 2011b, no. 2.

Minsky H.P., *The Financial Instability Hypothesis*, "Working Paper" 1992, no. 74, The Jerome Levy Economics Institute of Bard College.

Morris S., Shin H.S., *Rethinking Multiple Equilibria in Macroeconomic Modeling*, "NBER Macroeconomics Annual" 2001, vol. 15.

Muth J.F., *Rational Expectations and the Theory of Price Movements*, "Econometrica" 1961, vol. 29.

NBP, *Średniookresowa strategia polityki pieniężnej na lata 1999–2003*, Rada Polityki Pieniężnej, Warszawa, wrzesień 1998.

NBP, *Raport o inflacji*, Warszawa, lipiec 2013.

NBP, *Raport roczny 2013*, Warszawa 2014.

Niemann S., von Hagen J., *Coordination of monetary and fiscal policies: A fresh look at the issue*, "Swedish Economic Policy Review" 2008, vol. 15.

Obstfeld M., Rogoff K., *Foundations of International Macroeconomics*, MIT Press, Cambirdge 1997.

Okun A.M., *Inflation: The problems and prospects before us*, w: *Inflation: The Problems it Creates and the Policies It Requires*, red. H. Fowler, M. Gilbert, New York University Press, New York 1970.

Okun A.M., Peck M., Smith W., *The 1969 Annual Report of the Council of Economic Advisors*, Government Printing Office, Washington 1969.

Persson T., Tabellini G., *Political Economics and Macroeconomic Policy*, Institute for International Economics, Stockholm 1997, mimeo.

Phelps E.S., Curves Ph., *Expectations of Inflation and Optimal Employment Over Time*, "Economica" 1967, vol. XXXVII.

Phelps E.S., *Money wage dynamics and labor market equilibrium*, "Journal of Political Economy" 1968, vol. 76, no. 4.

Pigou A., *The Veil of Money*, MacMillan, London 1949.

Pissarides C.A., *Equilibrium Unemployment Theory*, Basil Blackwell, Cambridge-Oxford 1990.

Plosser C.I., *Understanding real business cycles*, "Journal of Economic Perspectives" 1989, no. 3(3).

Prescott E., *Theory ahead of business cycle measurement*, "FRB of Minneapolis Quarterly Review" 1986, no. 10(4).

Reinhart C.M., Reinhart V., *Pride Goes Before a Fall: Federal Reserve Policy and Asset Market*, "NBER Working Paper" 2011, no. 16815.

Rogoff K., *The Optimal Degree of Commitment to an Intermediate Monetary Target*, "Quarterly Journal of Economics" 1985, vol. 100, no. 4.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rogoff K., *Costs and benefits to phasing out paper currency*, "NBER Working Paper", May 2014.

Romer D., *The New Keynesian Synthesis*, "Journal of Economic Perspectives" 1993, vol. 7, no. 1.

Rosati D.K., *Dylematy polityki pieniężnej w warunkach swobody przepływów kapitału*, „Bank i Kredyt" 1999, nr. 7–8.

Rosati D.K., *Lot Ikara?*, „Gazeta Bankowa", 25 stycznia 2000.

Rosati D.K., *Przyczyny i mechanizm kryzysu finansowego w USA*, „Ekonomista" 2009, nr 3.

Rosati D.K., *Regulacje makroostrożnościowe a stabilność sektora finansowego*, „Bank i Kredyt" 2014, nr 5.

Rotemberg J., *The New Keynesian microfoundations*, "NBER Macroeconomics Annual" 1987, vol. 2.

Rotemberg J., Woodford M., *An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy*, "NBER Macroeconomics Annual" 1997, vol. 12.

Rothbard M.N., *An Austrian Perspective on the History of Economic Thought*, w: *Economic Thought Before Adam Smith*, vol. 1, Edward Elgar Publishing, Cheltenham 1995.

Rothbard M.N., *Taking Money Back*, "The Freeman", September–October 1995b.

Rozkrut M., *Polityka informacyjna banków centralnych*, w: *Polityka pieniężna*, red. A. Sławiński, C.H. Beck, Warszawa 2011.

Rzońca A., *Kryzys banków centralnych. Skutki stopy procentowej bliskiej zera*, C.H. Beck, Warszawa 2014.

Rzońca A., Ciżkowicz P. *The effects of unconventional monetary policy: what do central banks not include in their models?*, "CASE Seminar Proceedings" 2014, no. 131/2014.

Salvatore D., *International Economics*, wydanie 8, Wiley, Prentice Hall 2004.

Sałata P., *Krzywa dochodowości*, w: *Polityka pieniężna*, red. A. Sławiński, C.H. Beck, Warszawa 2011.

Samuelson P.A., *Economics: An Introductory Analysis*, wydanie 7, McGraw-Hill, New York 1967.

Sargent T.J., *Rational Expectations and Inflation*, Harper and Row, New York 1986.

Sargent T.J., Wallace N., *Rational expectations, the optimal money supply instrument, and the optimal money supply rule*, "Journal of Political Economy" 1975, no. 83(2).

Sargent T.J., Wallace N., *Rational expectations and the theory of economic policy*, "Journal of Monetary Economics" 1976, vol. 2.

Sargent T.J., Wallace N., *Some Unpleasant Monetary Arithmetic*, "Federal Reserve Bank of Minneapolis Quarterly Review" 1981.

Schmitt-Grohé S., Uribe M., *The Case for Temporary Inflation in the Eurozone*, "CEPR Discussion Paper" 2012, no. 9133.

Shapiro C., Stiglitz J.E., *Equilibrium Unemployment as a Worker Discipline Device*, "American Economic Review" 1984, vol. 74.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Sheedy K.D., *Debt and Incomplete Financial Markets: A Case for Nominal GDP Targeting*, "Brookings Papers on Economic Activity" 2014, Spring.

Sims C., *Macroeconomics and Reality*, "Econometrica" 1980, no. 48.

Sławiński A., *Polityka pieniężna*, C.H. Beck, Warszawa 2011.

Stanisławska E., *Differences in consumers' inflation expectations – the case of Poland*, w: *Inflation Expectations. A View from Various Perspectives*, red. E. Stanisławska, E. Tomczyk, Oficyna Wydawnicza SGH, Warszawa 2010.

Summers L., *Some Sceptical Observations on Real Business Cycle Theory*, "Federal Reserve Bank of Minneapolis Review" 1986, vol. 10, no. 4.

Summers L., *Relative Wages, Efficiency Wages and Keynesian Unemployment*, "American Economic Review" 1988, vol. 78.

Summers L., *How Should Long-Term Monetary Policy Be Determined?*, "Journal of Money, Credit and Banking" 1991, vol. 23, no. 3.

Sumner S.B., *Re-Targeting the Fed*, "National Affairs" 2011, no. 9.

Sumner S.B., *Nominal GDP Targeting: A Simple Rule to Improve Fed Performance*, "Cato Journal" 2014, vol. 34, no. 2.

Szpunar P., *Polityka pieniężna. Cele i warunki skuteczności*, PWE, Warszawa 2000.

Szpunar P., *Rola polityki makroostrożnościowej w zapobieganiu kryzysom finansowym*, „Materiały i Studia" 2012, nr 278, NBP.

Taylor J.B., *Aggregate dynamics and staggered contracts*, "Journal of Political Economy" 1980, vol. 88, no. 1.

Taylor J.B., *Discretion versus policy rules in practice*, "Carnegie-Rochester Conference Series on Public Policy" 1993a, no. 39.

Taylor J/B., *Macroeconomic Policy in a World Economy. From Econometric Design to Practical Operation*, W.W. Norton, New York 1993b.

Taylor J.B., *Reassessing Discretionary Fiscal Policy*, "Journal of Economic Perspectives" 2000, vol. 30, no. 3.

Taylor J.B., *Getting off Track*, Hoover Press, Stanford 2009.

Tobin J., *The cruel dilemma*, w: *The Battle Against Unemployment*, red. A.M. Okun, W.W. Norton and Company, New York 1972.

Tobin J., *The New Economics One Decade Older*, Princeton University Press, Princeton–New York 1974.

Tomczyk E., *Racjonalność oczekiwań polskich przedsiębiorców: reguła czy wyjątek? Wnioski z testu koniunktury IRG SGH*, w: *Koniunktura gospodarcza – 20 lat doświadczeń Instytutu Rozwoju Gospodarczego SGH*, red. E. Adamowicz, Warszawa 2008.

Turner A, *The Case for Monetary Finance – an Essentially Political Issue*, IMF Jacques Polak Research Conference, 5 November, 2015.

Tymoczko D., *Wybrane aspekty funkcjonowania polskiego systemu finansowego*, „Zeszyty Naukowe Kolegium Gospodarki Światowej" 2015, nr 34.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Werning I., *Managing a Liquidity Trap: Monetary and Fiscal Policy*, "NBER Working Paper" 2011, no. 17344.

Williams J.C., *Cash Is Dead! Long Live Cash!*, An Essay by President and CEO, Federal Reserve Bank of San Francisco, Annual Report, 2012.

Williams J.C., *Monetary Policy at the Zero Lower Bound. Putting Theory into Practice*, Hutchins Centre on Fiscal and Monetary Policy at Brookings, 16 January, 2014.

Wojtyna A., *Szkice o niezależności banku centralnego*, Wydawnictwo Naukowe PWN, Warszawa 1998.

Wojtyna A., *Kryzysy finansowe a skuteczność polityki makroekonomicznej*, „Bank i Kredyt" 1999, nr. 7–8.

Wojtyna A., *Ewolucja keynesowska a główny nurt ekonomii*, Wydawnictwo Naukowe PWN, Warszawa 2000.

Wojtyna A., *Demokratyczna kontrola nad polityką banku centralnego – kontrowersje teoretyczne i wnioski dla Polski*, „Ekonomista" 2002, no. 5.

Wojtyna A., *Banki centralne po kryzysie: czy konieczne są zmiany w strategii bezpośredniego celu inflacyjnego?*, „Gospodarka Narodowa" 2012, nr. 7–8.

Woodford M., *Fiscal Requirements for Price Stability*, "Journal of Money, Credit and Banking" 2001, vol. 33.

Woodford M., *Interest and Prices; Foundations of a Theory of Monetary Policy*, University Press, Princeton–Oxford 2003.

Woodford M., *Methods of Policy Accommodation at the Interest-Rate Lower Bound*, paper presented at the Federal Reserve Bank of Kansas City Symposium on "The Changing Policy Landscape", Jackson Hole, 31 August, 2012.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozdział 6

POLITYKA KURSU WALUTOWEGO

6.1. Kurs walutowy: podstawowe pojęcia

6.1.1. Kurs walutowy jako cena waluty obcej

Kurs walutowy jest ceną waluty zagranicznej wyrażoną w walucie krajowej (tzw. definicja europejska) lub ceną waluty krajowej wyrażoną w walucie zagranicznej (tzw. definicja brytyjska). Polityka kursu walutowego stanowi więc świadome działanie władzy gospodarczej – banku centralnego, rządu, ministra finansów – mające na celu ustalanie ceny waluty krajowej wobec walut zagranicznych na określonym poziomie. Kształtowanie tej ceny – tj. kursu waluty – może odbywać się w sposób bezpośredni, poprzez administracyjne wyznaczenie określonego poziomu kursu, lub w sposób pośredni, poprzez wpływanie na warunki wymiany walut i czynniki określające cenę waluty, takie jak na przykład ustalanie zakresu wymienialności waluty krajowej na waluty obce lub ustalanie opłat związanych z wymianą.

Technicznie rzecz biorąc, kurs walutowy można wyznaczyć na dwa sposoby. Ogromna większość krajów – w tym Polska – stosuje tzw. europejską definicję kursu. W tym ujęciu kurs walutowy wyrażony jest jako liczba jednostek waluty narodowej przypadających na jednostkę waluty obcej. Na przykład 28 listopada 2014 roku oficjalnie ogłoszony przez Narodowy Bank Polski kurs dolara amerykańskiego w kraju wynosił 3,3605 PLN/USD, a kurs euro 4,1814 PLN/EUR. Z kolei Wielka Brytania tradycyjnie stosuje brytyjską metodę określania kursu. W tym przypadku kurs walutowy wyrażony jest jako liczba jednostek waluty obcej przypadających na jednostkę waluty krajowej. Przykładowo 28 listopada 2014 roku kursy te, ogłoszone przez Bank Anglii, wyniosły odpowiednio 1,5713 USD/GBP i 1,2628 EUR/GBP.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W tabelach zawierających notowania kursów wykorzystuje się powszechnie przyjęte międzynarodowe kody walut lub ich symbole. Kody walut narodowych składają się z trzech liter, z których dwie pierwsze stanowią skrót nazwy danego kraju, a trzecia stanowi najczęściej pierwszą literę nazwy danej waluty[1]. W powszechnym użyciu pozostają także tradycyjne jedno- lub dwuznakowe symbole walut narodowych, z których najbardziej znane są takie jak $, €, ¥ oraz £, oznaczające odpowiednio dolara amerykańskiego, euro, jena japońskiego i funta brytyjskiego[2].

Kurs walutowy może być ustalony administracyjnie przez władze danego kraju lub kształtować się swobodnie na rynku walutowym jako cena równoważąca popyt na waluty obce z ich podażą. W pierwszym przypadku mamy do czynienia z kursem stałym (lub sztywnym), w drugim zaś z kursem płynnym. Zmiana kursu stałego, dokonana oficjalnie przez władze, nosi miano dewaluacji (kiedy cena waluty obcej rośnie) lub rewaluacji (kiedy cena waluty obcej spada). Z kolei zmiana kursu płynnego, dokonująca się za sprawą oddziaływania mechanizmu rynkowego, czyli pod wpływem zmian podaży walut i popytu na waluty, nazywana jest deprecjacją (kiedy cena waluty obcej rośnie) lub aprecjacją (kiedy cena waluty obcej spada).

Poziom kursu walutowego określa opłacalność transakcji międzynarodowych. Ogromna większość krajów świata posługuje się własną walutą, w której wyrażone są wewnętrzne ceny i płace, oraz dokonywane są płatności. Eksporterzy, sprzedający swoje produkty za granicę, otrzymują jednak z reguły należność w walutach obcych, które następnie są przeliczane na walutę krajową po kursie obowiązującym w momencie rozliczenia transakcji. Z kolei importerzy, chcąc nabyć zagraniczne towary lub usługi, muszą zakupić w tym celu walutę obcą po określonym kursie. Podobnie dzieje się w przypadku zagranicznej sprzedaży lub zakupu aktywów. We wszystkich tych przypadkach kurs walutowy pozwala określić opłacalność poszczególnych transakcji. Na przykład w przypadku eksportu wzrost kursu walut obcych oznacza większe przychody w walucie krajowej dla eksporterów, a więc poprawia ich sytuację finansową. Zarazem ten sam wzrost kursu walut obcych oznacza jednak wyższe koszty zakupu po stronie importerów, a więc czyni import mniej opłacalnym. Odmienny w skutkach jest zaś spadek kursu waluty obcej.

---

[1] Pełną listę standardowych kodów walut narodowych zawiera norma międzynarodowa ISO 4217, ustalona i podana do informacji publicznej przez Międzynarodową Organizację Standaryzacji (The International Organization for Standardization).

[2] Tradycyjnym symbolem złotego jest skrót „zł". W większości oficjalnych i bankowych dokumentów wykorzystuje się jednak trzyliterowe kody standardowe ISO.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Trzeba również zauważyć, że zmiany te mają bezpośredni wpływ na poziom cen krajowych, koszty i rentowność produkcji, stopę przychodu z zainwestowanego kapitału, a pośrednio także na poziom płac i zatrudnienie. Dlatego też przedsiębiorstwa prowadzące wymianę zagraniczną są zawsze żywo zainteresowane kształtowaniem się kursu. Dotyczy to zresztą nie tylko eksporterów oraz importerów dóbr i usług, ale także wszystkich uczestników rynku, którzy dokonują transakcji wymagających wymiany walut obcych na walutę krajową, i odwrotnie. Transakcje te obejmują zarówno eksport i import oraz zakup i sprzedaż aktywów, jak i udzielanie oraz zaciąganie kredytów zagranicznych, a także ich spłatę, zakup aktywów za granicą, bezpośrednie inwestycje zagraniczne oraz międzynarodowe transfery dochodów i czynników produkcji, takich jak wynagrodzenia uzyskane za granicą czy transfery zysków, dywidend i płatności odsetkowych. Do transakcji tych należą również wydatki na turystykę zagraniczną i transfery jednostronne, na przykład wypłaty środków z funduszy strukturalnych i spójnościowych z budżetu Unii Europejskiej dla poszczególnych państw członkowskich[3]. Ze względu na istotny wpływ na opłacalność produkcji krajowej, a w konsekwencji na poziom produkcji, inwestycji oraz zatrudnienia, kurs walutowy pozostaje także przedmiotem zainteresowania władz gospodarczych kraju.

6.1.2. Kurs nominalny, realny i efektywny

Kurs walutowy przedstawiony jako cena jednostki waluty obcej wyrażona w jednostkach waluty krajowej nazywa się kursem nominalnym. Sam kurs nominalny (lub jego zmiana) nie daje jednak pełnego obrazu opłacalności transakcji międzynarodowej, ponieważ zależy ona także od wzajemnych relacji cen krajowych i zagranicznych. Dopiero po przeliczeniu cen zagranicznych na ceny krajowe można stwierdzić, czy transakcja eksportowa lub importowa jest w danym przypadku opłacalna. Na przykład, gdy cena krajowa określonego towaru wynosi 100 PLN, a jego cena zagraniczna 25 EUR, transakcja eksportowa stanie się opłacalna pod warunkiem, że kurs walutowy osiągnie wartość co najmniej 4 PLN/EUR. Podobnie zresztą wahania kursu nominalnego niewiele wykażą, jeśli równocześnie nie weźmie się pod uwagę zmian w zakresie cen krajowych i zagranicznych. Przykładowo spadek kursu waluty krajowej w ciągu roku z 4 PLN/EUR na 4,20 PLN/EUR

[3] Oczywiście kurs walutowy odnosi się w tym przypadku do państw członkowskich nienależących do obszaru wspólnej waluty.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

– czyli o 5% – nie spowoduje poprawy opłacalności eksportu, jeśli w tym samym czasie ceny krajowe produktów eksportowanych wzrosną w wyniku inflacji o 10%. Wówczas opłacalność sprzedaży nakierowanej na eksport ulegnie pogorszeniu w relacji do opłacalności sprzedaży na rynku krajowym. Podobnie dzieje się, jeśli w tym samym czasie ceny zagraniczne produktów eksportowanych spadną o 10% – wtedy opłacalność eksportu ulega pogorszeniu, a nie poprawie.

Dlatego też należy dokonać rozróżnienia między kursem nominalnym (*nominal exchange rate*) a kursem realnym (*real exchange rate*). Kurs nominalny to cena waluty zagranicznej wyrażona w walucie krajowej. Z kolei kurs realny jest po prostu stosunkiem cen krajowych P do cen zagranicznych P^*, czyli:

$$e_R = \frac{P}{P^*}. \tag{6.1}$$

Związek między kursem nominalnym i kursem realnym określają relacje cen krajowych i zagranicznych. Oznaczając kurs nominalny jako e_N, kurs realny e_R można też przedstawić następująco:

$$e_R = e_N \left(\frac{P^*}{P} \right). \tag{6.1a}$$

Gdy cena zagraniczna P^* przeliczona na walutę krajową po kursie e_N równa się dokładnie cenie krajowej P, kurs realny e_R będzie równy jedności. Kiedy cena krajowa będzie jednak rosła przy danej cenie zagranicznej i stałym kursie nominalnym, wówczas wartość e_R spadnie. Mówi się wówczas o aprecjacji realnego kursu waluty krajowej (lub o deprecjacji realnego kursu waluty zagranicznej). Również odwrotnie, gdy cena krajowa spada przy danej cenie zagranicznej i stałym kursie nominalnym wartość e_R wzrasta i mamy wówczas do czynienia z deprecjacją realnego kursu waluty krajowej (lub aprecjacją realnego kursu waluty zagranicznej).

Wykorzystanie formuł (6.1) i (6.1a) do wyliczania kursu realnego pozostaje jednak kłopotliwe, ponieważ trudno jest przedstawić rozmaite ceny krajowe i ceny zagraniczne w zagregowanej postaci jednej liczby. Dlatego też w praktyce stosuje się wskaźniki zmian cen i zmian kursu. Kurs realny jest wtedy iloczynem *wskaźnika* zmian kursu nominalnego oraz relacji *wskaźnika* zmian cen krajowych do *wskaźnika* zmian cen zagranicznych. Oznaczając wskaźnik zmian poziomu cen krajowych wyrażony w walucie krajowej i wskaźnik zmian poziomu cen zagranicznych wyrażony w walucie zagranicznej odpowiednio jako $\dot{P}$ oraz $\dot{P}^*$, a wskaźniki zmian

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kursu nominalnego i kursu realnego odpowiednio jako $\dot{e}_N$ oraz $\dot{e}_R$, otrzymujemy definicję kursu realnego w postaci wskaźnika:

$$\dot{e}_R = \dot{e}_N \left(\frac{\dot{P}}{\dot{P}^*} \right). \tag{6.1b}$$

Równanie w postaci wskaźnikowej (6.1b) jest użyteczne, ponieważ w praktycznym zastosowaniu istotne są przede wszystkim zmiany realnego kursu w czasie, a nie sama jego wartość w danym momencie. Załóżmy, że kurs nominalny euro rośnie w ciągu roku z 4 PLN/EUR do 4,20 PLN/EUR, czyli o 5%, z kolei ceny zagraniczne utrzymują się na stałym poziomie. Jeśli w tym samym czasie ceny krajowe nie wzrosną, kurs realny euro zwiększy także o 5%, ponieważ stosunek wskaźników zmiany kursu nominalnego i zmiany cen wyniesie $1{,}05\frac{1{,}00}{1{,}00} = 1{,}05$ (równanie 6.1b). W takim przypadku mówi się o aprecjacji realnego kursu euro (ponieważ relatywna wartość nabywcza euro wzrosła) albo – co oznacza to samo – o deprecjacji realnego kursu złotego (ponieważ relatywna wartość nabywcza złotego zmalała). Jeśli w tym samym czasie ceny krajowe także wzrosną o 5%, kurs realny nie ulegnie zmianie, gdyż $1{,}05\frac{1{,}00}{1{,}05} = 1{,}00$. Wreszcie, kiedy ceny krajowe w tym samym czasie wzrosną o więcej niż 5%, na przykład o 10%, kurs realny euro spadnie o około 4,5%, ponieważ $1{,}05\frac{1{,}00}{1{,}10} \simeq 0{,}955$. Mamy wówczas do czynienia z deprecjacją realnego kursu euro lub z aprecjacją realnego kursu złotego. Należy jednak zawsze pamiętać o tym, że kurs realny jest miarą relatywną – i to w podwójnym sensie: ukazuje bowiem zmiany kursu nominalnego w relacji do zmian cen krajowych oraz w relacji do okresu bazowego. Wskaźnik ten dostarcza więc bardziej wiarygodnej informacji na temat zmian opłacalności eksportu i importu w danym okresie.

Kursy realne obliczane są najczęściej jako zmiany kursów nominalnych w relacji do zmian krajowych i zagranicznych cen towarów konsumpcyjnych, lecz można je także wyliczyć w relacji do zmian cen produkcji przemysłowej, wskaźników inflacji bazowej czy też w relacji do wskaźnika jednostkowych kosztów pracy. Interpretacja

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zmian kursu realnego będzie w każdym z tych przypadków nieco inna, lecz sens ekonomiczny pozostanie zawsze ten sam.

Inne ważne rozróżnienie dotyczy kursu nominalnego i kursu efektywnego. Gdy przychodzi ocenić opłacalność całkowitej wartości eksportu danego kraju, kurs walutowy odnoszący się tylko do jednej waluty obcej, na przykład do dolara lub euro, okazuje się miarą niewystarczającą, ponieważ eksport danego kraju kierowany jest na różne rynki i rozliczany w różnych walutach obcych. Kursy tych walut w stosunku do złotego mogą zmieniać się inaczej niż kursy dolara lub euro, co oznacza, że opłacalność transakcji rozliczanych w innych walutach obcych będzie większa lub mniejsza niż opłacalność transakcji rozliczanych w dolarach lub euro. Dlatego też w analizie opłacalności handlu zagranicznego, lub szerzej – konkurencyjności międzynarodowej danej gospodarki, często posługujemy się pojęciem kursu efektywnego. Nominalny kurs efektywny (*nominal effective exchange rate* – NEER) to średni kurs waluty krajowej wobec „koszyka" walut obcych wykorzystywanych w handlu zagranicznym, przy czym „wagi" poszczególnych walut znajdujących się w „koszyku" pozostają równe udziałom tych walut w ogólnej sumie rozliczeń wszystkich transakcji w eksporcie oraz imporcie dóbr i usług. Z kolei realny kurs efektywny (*real effective exchange rate* – REER) oznacza nominalny kurs efektywny skorygowany zmianami cen krajowych i zagranicznych zgodnie z formułą (6.1).

Wykres 6.1. Zmiany nominalnego efektywnego kursu walutowego (NEER) i kursu nominalnego EUR/PLN w latach 1999–2014 (2005 = 100, wzrost wskaźnika oznacza aprecjację złotego)

Uwaga: NEER obliczony w stosunku do walut 42 najważniejszych krajów partnerskich.

Źródło: opracowanie własne na podstawie danych Eurostatu i NBP.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Na wykresie 6.1 przedstawiono zmiany nominalnego efektywnego kursu złotego na tle przeobrażeń wskaźnika nominalnego kursu euro do złotego (EUR/PLN). Biorąc pod uwagę cały piętnastoletni okres objęty analizą, należy stwierdzić, że kurs złotego nie ulegał wówczas systematycznej aprecjacji ani deprecjacji, lecz w jego ramach mieliśmy do czynienia z kilkuletnimi cyklami umacniania się lub osłabiania kursu. W szczególności zaś w latach 2004–2008 doszło do silnej aprecjacji złotego, a po wybuchu kryzysu finansowego nastąpił gwałtowny spadek kursu. Jednocześnie występowały zauważalne różnice między kursem NEER a kursem rynkowym. W latach 1999–2002 kurs NEER był niższy niż kurs rynkowy głównie ze względu na silniejszy kurs dolara wobec euro, co z kolei rzutowało na słabszą pozycję złotego w stosunku do koszyka walut wykorzystywanych w handlu zagranicznym w porównaniu z kursem wobec euro. Z kolei w latach 2009–2014 kurs NEER był wyższy niż kurs rynkowy, ponieważ w tym czasie kurs dolara wobec euro uległ osłabieniu.

6.1.3. Kurs parytetu siły nabywczej

Omawiając poszczególne rodzaje kursów walutowych, należy wreszcie wspomnieć o kursie parytetu siły nabywczej (*purchasing power parity* – PPP). Kurs PPP to stosunek poziomów cen porównywalnych i reprezentatywnych koszyków takich samych dóbr i usług w kraju oraz za granicą. W przeciwieństwie do kursu rynkowego, który odzwierciedla bieżące relacje popytu i podaży walut obcych, kurs PPP wyraża relację faktycznej siły nabywczej obu walut, mierzonej zdolnością do nabycia tego samego (szerokiego) zestawu dóbr i usług w kraju oraz za granicą. Kurs rynkowy z reguły odbiega od kursu PPP, ponieważ relacje popytu i podaży walut obcych zależą głównie od bieżących wielkości eksportu oraz importu, a także od pozostałych bieżących transakcji finansowych. Kurs rynkowy odzwierciedla więc w tendencji relacje cen krajowych i cen zagranicznych towarów wymiennych (*tradables*), czyli będących przedmiotem wymiany międzynarodowej, skorygowane relacją między napływem i odpływem kapitału finansowego. Tymczasem kurs PPP obejmuje porównanie cen wszystkich towarów, w tym niewymiennych (*nontradables*), takich przede wszystkim jak usługi krajowe (transport miejski, czynsze, usługi remontowe, handel detaliczny, usługi gastronomiczne oraz związane z zakwaterowaniem itp.). Dzięki temu kurs PPP stanowi szerszą i bardziej kompletną miarą rzeczywistej siły nabywczej waluty krajowej niż kurs rynkowy.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ponieważ spora część wytwarzanego PKB nie jest przedmiotem wymiany międzynarodowej, ceny dóbr i usług nieuczestniczących w wymianie nie podlegają arbitrażowi w skali międzynarodowej, a zatem nie wyrównują się pomiędzy poszczególnymi krajami. Dotyczy to przede wszystkim usług niewymiennych, a także w znacznej mierze podstawowych produktów rolno-spożywczych. Ich ceny zależą od stopnia rozwoju gospodarczego danego kraju, a ściślej rzecz biorąc, od poziomu dochodów – im wyższy poziom dochodu na mieszkańca, tym zasadniczo wyższy poziom cen towarów niewymiennych. Dlatego też w krajach znajdujących się na niższym poziomie rozwoju gospodarczego ceny towarów niewymiennych są odpowiednio niższe od cen tych towarów w krajach rozwiniętych. Oznacza to jednak, że jednostka waluty krajowej w kraju mniej rozwiniętym ma w rzeczywistości większą siłę nabywczą, niż wynikałoby to z przeliczenia jednostki waluty obcej na walutę krajową po kursie rynkowym. Ramka 6.1 przedstawia prosty przykład liczbowy, który pozwoli lepiej zrozumieć tę różnicę.

Ramka 6.1. Kurs parytetu siły nabywczej – przykład liczbowy

Załóżmy, że kraj A jest krajem rozwiniętym (na przykład Niemcy), kraj B to natomiast kraj szybko rozwijający się, który „dogania" kraje rozwinięte (na przykład Polska), i że poziom PKB na mieszkańca w kraju B stanowi, po przeliczeniu po kursie rynkowym, 40% poziomu osiągniętego w kraju A. Przyjmijmy też, że reprezentatywny koszyk konsumpcyjny w kraju A ma wartość 1000 EUR, z czego 500 EUR stanowią produkty wymienne, takie jak odzież, obuwie, sprzęt RTV, komputery czy samochody, oraz 500 EUR produkty niewymienne, takie jak usługi transportowe, dostawy wody i ciepła, czynsze, usługi fryzjerskie, naprawa sprzętu RTV czy usługi remontowe. O ile ceny produktów wymiennych wyrównują się w skali międzynarodowej dzięki swobodzie wymiany zagranicznej (arbitraż cenowy), o tyle ceny towarów niewymiennych ustalane są przez dostawców na poziomie zbliżonym do jednostkowych kosztów produkcji, głównie wynagrodzeń. Wartość towarów wymiennych w reprezentatywnym koszyku w cenach kraju B wynosi 2000 PLN, zgodnie z kursem rynkowym e^R = 2000 PLN / 500 EUR = 4 PLN/EUR. Przyjmijmy obecnie, że w związku z różnicą utrzymującą się na poziomie dochodów (i wynagrodzeń) ceny towarów niewymiennych w kraju B stanowią 40% cen obowiązujących w kraju A, czyli 0,4*500 EUR = 200 EUR. Wówczas wartość reprezentatywnego koszyka konsumpcyjnego w cenach kraju B wyniesie 500 EUR + 200 EUR = 700 EUR.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Oznaczając ceny towarów wymiennych i niewymiennych w kraju A i w kraju B odpowiednio jako P_T^A, P_{NT}^A, P_T^B, P_{NT}^B, możemy zapisać kurs parytetu siły nabywczej w następującej postaci:

$$PPP = \frac{P_T^B + P_{NT}^B}{P_T^A + P_{NT}^A} = \frac{(500+200)4zł/€}{500+500} = \frac{2800\ zł}{1000\ €} = 2{,}8\ zł/€.$$

Oznacza to, że jednostka dochodu w kraju B ma większą siłę nabywczą niż jednostka dochodu w kraju A, czyli za 4 PLN można kupić więcej dóbr i usług w kraju B niż za 1 EUR w kraju A. Im większe odchylenie w dół kursu PPP do kursu rynkowego, tym większa jest faktyczna siła nabywcza waluty krajowej w relacji do waluty zagranicznej. W rozpatrywanym przypadku mamy $\frac{PPP}{e^R} = \frac{2{,}8}{4{,}0} = 0{,}7$. Ogólnie można więc zapisać:

$$PPP = \left[\alpha P_T + (1-\alpha)P_{NT}\right] / \left[\alpha^* P_T^* + \left(1-\alpha^*\right)P_{NT}^*\right], \tag{6.2}$$

gdzie α oznacza udział towarów wymiennych w całkowitej konsumpcji, a gwiazdka (*) symbolizuje odpowiednie wartości odnośnie zagranicy.

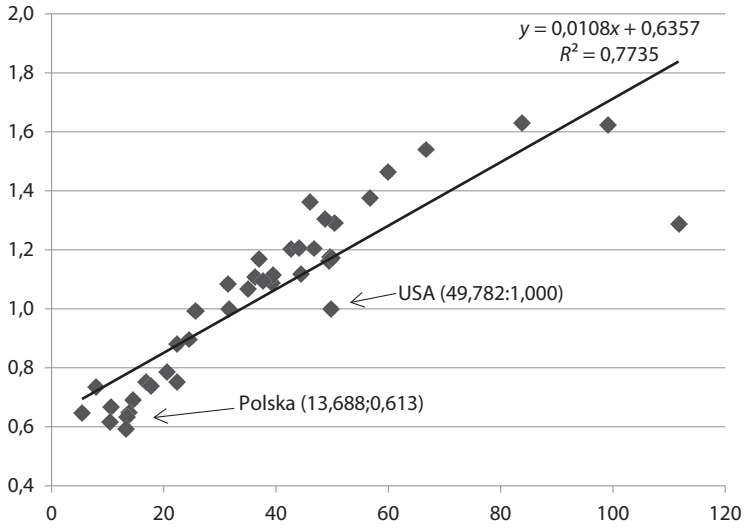
Wykres 6.2 przedstawia relację kursu PPP do kursu oficjalnego w wybranych krajach. Jak widać, występuje wyraźna korelacja dodatnia między stosunkiem PPP i kursu rynkowego a poziomem rozwoju gospodarczego mierzonym wielkością PKB na mieszkańca (współczynnik determinacji $R^2 = 0{,}7735$). Kurs oficjalny zaniża rzeczywistą wartość walut obowiązujących w mniej zamożnych państwach, co oznacza, że siła nabywcza dochodów osiąganych w tych krajach jest w rzeczywistości wyższa, niż wynikałoby to z przeliczenia po kursie oficjalnym. Najwyższe wartości relacji PPP/e^R notują kraje najbogatsze (Norwegia, Szwajcaria), a najniższe kraje najmniej zamożne (Turcja, Rosja, a także Polska).

Ciekawą ilustracją zależności przedstawionej na wykresie 6.2 jest także popularny test stosowany przez tygodnik brytyjski "The Economist", polegający na porównaniu cen standardowej kanapki Big Mac sprzedawanej w restauracjach McDonald's w różnych krajach. Ponieważ Big Mac jest produktem wystandaryzowanym, przygotowywanym na podstawie tej samej receptury (zawierającym identyczne składniki dostarczane w tych samych ilościach), różnice w cenie tego produktu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w poszczególnych krajach odzwierciedlają przede wszystkim różnice w pozostałych kosztach produkcji. Należą do nich głównie wynagrodzenia pracowników, czynsze, ceny energii elektrycznej, dostaw wody, wywozu odpadów, sprzątania, a także koszty napraw i remontów. W przeważającej mierze są to różnice w cenach dóbr i usług niewymiennych.

Wykres 6.2. Stosunek kursu PPP do kursu oficjalnego a poziom PKB na mieszkańca ('000 USD) w 2012 roku[a]



a) Próba obejmuje 37 krajów (34 kraje OECD oraz Chiny, Rosja i Republika Południowej Afryki), a także EU 27, strefę euro oraz OECD jako całość. Relacja PPP/e^R w przypadku USA przyjęta na poziomie równym jedności.

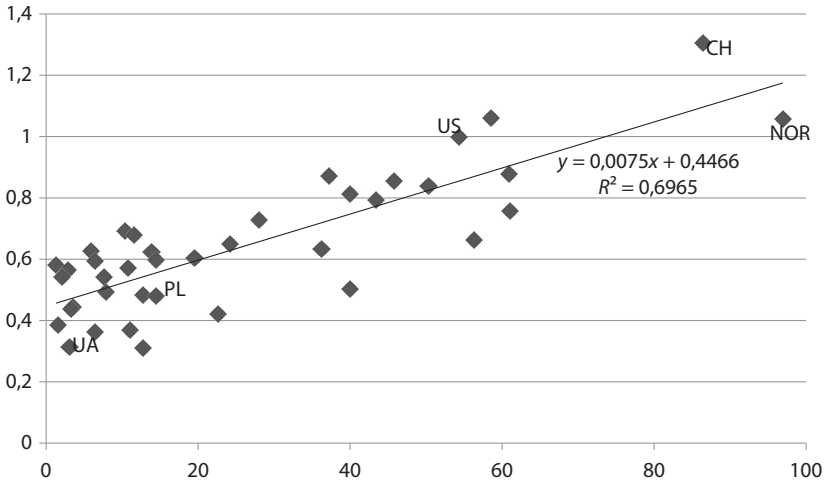
Źródło: opracowanie własne na podstawie danych OECD.

Zależność między indeksem *Big Mac* a poziomem PKB w przeliczeniu na mieszkańca kraju wśród 40 ważniejszych państw ukazana została na wykresie 6.3, gdzie na osi poziomej odłożono poziom PKB na mieszkańca wyrażony w tysiącach dolarów USA, a na osi pionowej wartość indeksu *Big Mac*, definiowanego jako stosunek ceny Big Maca w danym kraju przeliczonej po kursie rynkowym na dolary amerykańskie i ceny Big Maca w USA. Jak widać, wykres ten wykazuje bardzo duże podobieństwo z poprzednim wykresem, co potwierdza, że jednostka waluty krajowej w państwie mniej rozwiniętym gospodarczo ma relatywnie wyższą siłę nabywczą. W styczniu 2016 roku cena Big Maca w jednym z najbogatszych krajów objętych analizą (Szwajcaria) przeliczona po kursie oficjalnym na dolary amerykańskie była ponad czterokrotnie wyższa niż w jednym z najuboższych krajów włączonych do

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

badania (Ukraina)[4] – co oznacza, że za tę samą kwotę dolarów można było kupić jednego Big Maca w Szwajcarii i aż cztery takie kanapki na Ukrainie.

Wykres 6.3. Wskaźnik *Big Mac* a PKB per capita (wg World Bank) w 39 ważniejszych krajach (dane cenowe ze stycznia 2016 roku, dane PKB na mieszkańca w '000 USD z 2014 roku)



Źródło: opracowanie własne na podstawie danych "The Economist", http://www.economist.com/content/big-mac--index (12.03.2016) oraz danych Banku Światowego.

6.1.4. Krótka historia kursów walutowych

Ponieważ przez wiele stuleci pieniądz występował w postać kruszcowej (złoto, srebro), wzajemne relacje wartości jednostek pieniądza emitowanego w różnych krajach były określone na podstawie stosunku zawartości kruszcu w różnych jednostkach pieniężnych. W tym sensie nie istniała polityka kursu walutowego, rozumiana jako świadome działanie władzy gospodarczej służące określaniu wartości jednostki waluty krajowej wobec innych walut. Kurs kształtował się samorzutnie na rynku wymiany pieniądza, wyłącznie na podstawie zawartości kruszcu w jednostce waluty i po uwzględnieniu kosztów związanych z transportem pieniądza i zapewnieniem bezpieczeństwa transakcji.

[4] Próba krajów uwzględnianych przez "The Economist" w szacowaniu *Big Mac Index* różni się składem od próby OECD przedstawionej na wykresie – zob. http://www.economist.com/content/big-mac--index (12.03.2016).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W połowie XIX wieku, kiedy upowszechnił się pieniądz papierowy, kursy poszczególnych walut nadal były określone parytetami wobec złota. W ramach systemu waluty złotej (*gold standard*) jednostka waluty krajowej w postaci pieniądza papierowego reprezentowała bowiem ściśle określoną ilość jednostek wagowych złota (parytet) i stanowiła jednocześnie zobowiązanie emitenta do bezwarunkowej wymiany pieniądza papierowego na złoto według tego parytetu. Kursy walutowe w tym systemie określone były przez relację parytetu złota jednostki pieniężnej jednego kraju do parytetu złota jednostki pieniężnej drugiego kraju. Był to więc system stałych kursów parytetowych (wartości parytetowych – *par values*). Funkcjonowanie systemu waluty złotej wymagało od poszczególnych państw posiadania rezerw złota w odpowiedniej wysokości, tak aby w pełni pokryć wartość pieniądza papierowego znajdującego się w obiegu. System waluty złotej posiadał wszystkie zalety systemu pieniądza kruszcowego – zapewniał stabilną wartość pieniądza i stabilny kurs wymiany – a zarazem był wygodniejszy i tańszy w zastosowaniu. Główna różnica polegała na tym, że system *gold standard* wymagał od władz krajowych dyscypliny monetarnej, czyli rygorystycznego przestrzegania zasady pełnego pokrycia obiegu pieniężnego złotem. Był to warunek trwałości i wiarygodności systemu.

Ważną zaletą systemu waluty złotej był samoczynny mechanizm dostosowawczy, który pozwalał na korygowanie nierównowagi płatniczej występującej między krajami (zob. McKinnon 1963, Eichengreen 1992). Nierównowaga płatnicza w postaci deficytu na rachunku obrotów bieżących w danym kraju powodowała odpływ złota za granicę, a to oznaczało spadek obiegu pieniężnego. Pociągało to za sobą spadek cen krajowych i wzrost konkurencyjności produkcji krajowej, co z kolei pozwalało zwiększyć eksport i obniżyć import. W rezultacie mechanizm quasi-automatycznych przepływów pieniądza i kruszców (*price-specie flow mechanism*) wywoływał zmiany cen relatywnych i zapewniał powrót bilansu płatniczego do równowagi. System ten sprzyjał szybkiemu rozwojowi handlu międzynarodowego w drugiej połowie XIX i na początku wieku XX, dając początek tzw. pierwszej globalizacji.

Pierwsza wojna światowa zburzyła system waluty złotej i położyła kres pierwszej globalizacji. Wielka Brytania, a wraz z nią i jej waluta, utraciły wiodącą pozycję w handlu międzynarodowym, przy jednoczesnym wzroście roli Stanów Zjednoczonych i dolara amerykańskiego. Po zakończeniu wojny kraje rozwinięte podjęły wysiłki w celu ustabilizowania światowych finansów, lecz próby odbudowy systemu *gold standard* po 1925 roku nie dały trwałych rezultatów. Funkcjonowanie mechanizmu dostosowawczego uległo zakłóceniu na skutek protekcjonistycznych i merkantylistycznych praktyk stosowanych przez niektóre państwa, gromadzące

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zapasy złota i administracyjnie ograniczające jego odpływ (USA, Francja). Wielki kryzys lat 1929–1933 wywołał kolejne napięcia. Poszczególne kraje, dążąc do pobudzenia swojej gospodarki i walcząc z masowym bezrobociem, angażowały się w ekspansywną politykę pieniężną i fiskalną, a wobec narastającej nierównowagi w bilansie płatniczym dewaluowały swoje waluty i ograniczały, a następnie zawieszały wymienialność na złoto. Oznaczało to *de facto* wyjście z systemu *gold standard*. W efekcie stałe kursy walutowe uległy załamaniu i w połowie lat 30. system waluty złotej przestał istnieć.

Po zakończeniu II wojny światowej kraje rozwinięte chciały powrócić do systemu stałych kursów walutowych i do pełnej wymienialności walut. Na międzynarodowej konferencji w amerykańskim kurorcie Bretton Woods w lipcu 1944 roku podjęto decyzję o budowie nowego międzynarodowego systemu walutowego, którego podstawą miał być dolar amerykański. Waluty poszczególnych państw miały mieć stałe kursy wymiany w stosunku do dolara, a ten z kolei powinien być w pełni wymienialny na złoto po stałym parytecie, ustalonym na poziomie 35 USD za uncję. Istniejące ograniczenia w zakresie wymienialności miały być stopniowo znoszone, a o stabilność systemu powinien dbać Międzynarodowy Fundusz Walutowy, pełniący rolę pożyczkodawcy „ostatniej szansy" w przypadku państw, które cierpiały na nierównowagę płatniczą. Nowy system nie oznaczał więc całkowitego odejścia od systemu waluty złotej. Miało ono pełnić nadal rolę stabilizującą, lecz w sposób pośredni – poszczególne waluty powinny posiadać stałe parytety wobec dolara, podczas gdy dolar miał mieć stały parytet wobec złota (*gold exchange standard*). Stał się on w ten sposób główną międzynarodową walutą rezerwową, zastępując w tej funkcji zarówno złoto, jak i waluty innych krajów, głównie Wielkiej Brytanii.

Nowy międzynarodowy system walutowy zaczął funkcjonować w trudnych warunkach. Państwa europejskie, wyniszczone wojną, zmagały się z ustawicznym deficytem płatniczym, a ich waluty pozostawały niewymienialne. Ogromne potrzeby importowe mogły być finansowane tylko dolarami, lecz rosnący popyt na tę walutę nie był w pełni zaspokojony, ponieważ rezerwy złota w USA nie rosły na tyle szybko, aby zapewnić odpowiedni wzrost podaży pieniądza. Również utrzymanie na wiarygodnym poziomie i obrona stałych kursów wymagały od poszczególnych państw przechowywania znacznych rezerw walutowych, głównie w postaci dolarów i złota. Ostry brak płynności międzynarodowej hamował więc rozwój handlu i spowalniał wzrost gospodarczy na świecie. Chroniczny niedobór dolarów i złota jako aktywów rezerwowych w latach 60. skłonił Międzynarodowy Fundusz Walutowy do podjęcia w 1969 roku decyzji o emisji syntetycznego aktywu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

rezerwowego w postaci tzw. specjalnych praw ciągnienia, czyli SDR (*special drawing rights* – zob. ramka 6.2).

Ramka 6.2. Specjalne prawa ciągnienia – SDR

Międzynarodowy Fundusz Walutowy przyznawał określone kwoty SDR-ów poszczególnym państwom członkowskim, według klucza określającego ich udziały w Funduszu. Formalnie rzecz biorąc, SDR-y nie były ani pieniądzem, ani też wierzytelnością wobec MFW – stanowiły raczej szczególny rodzaj aktywów rezerwowych, które każde państwo członkowskie mogło wykorzystać w celu nabycia waluty innego państwa członkowskiego, aby sfinansować powstały deficyt płatniczy.

Wartość jednostki SDR została początkowo ustalona na poziomie równym jednemu dolarowi amerykańskiemu, co odpowiadało równowartości 0,888671 grama czystego złota. Po rozpadzie systemu Bretton Woods w 1973 roku wartość jednostki SDR została ustalona na podstawie koszyka pięciu głównych walut rezerwowych funkcjonujących w owym czasie (dolar amerykański, marka niemiecka, funt brytyjski, jen japoński i frank francuski). Od tamtej pory MFW publikuje codziennie aktualny kurs SDR wobec dolara amerykańskiego, który zmienia się w zależności od zmian kursów walut wchodzących w skład koszyka. Struktura koszyka ustalana jest raz na pięć lat, z uwzględnieniem zmieniającej się roli poszczególnych walut w wymianie międzynarodowej[5]. W tabeli 6.1 przedstawione zostały struktury walutowe koszyka SDR ustalone na początku każdego okresu pięcioletniego w latach 2001–2015. W 2015 roku MFW zdecydował o rozszerzeniu koszyka i włączeniu do niego chińskiego juana. Nowy koszyk wszedł w życie 1 października 2016 roku. Jak widać, włączenie nowej waluty odbyło się głównie kosztem obniżenia udziału euro w koszyku.

MFW przydziela kwoty SDR w miarę ewolucji potrzeb płynnościowych w gospodarce światowej. Pierwsza alokacja z lat 1970–1972 opiewała na kwotę 9,3 mld SDR, a druga alokacja mająca miejsce w latach 1979–1981 na kwotę 12,1 mld SDR. Trzecia alokacja została przeprowadzona w sierpniu 2009 roku w odpowiedzi na kryzys finansowy i objęła kwotę 161,2 mld SDR. Dodatkowo, w tym samym czasie

[5] Określając udziały poszczególnych walut w koszyku SDR, MFW bierze pod uwagę wielkość eksportu danego kraju, wielkość obrotów na rynku walutowym tego kraju, poziom rezerw międzynarodowych utrzymywanych przez kraje członkowskie w danej walucie oraz wielkość bankowych zobowiązań międzynarodowych i papierów wartościowych denominowanych we wskazanej walucie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

MFW dokonał specjalnej, jednorazowej alokacji na kwotę 21,5 mld SDR, przeznaczonej głównie dla krajów, które dołączyły do MFW po 1981 roku i nie korzystały wcześniej z dostępu do SDR. Posiadane przez poszczególne państwa kwoty SDR na rachunkach w MFW pozostają oprocentowane (na koniec 2015 roku stopa oprocentowania wynosiła 0,5%).

Tabela 6.1. Początkowa struktura koszyka walutowego SDR w latach 2001–2015 (dane w %)

Waluta	Symbol	2001–2005	2006–2010	2011–2015	2016–2020[a]
Euro	EUR	30,76	33,95	36,83	30,93
Jen japoński	JPY	13,96	10,91	9,66	8,33
Funt brytyjski	GBP	11,21	10,89	11,19	8,09
Dolar amerykański	USD	44,08	44,25	42,97	41,73
Juan chiński	CNY	–	–	–	10,92
Cały koszyk		100,00	100,00	100,00	100,00

a) Zgodnie z decyzją MFW z listopada 2015 roku.

Źródło: obliczenia własne na podstawie danych MFW.

Dopiero w połowie lat 60. Stany Zjednoczone zdecydowały się na stopniowe zwiększanie podaży pieniądza ponad poziom posiadanych rezerw złota, głównie po to jednak, aby sfinansować rosnący deficyt budżetowy związany z prowadzeniem wojny w Wietnamie. Jak się potem okazało, to naruszenie związku między wielkością obiegu pieniężnego a rezerwami złota w USA stało się początkiem końca międzynarodowego systemu walutowego opartego na dolarze jako walucie rezerwowej. Szybki wzrost podaży dolarów w drugiej połowie lat 60. i w latach 70. spowodował wzrost inflacji w USA, która w ramach systemu sztywnych kursów walutowych była przenoszona do innych krajów, pociągając za sobą między innymi wzrost światowych cen surowców i paliw (głównie ropy naftowej). Kraje europejskie, których waluty były powiązanie z dolarem, zostały zmuszone do stosowania polityki deflacyjnej (czyli restrykcyjnej polityki pieniężnej mającej na celu hamowanie inflacji), której skutkiem stało się spowolnienie wzrostu gospodarczego i wzrost bezrobocia. Spadek wartości dolara sprawił, że niektóre z państw (m.in. Francja) zwróciły się do rządu USA z żądaniem wymiany części rezerw dolarowych na złoto. W tej sytuacji 15 sierpnia 1971 roku prezydent Nixon ogłosił decyzję o zawieszeniu wymienialności dolara na złoto.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Nadal próbowano jednak ratować system oparty na dolarze. 18 grudnia 1971 roku kraje rozwinięte podpisały tzw. Smithsonian Agreement, który przewidywał obniżenie (dewaluację) parytetu dolara wobec złota do 37 USD za uncję i zwiększenie dopuszczalnych wahań kursów poszczególnych walut wokół parytetów centralnych. Były to jednak drobne korekty, które nie zlikwidowały fundamentalnych przyczyn powstałej nierównowagi. W sytuacji, kiedy polityka pieniężna i fiskalna w USA była podporządkowana wyłącznie wewnętrznym celom politycznym i gospodarczym, utrzymanie stabilności międzynarodowego systemu walutowego okazało się niemożliwe. W marcu 1973 roku, gdy w obawie przed inflacją i wzrostem kosztów importowanej ropy naftowej większość krajów postanowiła uniezależnić swoje waluty od dolara i upłynnić kursy, system z Bretton Woods rozpadł się ostatecznie.

Kraje Europy Zachodniej postanowiły w tej sytuacji stworzyć własny, regionalny system stałych kursów walutowych, w którym rolę centralnej waluty rezerwowej miała pełnić marka niemiecka. Już w 1972 roku powołano do życia Europejski System Walutowy (*European Monetary System* – EMS), który składał się z trzech elementów: systemu stałych parytetów poszczególnych walut wobec marki, a za pośrednictwem marki także wobec wszystkich pozostałych walut uczestniczących w EMS, Europejskiego Funduszu Walutowego, mającego (na wzór Międzynarodowego Funduszu Walutowego) udzielać pomocy krajom notującym duże deficyty w bilansach płatniczych, oraz Europejskiej Jednostki Rozrachunkowej (*European Unit of Account* – EUA), później przemianowanej na Europejską Jednostkę Walutową (*European Currency Unit* – ECU), która miała stać się pierwowzorem przyszłej wspólnej waluty europejskiej. Kluczowym elementem był system wzajemnie ustalanych stałych kursów walut europejskich względem siebie, zwany siatką kursową (*grid*). Dążąc do zapewnienia w jego funkcjonowaniu pewnego zakresu elastyczności, przewidziano możliwość odchyleń kursów bieżących od parytetów o +/–2,25% (a w szczególnych sytuacjach o +/–6%), tworząc w ten sposób wąski „korytarz" dopuszczalnych wahań kursowych (dlatego w początkowym okresie obowiązywania ten nowy system kursowy nazywano „wężem w tunelu").

Gdy wzajemna wymienialność walut europejskich w zakresie transakcji kapitałowych była ograniczona, system funkcjonował bez większych zakłóceń. Kraje, które doświadczały systematycznego odpływu kapitału na skutek chronicznych deficytów handlowych, zmuszone zostały do dewaluacji swoich walut wobec marki niemieckiej (Francja, Włochy, Wielka Brytania), co zawsze pozostawało operacją politycznie kosztowną, lecz umożliwiało jednocześnie odzyskać równowagę

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

płatniczą (choć z reguły nie na długo). Gdy zakres wymienialności uległ poszerzeniu na transakcje kapitałowe, obrona stałych kursów wymagała zaostrzania polityki pieniężnej, a to odbijało się negatywnie na wzroście gospodarczym. Skutki obrony kursów były tym dotkliwsze, że Niemcy odmawiały symetrycznego łagodzenia polityki pieniężnej (obniżania własnych stóp procentowych), kierując się troską o utrzymanie krajowej inflacji pod kontrolą. W rezultacie system EMS stał się źródłem ciągłych napięć i wzajemnych sporów między krajami członkowskimi. W tych warunkach szybko dojrzewała myśl o utworzeniu w Europie unii walutowej.

6.1.5. Podstawowe decyzje z zakresu polityki kursowej

O ile polityka pieniężna odpowiada za wartość pieniądza krajowego w odniesieniu do nabywanych dóbr i usług, o tyle polityka kursowa decyduje o wartości pieniądza krajowego w stosunku do innych walut. Oba te wymiary wartości pieniądza krajowego – wewnętrzny i zewnętrzny – są jednak ze sobą bardzo ściśle powiązane. Obie miary nie mogą bowiem w dłuższym okresie podążać w różnych kierunkach – jeśli pieniądz traci wartość na skutek inflacji krajowej, prędzej czy później również jego kurs wobec walut obcych zacznie słabnąć, i odwrotnie. Z tego powodu, a także ze względu na to, że obie odmiany polityki odnoszą się do pieniądza krajowego, traktuje się najczęściej politykę kursową jako składową polityki pieniężnej. W niniejszej książce zdecydowano się jednak omawiać politykę kursową w odrębnym rozdziale ze względu na jej specyfikę, wynikającą z bezpośredniego związku z zagranicą, oraz ze względu na większą przejrzystość wykładu[6].

W ramach polityki kursowej władza gospodarcza rozstrzyga o trzech istotnych kwestiach. Po pierwsze, dokonuje wyboru systemu (reżimu) kursowego. W tym zakresie kraje mają do wyboru system kursu stałego (sztywnego), system kursu płynnego oraz całą gamę wariantów pośrednich między tymi skrajnymi rozwiązaniami (takich jak kursy okresowo stałe, kursy płynne kontrolowane, stałe i ruchome przedziały wahań kursowych, a także wiele innych). Należy zauważyć, że w pewnych skrajnych wariantach kraj rezygnuje w ogóle z polityki kursowej. Dzieje się tak wówczas, gdy dane państwo odstępuje od stosowania własnej waluty i decyduje się posługiwać walutą innego kraju. Ma to również miejsce w sytuacji, gdy kraj wybiera system w pełni płynnego kursu walutowego, a zarazem w żaden

[6] Taka konwencja jest zresztą powszechnie stosowana w większości podręczników do polityki gospodarczej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

sposób nie ogranicza wymienialności własnej waluty. Wówczas dane państwo nie podejmuje decyzji dotyczących kursu walutowego, o wszystkim decyduje bowiem mechanizm rynkowy.

Dlatego też drugą ważną kwestią, o jakiej musi rozstrzygnąć władza gospodarcza, jest zakres wymienialności waluty krajowej na waluty obce – czyli określenie katalogu transakcji międzynarodowych, dla których zezwala się na swobodną wymienialność walut – wraz z ustaleniem przepisów określających dostęp do walut obcych w przypadku krajowych podmiotów. Decyzja o zakresie wymienialności definiuje ramy powiazań gospodarczych i finansowych kraju z zagranicą. W tej kwestii państwo ma do wyboru całkowity brak wymienialności, jeśli z jakichś powodów chce całkowicie odizolować własną gospodarkę od rynków zewnętrznych, poprzez selektywne ograniczenie wymienialności poszczególnych rodzajów transakcji międzynarodowych lub określonych podmiotów, aż po pełną wymienialność bez żadnych ograniczeń.

Wreszcie trzecia kwestia odnosi się do ustalenia określonego poziomu kursu walutowego w sytuacji, gdy kraj decyduje się na wprowadzenie systemu kursu stałego lub jakiejś jego odmiany zakładającej pewną formę kontroli. Ogólnie rzecz biorąc, wybrany kurs walutowy powinien zapewniać równowagę między podażą i popytem na rynku walut obcych oraz względną stabilność kursu realnego w średnim okresie w warunkach pełnego wykorzystania krajowych zasobów produkcyjnych. Ponieważ poziom kursu określa bieżącą opłacalność transakcji międzynarodowych, jego wybór ma zawsze duże znaczenie ekonomiczne i polityczne. Ustalając stały poziom kursu, każdy kraj zobowiązuje się tym samym do jego obrony w sytuacji, gdy popyt na waluty obce odchyla się od podaży. Wymaga to przyjęcia i stosowania określonych zasad interwencji na rynku walutowym. W dalszych częściach rozdziału omówimy kolejno te trzy kwestie.

6.2. Wybór systemu kursowego

6.2.1. Wybór między kursem stałym a kursem płynnym

Czy można stwierdzić, że istnieje jeden optymalny system kursowy? Jeśliby tak było, to w większości krajów powinien obowiązywać ten sam, identyczny system kursowy. Tak jednak nie jest– wystarczy spojrzeć na zróżnicowaną gamę reżimów kursowych stosowanych w różnych krajach. Ta rozmaitość sugeruje, że poszczególne

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

państwa, rozstrzygając o wyborze systemu kursowego, biorą pod uwagę odmienne czynniki,. Czy jesteśmy w stanie coś o nich powiedzieć?

Wahania kursu walutowego są źródłem niepewności dla uczestników rynku, ponieważ powodują zmiany relacji cen w kraju oraz za granicą i w ten sposób wpływają na opłacalność transakcji międzynarodowych w krótkim okresie. Zmiany te stają się tym silniejsze, im głębsze są wahania kursu oraz im większą rolę w gospodarce odgrywają transakcje międzynarodowe. Dlatego też na poziomie najbardziej podstawowym wybór reżimu kursowego zależy od wielkości gospodarki i stopnia jej otwartości. Małe kraje o gospodarce otwartej, gdzie relacja obrotów handlu zagranicznego do PKB jest wysoka, znajdują istotne powody, aby stosować kursy stałe. W takich krajach kurs walutowy stanowi najważniejszą cenę jednostkową, ponieważ wpływa silnie – ze względu na duży udział obrotów międzynarodowych w produkcji i konsumpcji – na poziom pozostałych cen, a także – pośrednio – na wynagrodzenia czynników produkcji (płace, zyski). Wahania kursu przenoszą się więc w sposób bardzo odczuwalny na krajowe ceny i płace. Dlatego też kraje o wysokim udziale transakcji międzynarodowych w PKB chętnie stosują stałe kursy walutowe, ponieważ stabilizują one warunki gospodarowania i chronią gospodarkę przed szokami zewnętrznymi. Do krajów takich należą m.in. Belgia, Holandia, Luksemburg, Irlandia, Estonia, czy Czechy, w których relacja sumy eksportu oraz importu dóbr i usług w stosunku do PKB przekracza 150%.

Oprócz wielkości i stopnia otwarcia gospodarki również inne względy mogą przemawiać na rzecz stosowania kursu stałego (zob. Frankel 1999). Na przykład kraje wysoko zadłużone w walutach obcych powinny dążyć raczej do stabilizowania kursu, aby zminimalizować wpływ wahań kursowych na koszty obsługi długu ponoszone w walucie krajowej. Podobnie też kraje zwalczające wysoką inflację stosują najczęściej kursy stałe wobec głównych walut wymienialnych, wykorzystując w ten sposób kurs walutowy jako „nominalną kotwicę" hamującą wzrost cen i stabilizującą oczekiwania inflacyjne. Taką strategię stosowały wielokrotnie kraje Ameryki Łacińskiej (Argentyna, Brazylia) w walce z hiperinflacją w latach 80. i 90., a także Polska w latach 1990–1992 w trakcie realizacji programu stabilizacyjnego.

Najtwardszą i najbardziej wiarygodną odmianą systemu kursu sztywnego jest rozwiązanie nazywane urzędem walutowym (*currency board*), w ramach którego obowiązuje całkowite pokrycie krajowego obiegu pieniężnego rezerwami walut wymienialnych, a parytet wymiany jest zapisany w konstytucji lub co najmniej na poziomie ustawowym (zob. ramka 6.3). Z kolei najbardziej elastyczną formą reżimu kursowego jest kurs w pełni płynny (*free float*), gdzie kurs walutowy jest

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kształtowany wyłącznie poprzez relację podaży waluty obcej i popytu na nią, a władza gospodarcza nie interweniuje na rynku walutowym. Między tymi skrajnymi rozwiązaniami zawiera się wiele reżimów pośrednich, o różnym stopniu swobody w zakresie kształtowania kursu i różnej skali interwencji.

Ramka 6.3. Urząd walutowy (*currency board*)

Urząd walutowy stanowi odmianę specjalnego systemu pieniężno-kursowego, w ramach którego zachodzi pełna wymienialność waluty krajowej na walutę obcą, pełniącą funkcję waluty rezerwowej, po stałym kursie oraz zapewnione jest pełne pokrycie (co najmniej w 100%) podaży krajowego pieniądza rezerwowego walutą rezerwową. W celu umocnienia wiarygodności zobowiązań przyjętych w ramach urzędu walutowego jego podstawą prawną jest z reguły zapis konstytucyjny lub ustawowy. Ogranicza to lub praktycznie uniemożliwia manipulowanie przy systemie kursowym przez rząd lub bank centralny. Trzy filary systemu *currency board* – stały kurs walutowy, swobodna międzynarodowa wymienialność waluty krajowej w warunkach pełnego pokrycia obiegu pieniężnego rezerwami walutowymi oraz bardzo mocne zobowiązanie prawne do utrzymania kursu – zapewniają wiarygodność i długookresową stabilność kursu waluty.

System typu *currency board* jest stosowany z reguły przez małe kraje o otwartych gospodarkach, chcące zwiększyć wiarygodność stałego kursu waluty. Według danych Międzynarodowego Funduszu Walutowego, na koniec 2014 roku 11 państw na świecie stosowało system urzędu walutowego, w tym dwa kraje europejskie – Bułgaria oraz Bośnia i Hercegowina[7].

Hong Kong to przykład kraju, w którym urząd walutowy funkcjonował bez zakłóceń przez wiele dziesięcioleci, będąc – w opinii ekonomistów – ważnym czynnikiem kształtowania konkurencyjnej, otwartej gospodarki. Z kolei krajem, w którym miało miejsce najbardziej spektakularne załamanie systemu urzędu walutowego, jest Argentyna (kryzys walutowy w 2001 roku).

W ramach systemu *currency board* kompetencje banku centralnego są bardzo ograniczone. Jego zadaniem jest przede wszystkim dostarczać walutę krajową lub walutę rezerwową na każde żądanie uczestników rynku. W skrajnym wariancie

[7] Pozostałe kraje to Hongkong SAR, Antigua i Barbuda, Djibouti, Dominica, Grenada, St. Kitts i Nevis, St. Lucia, St. Vincent i the Grenadines oraz Brunei Darussalam.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

bank centralny nie ustala nawet stóp procentowych, które kształtują się swobodnie na rynku międzybankowym. Urząd walutowy pozostaje najbardziej wiarygodnym, stabilnym i zarazem najsztywniejszym systemem kursu stałego w sytuacji, gdy dany kraj chce mimo wszystko nadal zachowywać własną walutę (Hanke i Schuler 1994, Hanke 2002).

Posługiwanie się stałym kursem walutowym jako instrumentem zapewniającym stabilne warunki gospodarowania niesie jednak także pewne zagrożenia. Przede wszystkim usztywnienie kursu nominalnego blokuje działanie ważnego mechanizmu dostosowawczego, jakim są zmiany kursu elastycznie reagującego na zmieniającą się sytuację w bilansie płatniczym i w sektorze realnym. Brak tego mechanizmu sprawia, że gospodarka dotknięta szokiem podażowym lub popytowym nie powraca automatycznie do równowagi. W rezultacie kurs realny odchyla się od kursu nominalnego, co może powodować kumulowanie się napięć, a w dłuższym okresie prowadzi ostatecznie nawet do kryzysu walutowego. Oznacza to, że kraje stosujące stałe kursy płacą za uzyskaną stabilność warunków działania określoną cenę w postaci zwiększonego ryzyka nierównowagi makroekonomicznej. Dlatego też wiele państw nie decyduje się na takie rozwiązanie, wybierając w to miejsce kursy elastyczne. Dotyczy to głownie dużych krajów o rozległym rynku wewnętrznym, gdzie udział eksportu i importu w PKB jest znacznie mniejszy, w związku z czym wahania kursu nie mają tak silnego wpływu na całą gospodarkę. Do krajów takich należą m.in. Stany Zjednoczone, Japonia, Kanada czy Australia, a także Wielka Brytania i strefa euro jako całość, w przypadku których relacja sumy eksportu oraz importu dóbr i usług do PKB kształtuje się na poziomie wyraźnie poniżej 100%.

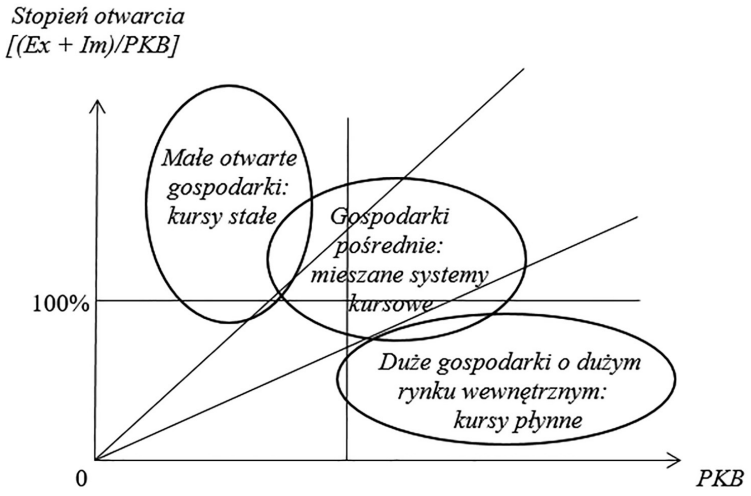
Rysunek 6.1 ukazuje ogólną postać zależności reżimu kursowego od wielkości gospodarki i stopnia jej otwarcia. Powierzchnia wykresu podzielona jest w przybliżeniu na trzy obszary. Na jednym krańcu znajdują się małe, otwarte gospodarki, na drugim zaś umieszczono gospodarki duże, o niewielkim stopniu otwarcia, a pomiędzy nimi zawierają się gospodarki średniej wielkości. Ogólna prawidłowość polega na tym, że im mniejsza jest gospodarka oraz im większy stopień otwarcia w relacji do PKB, tym większe staje się uzasadnienie dla zastosowania kursu stałego, i odwrotnie.

Z kolei wykres 6.4 przedstawia reżimy kursowe obowiązujące zarówno w krajach UE posiadających własne waluty, jak i w pozostałych najważniejszych gospodarkach świata. Jak widać, duże kraje o rozległym rynku wewnętrznym – takie jak

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

USA, Kanada, Australia, Japonia, Wielka Brytania, a także strefa euro – stosują kursy płynne. Z kolei małe kraje o otwartej gospodarce w większym stopniu starają się ograniczać wahania kursów, stosując albo kursy stałe (Bułgaria, Litwa, Dania), albo kursy zmienne kontrolowane (Węgry, Rumunia, Chorwacja). Te obserwacje w dużym stopniu pokrywają się ze wskazaniami teoretycznymi – kraje o małych otwartych gospodarkach unikają w pełni płynnego kursu (odstępstwem od tej reguły są Czechy, gdzie stosuje się kurs płynny, choć gospodarka czeska jako „mała otwarta gospodarka" bardziej „pasuje" do systemu kursu stałego)[8]. Z kolei Polska i Szwecja reprezentują typ pośredni, gdzie oba skrajne rozwiązania teoretycznie znalazłyby zastosowanie.

Rysunek 6.1. Wielkość i stopień otwarcia gospodarki a system kursowy



Źródło: opracowanie własne.

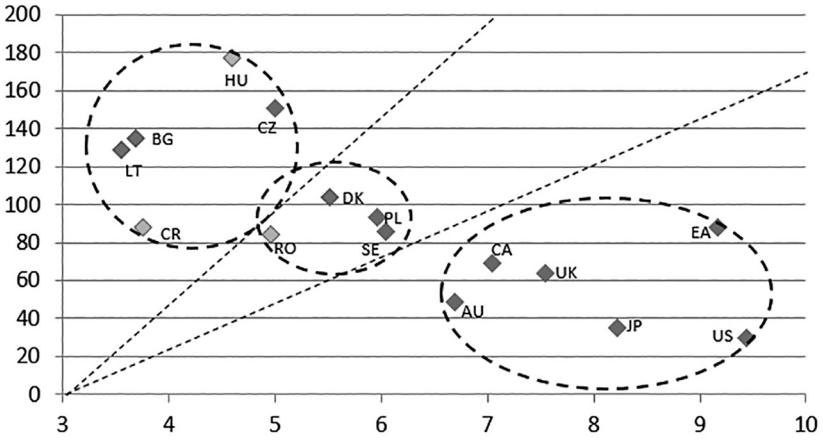
Kurs stały eliminuje ryzyko kursowe w transakcjach z opóźnionymi płatnościami i w transakcjach terminowych. Zapewnia w ten sposób stabilność warunków działania i pozwala bezpiecznie planować przedsięwzięcia gospodarcze. Lecz to, co pozostaje największą zaletą kursu stałego, jest równocześnie jego istotną wadą. Brak ryzyka kursowego zachęca do spekulacji, zwłaszcza wówczas, gdy występują różnice w poziomie stóp procentowych miedzy krajem a zagranicą. Inwestorzy mogą wtedy bezpiecznie inwestować w papiery o wyższym oprocentowaniu, nie

[8] Od jesieni 2014 roku w Czechach obowiązuje kurs płynny kontrolowany.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

obawiając się zmiany kursu. Sztywny kurs oznacza, w kategoriach równania parytetu stóp procentowych (zob. rozdział 5, równanie 5.72), że $de/e = 0$. Jeśli zarazem $i \neq i^*$, to będziemy mieli do czynienia z masowym napływem lub odpływem kapitału, co może silnie ograniczać autonomię polityki pieniężnej w danym kraju (zob. dalej). Duże przepływy kapitału w ramach systemu stałego kursu wymagają utrzymywania dużych rezerw walutowych, aby móc skutecznie interweniować na rynku walutowym i bronić ustalonego poziomu kursu. Lecz utrzymywanie dużych rezerw jest kosztowne, a co ważniejsze, nawet największe rezerwy nie są w stanie w pełni uchronić państwa przed ryzykiem kryzysu walutowego w sytuacji masowego odpływu kapitału. Zagrożenia te nasiliły się po 1990 roku, kiedy nastąpił wzrost mobilności kapitałów w skali międzynarodowej, związany z usuwaniem barier w zakresie wymienialności walut i otwieraniem rachunków kapitałowych w poszczególnych krajach. Dlatego też w latach 90. wiele państw zrezygnowało z systemu stałego kursu walutowego.

Wykres 6.4. Reżimy kursowe w zależności od wielkości gospodarki (oś pozioma – logarytm naturalny z PKB, dane w mld EUR) i relacji sumy eksportu oraz importu do PKB w wybranych krajach (oś pionowa) w 2013 roku



Uwaga: czarnym kolorem zaznaczono kraje z kursem płynnym, ciemnoszarym – kraje z kursem sztywnym, a jasnoszary kolor oznacza kraje z systemem kursu zmiennego kontrolowanego (*managed float*).

Źródło: opracowanie własne na podstawie danych Eurostatu i OECD.

Kurs płynny jest lustrzanym odbiciem kursu stałego, w związku z czym wady kursu stałego stają się zaletami kursu płynnego, i odwrotnie. Choć płynny kurs zwiększa niepewność dotyczącą przyszłości, to stanowi zarazem skuteczny mechanizm

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

dostosowawczy, przywracający naruszoną równowagę. Dzięki temu w systemie kursu płynnego ryzyko kryzysu walutowego radykalnie maleje. Wprawdzie wahania kursu zwiększają ryzyko kursowe, co przekłada się na wzrost rynkowych stóp procentowych, lecz wzrost ryzyka kursowego ogranicza zarazem skłonność inwestorów do podejmowania prób spekulowania kursem.

Jednocześnie w systemie kursu płynnego polityka pieniężna staje się bardziej autonomiczna i skuteczniejsza niż w systemie kursu stałego. Gdy bank centralny obniża stopy procentowe w celu pobudzenia gospodarki, rosną wydatki krajowe, w tym te przeznaczane na import, co pogarsza bilans obrotów bieżących. Spadek stóp procentowych powoduje zarazem odpływ kapitału za granicę, co pogarsza saldo bilansu obrotów finansowych. W ramach systemu kursu stałego skutkiem obu tych zjawisk jest odpływ oficjalnych rezerw walutowych za granicę, co zmniejsza jednak podaż pieniądza krajowego i pociąga za sobą wzrost krajowych stóp procentowych. Polityka pieniężna banku centralnego zmierzająca do obniżenia stóp okazuje się więc nieskuteczna. W systemie kursu płynnego reakcja gospodarki staje się zasadniczo odmienna. Obniżka stóp procentowych przez bank centralny również przekłada się na wzrost importu i odpływ kapitału, lecz skutkiem tego jest osłabienie kursu waluty krajowej, a nie odpływ oficjalnych rezerw procentowych. W efekcie podaż pieniądza krajowego nie spada, a deprecjacja kursu wywołuje wzrost eksportu i przyspiesza rozwój gospodarczy. Polityka pieniężna pozostaje w tej sytuacji skuteczna. Powrócimy do tego zagadnienia w dalszej części rozdziału.

Wcześniej zaś w rozdziale 5 omawialiśmy zjawisko „niemożliwego trójkąta" i pokazaliśmy, że w warunkach pełnej swobody transgranicznych przepływu kapitałów bank centralny nie może jednocześnie ustalać celu inflacyjnego i kursu walutowego. Oznacza to, że jeśli kraj stosuje strategię bezpośredniego celu inflacyjnego, to nie jest w stanie kontrolować zarazem kursu walutowego. Jedynym systemem kursowym w pełni spójnym ze strategią BCI jest więc system kursu płynnego. Wobec tego państwa, które chcą kontrolować poziom kursu, muszą zrezygnować ze skutecznej polityki kontrolowania inflacji.

6.2.2. Rozwiązania pośrednie

Jak już wspomniano, między rozwiązaniami skrajnymi – kursem sztywnym i kursem płynnym – istnieje wiele rozwiązań pośrednich, które zapewniają dowolny stopień wymienności między stabilnością oraz elastycznością systemu kursowego w zależności od indywidualnych potrzeb danego kraju. W tabeli 6.2 przedstawiono

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

klasyfikację różnych odmian systemu kursowego stosowaną przez Międzynarodowy Fundusz Walutowy (IMF 2007, IMF 2014).

Tabela 6.2. Klasyfikacja różnych odmian systemu kursu walutowego

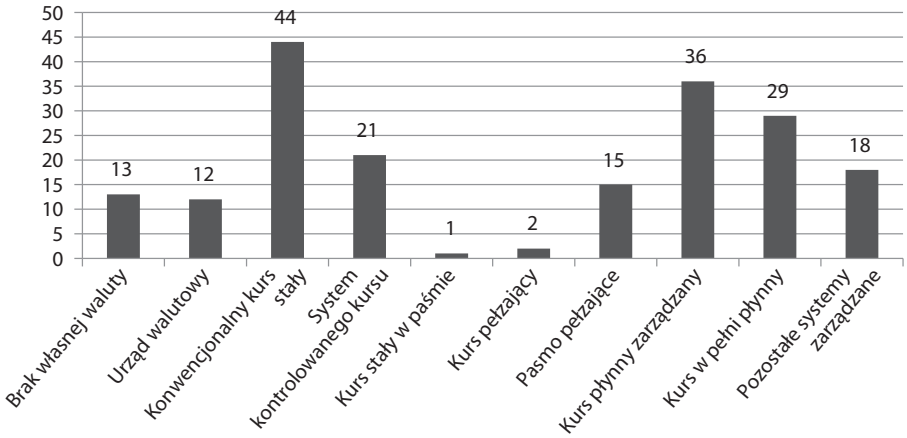
Kategoria ogólna	System kursowy	Odmiana systemu	Charakterystyka
Kurs sztywny (*hard peg*)	brak własnego środka płatniczego (*no separate legal tender*)		Waluta innego kraju jest jedynym legalnym środkiem płatniczym.
	zarząd walutą (*currency board*)		Pełne pokrycie obiegu pieniężnego rezerwami walut obcych i zobowiązanie konstytucyjne lub ustawowe do wymiany waluty krajowej na walutę obca po ustalonym kursie.
Kurs stały (*soft peg*)	konwencjonalny kurs stały (*conventional fixed peg*)		Kurs waluty krajowej pozostaje przez co najmniej trzy miesiące w paśmie +/–1% wobec waluty zagranicznej.
	kurs stały elastyczny (*intermediate peg*)	kurs stały w paśmie poziomym (*peg within horizontal bands*)	Kurs waluty krajowej fluktuuje w paśmie o szerokości większej niż +/–1%.
		kurs stały pełzający (*crawling peg*)	Kurs waluty krajowej ulega systematycznej dewaluacji z określoną częstotliwością i w określonym tempie w paśmie +/–1%.
		pełzające pasmo kursowe (*crawling band*)	Pasmo kursowe o szerokości większej niż +/–1% ulega systematycznej dewaluacji z określoną częstotliwością i w określonym tempie.
Kurs płynny (*floating*)	kurs płynny zarządzany (*managed*), bez ustalonej ścieżki zmian i bez określonego celu kursowego (*no pre-determinded path, no pre-determined target rate*)		Władza pieniężna interweniuje na dowolną skalę i z dowolną częstotliwością na rynku walutowym.
Kurs w pełni płynny (*independent float*)	kurs płynny rynkowy (*market-determined*) z ograniczonymi interwencjami		Kurs kształtuje się na rynku walutowym pod wpływem zmian podaży i popytu na walutę obcą. Ograniczone, okazjonalne są zaś interwencje władzy pieniężnej na rynku walutowym.
Inne	inne systemy pośrednie nie będące systemem ani kursu sztywnego, ani kursu płynnego		

Źródło: opracowanie własne na podstawie danych MFW.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 6.5, sporządzony na podstawie danych MFW, przedstawia rozkład różnych reżimów kursowych stosowanych przez poszczególne kraje w 2014 roku. Jak widać, kraje wybierają zasadniczo jedno z dwóch głównych rozwiązań – system kursu stałego lub system kursu płynnego, rzadziej stosują natomiast rozwiązania pośrednie (*intermediate regimes*). W 2014 roku na ogólną liczbę 191 krajów objętych oceną MFW tylko 13 spośród nich nie posiadało własnej waluty. Z kolei na 178 krajów posiadających własną walutę, w 77 obowiązywała jakaś odmiana kursu sztywnego, a 65 państw stosowało kurs zasadniczo płynny lub w pełni płynny. Ten spolaryzowany rozkład może sugerować, że próby łączenia obu celów – stabilności i elastyczności – w ramach mieszanego systemu kursowego wiążą się ze zwiększonym ryzykiem zaprzepaszczenia ich realizacji w zadawalającym stopniu. Wniosek ten potwierdziły badania wielu autorów (zob. np. Obstfeld i Rogoff 1995, Frankel 1999, Rogoff i in. 2005). Całkowita rezygnacja z własnej waluty i przyjęcie waluty innego kraju – bez korzyści wynikających z uczestnictwa w unii walutowej – okazuje się zarazem rozwiązaniem atrakcyjnym tylko dla nielicznej grupy bardzo małych gospodarek.

Wykres 6.5. Reżimy kursowe na świecie liczba stosujących je krajów w 2014 roku



Uwaga: państwa członkowskie strefy euro należą w tej klasyfikacji do kategorii „kursu płynnego".
Źródło: IMF 2014.

Należy zauważyć, że w okresie ostatnich trzech dziesięcioleci ubiegłego wieku doszło do zasadniczej zmiany w układzie systemów kursowych na świecie. Wprawdzie dane dotyczące poszczególnych lat nie są w pełni porównywalne ze względu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

na częste zmiany w klasyfikacji dokonywane przez MFW, lecz z całą pewnością można stwierdzić, że mieliśmy w tym czasie do czynienia z masowym odchodzeniem od systemu kursu stałego. Jak podają Calvo i Reinhart (2000), udział krajów z jakąś odmianą systemu kursu stałego (konwencjonalnego i zarządzanego) spadł z 97% w 1970 roku do 54% w 1999 roku na rzecz wzrostu znaczenia systemu kursu płynnego i w pełni płynnego. Trend ten był spowodowany rosnącym ryzykiem wystąpienia kryzysu walutowego w warunkach zwiększonej swobody przepływów kapitału w ramach obowiązywania systemu kursu stałego. Jednocześnie zmniejszył się udział krajów stosujących pośrednie systemy kursowe (Fischer 2001, Eichengreen i Razo-Garcia 2006).

Calvo i Reinhart (2000) zauważyli, że dane MFW mogą być niekiedy mylące. Zaintrygowało ich to, że w przypadku wielu krajów deklarujących stosowanie płynnych kursów walutowych mamy w rzeczywistości do czynienia z kursami w mniejszym lub większym zakresie kontrolowanymi. Calvo i Reinhart nazwali to zjawisko „obawą przed upłynnieniem" (*fear of floating*). Niechęć tych państw do pełnego upłynnienia kursu oraz skłonność do częstego interweniowania na rynku walutowym wynika z zasadniczo niskiej wiarygodności stosowanej w tych krajach polityki i związanego z tym potencjalnego ryzyka dużych wahań kursu. W takiej sytuacji wiele państw, zwłaszcza rozwijających się, interweniuje na rynku walutowym w celu zmniejszenia wahań i ustabilizowania kursu, mimo że formalnie obowiązują tam płynne kursy walutowe. Wynika stąd, że w rzeczywistości udział krajów stosujących płynne kursy walutowe jest mniejszy niż ten odnotowany przez MFW.

W kolejnych latach tendencja do wzrostu udziału systemu kursu płynnego kosztem systemu kursu stałego uległa zahamowaniu, a nawet częściowemu odwróceniu. W krajach uprzemysłowionych, gdzie rynki finansowe są „głębokie" i dobrze rozwinięte, kursy walut w ramach systemu kursów płynnych nie podlegały dużym wahaniom (zob. IMF 2007). Lecz w wielu innych państwach, zwłaszcza w przypadku gospodarek „wschodzących", korzystających w znacznej mierze z zewnętrznego finansowania rozwoju, płynne kursy stały się źródłem nadmiernej zmienności i niestabilności. Wobec tego w pierwszej dekadzie XXI wieku szereg krajów zdecydowało się na szeroki zakres interwencji na rynkach walutowych, próbując ograniczyć w ten sposób skalę wahań kursowych w ramach systemu nominalnie płynnego kursu. Szczególnie w okresie ostatniego kryzysu finansowego i bezpośrednio po nim mieliśmy do czynienia z tendencją do rezygnacji z kursów płynnych na rzecz kursów stałych elastycznie zarządzanych i na rzecz reżimów pośrednich. Pewne odmiany kursów płynnych „zarządzanych" mogły się stać bardziej atrakcyjne

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w okresie kryzysu, ponieważ zapewniają one bankom centralnym możliwość skuteczniejszego reagowania na szoki, bez konieczności obrony jakiegoś ustalonego poziomu kursu. Nowsze badania, uwzględniające doświadczenia ostatniego kryzysu finansowego, pokazują, że rozwiązania skrajne (*corner solutions*) nie gwarantują dostatecznej stabilności w warunkach kryzysowych, zwłaszcza w przypadku gospodarek „wschodzących". Z kolei systemy „zarządzane", zapewniające zarówno szeroki zakres płynności kursu, jak i możliwość interwencji (*managed floats*), gwarantują wystarczająco wysoki stopień elastyczności, a zarazem znacząco obniżają ryzyko wystąpienia kryzysu walutowego. Wskazane badania potwierdzają też jednocześnie, że kursy stałe są zasadniczo bardziej ryzykowne i częściej towarzyszą im epizodyczne kryzysy oraz załamania wzrostu gospodarczego (Ghosh, Ostry i Qureshi 2014).

6.3. Wpływ systemu kursowego na skuteczność polityki pieniężnej i fiskalnej

6.3.1. System kursowy a skuteczność polityki pieniężnej

Polityka makroekonomiczna służy utrzymaniu gospodarki na ścieżce zrównoważonego wzrostu, zapewniającej pełne zatrudnienie i stabilność cen. Jej głównym zadaniem jest więc przywracanie równowagi – czyli stabilizowanie gospodarki – w reakcji na cykliczne wahania koniunktury i jednorazowe szoki zewnętrzne. Poniżej rozpatrzymy skuteczność polityki makroekonomicznej w przywracaniu równowagi w warunkach różnych systemów kursowych, przy czym zajmiemy się głównie polityką *ekspansywną*, a więc taką, której celem jest wywoływanie wzrostu dochodu, produkcji i zatrudnienia[9]. W tym kontekście mówimy o skuteczności polityki makroekonomicznej lub jej braku. Stwierdzimy więc, że polityka makroekonomiczna okazuje się skuteczna wówczas, gdy jest w stanie spowodować pożądany wzrost produkcji i zatrudnienia. W przeciwnym razie uznamy, że polityka makroekonomiczna pozostaje nieskuteczna.

[9] Odchylenia od poziomu równowagi nie są przez władzę gospodarczą traktowane całkiem symetrycznie. Recesja gospodarcza jest z reguły uznawana za zjawisko groźniejsze niż „przegrzanie" gospodarki, ponieważ pociąga za sobą wzrost bezrobocia oraz istotne pogorszenie warunków życia społeczeństwa. Dlatego też politycy zawsze starają się, aby mieć do dyspozycji takie instrumenty polityki, które w obliczu recesji lub negatywnego szoku zewnętrznego zapewnią wzrost dochodu, produkcji i zatrudnienia.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W niniejszym podrozdziale zostanie również wykazane, że skuteczność polityki makroekonomicznej – zarówno pieniężnej, jak i fiskalnej – w znacznej mierze zależy od obowiązującego w danym kraju systemu kursowego. W tym celu najpierw omówimy skuteczność polityki pieniężnej w warunkach stałego i płynnego kursu waluty, a w następnym punkcie omówimy skuteczność polityki fiskalnej.

Najkrócej rzecz ujmując, skuteczność polityki pieniężnej w systemie kursu stałego jest niewielka lub żadna, podczas gdy w systemie kursu płynnego skuteczność pozostaje wysoka. Ponadto, skuteczność polityki pieniężnej zależy też od stopnia transgranicznej mobilności kapitału. Im wyższy jest stopień mobilności kapitału,, tym mniejsza stanie się skuteczność polityki pieniężnej w systemie kursu stałego, i tym większa stanie się jej skuteczność w systemie kursu płynnego.

Najpierw przeanalizujemy skuteczność polityki pieniężnej w systemie **kursu stałego**. Załóżmy, że gospodarka znalazła się w recesji i bank centralny podejmuje starania, aby doprowadzić do wzrostu produkcji i zatrudnienia. W tym celu obniża oficjalne stopy procentowe i zwiększa wielkość bazy monetarnej (rośnie podaż pieniądza rezerwowego), co powoduje spadek rynkowych stóp procentowych. Wówczas nastąpi wprawdzie wzrost inwestycji, lecz będzie on miał charakter przejściowy i po krótkim czasie inwestycje powrócą do poziomu początkowego. Aby zrozumieć, dlaczego tak się dzieje, należy przyjrzeć się zmianom w zakresie bilansu płatniczego. Po pierwsze, obniżenie stóp procentowych przekłada się na pogorszenie rachunku bieżącego bilansu płatniczego, ponieważ wraz ze wzrostem wydatków krajowych rosną także wydatki na import, co przy stałym poziomie eksportu (decyduje o nim poziom popytu zagranicznego, który nie zależy od zmian krajowej stopy procentowej) powoduje pogorszenie salda obrotów bieżących. Po drugie, pogarsza się również saldo rachunku finansowego bilansu płatniczego, ponieważ obniżenie stóp procentowych wywołuje odpływ kapitału za granicę. Pod wpływem działania obu tych mechanizmów rośnie popyt na waluty obce, co wywołuje presję na osłabienie kursu waluty krajowej. Lecz ze względu na to, że obowiązuje system stałego kursu walutowego, bank centralny musi bronić ustalonego poziomu kursu. Aby zaspokoić zwiększony popyt na waluty obce, interweniuje więc na rynku walutowym, sprzedając waluty z własnych rezerw w zamian za walutę krajową. Interwencje banku centralnego prowadzą jednak do spadku rezerw walutowych i ograniczenia podaży pieniądza krajowego, co z kolei powoduje wzrost rynkowych stóp procentowych i spadek inwestycji. W systemie stałego kursu bank centralny nie jest więc w stanie doprowadzić do trwałego wzrostu podaży pieniądza oraz trwałego wzrostu produkcji i zatrudnienia. W tej sytuacji zmiany w zakresie bazy monetarnej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

nie są wynikiem autonomicznych decyzji banku centralnego, lecz mają charakter endogeniczny. W systemie kursu stałego polityka pieniężna okazuje się więc zasadniczo nieskuteczna.

To, czy w ogóle wystąpi jakikolwiek krótkookresowy wzrost inwestycji, zależy od stopnia transgranicznej mobilności kapitału. Gdy istnieją ograniczenia w przepływie kapitału, wynikające na przykład z częściowej lub całkowitej niewymienialności waluty krajowej, pogorszenie bilansu płatniczego nastąpi z opóźnieniem i będzie stosunkowo niewielkie, proporcjonalne do wzrostu importu. W warunkach braku swobody przepływu kapitału stopień kontroli nad podażą pieniądza jest większy, a polityka pieniężna staje się w pewnym ograniczonym zakresie skuteczna. Z kolei, gdy brakuje barier w przepływie kapitału i jest on bardzo elastyczny względem stóp procentowych, pogorszenie bilansu płatniczego będzie znaczne i wystąpi natychmiast. W takiej sytuacji obniżenie stóp procentowych przez bank centralny nie przyniesie żadnych efektów. Dlatego też w warunkach stałego kursu i przy pełnej swobodzie przepływów kapitału polityka pieniężna pozostaje całkowicie nieskuteczna.

W systemie **kursu płynnego** gospodarka reaguje inaczej. Gdy bank centralny obniża oficjalne stopy procentowe i zwiększa podaż pieniądza rezerwowego, rynkowe stopy procentowe spadają, co wywołuje początkowy wzrost inwestycji, podobnie jak miało to miejsce w systemie kursu stałego. Tam miało jednak również miejsce osłabienie kursu waluty krajowej – na skutek pogorszenia zarówno rachunku bieżącego, jak i rachunku finansowego.

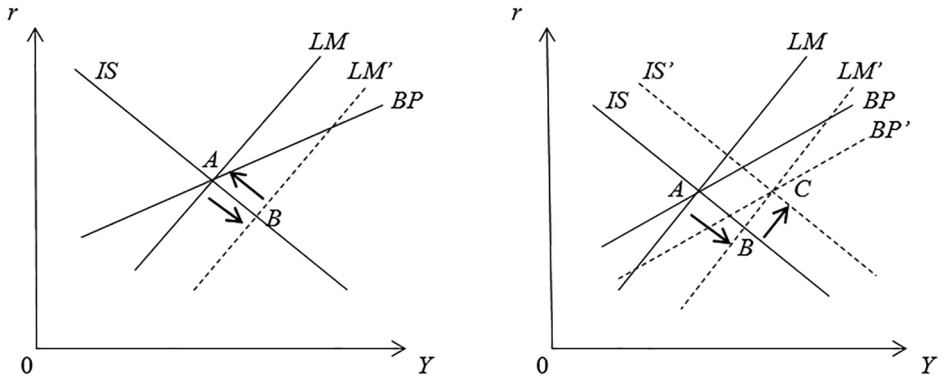
Przy niezmienionej podaży walut obcych równowaga na rynku walutowym jest przywracana poprzez deprecjację kursu, a rezerwy walutowe banku centralnego pozostają stałe. Efekt początkowego spadku stóp procentowych zostaje spotęgowany efektem deprecjacji kursu, co wywołuje dodatkowy wzrost produkcji, dochodu i zatrudnienia (rośnie eksport, spada import, wzrasta popyt zagraniczny, a popyt krajowy przesuwa się na produkcję krajową – tzw. *demand shift*). Efekt ten będzie tym silniejszy, im bardziej wrażliwe na zmiany stóp okażą się przepływy kapitału (mobilność), ponieważ silniejsza będzie wówczas deprecjacja kursu. W warunkach ograniczeń lub braku mobilności kapitału deprecjacja kursu staje się odpowiednio mniejsza, ponieważ wpływa na nią wyłącznie efekt pogorszenia rachunku bieżącego.

Skutki ekspansywnej polityki pieniężnej w warunkach stałego i płynnego kursu walutowego wygodnie jest przedstawić graficznie za pomocą modelu Mundella-Fleminga (zob. rozdział 2). Lewa część rysunku 6.2 ilustruje skutki ekspansji pieniężnej w systemie stałego kursu. W punkcie wyjścia gospodarka znajduje

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

się w miejscu przecięcia krzywej *LM*, krzywej *IS* oraz krzywej równowagi bilansu płatniczego *BP* (punkt *A*). Nachylenie krzywej *BP* zależy od stopnia elastyczności przepływów kapitału względem zmian stopy procentowej – im bardziej mobilny okaże się kapitał, tym bardziej płaska stanie się krzywa *BP*. Wzrost bazy monetarnej i podaży pieniądza przesuwa krzywą *LM* w prawo (następuje spadek rynkowych stóp procentowych), a to prowadzi do wzrostu inwestycji oraz zwiększenia poziomu produkcji z punktu *A* do *B*. Jednocześnie pojawia się deficyt w bilansie płatniczym i występuje presja na osłabienie kursu waluty. W warunkach stałego kursu zmusza to bank centralny do sprzedaży walut z rezerw, co pociąga za sobą spadek bazy monetarnej. Krzywa *LM* przesuwa się z powrotem do początkowego położenia i gospodarka powraca do punktu *A*. Prowadzona polityka pieniężna jest zasadniczo nieskuteczna. Skala i tempo początkowego przesunięcia z *A* do *B* zależą od stopnia mobilności kapitału. Przy wysokiej mobilności skala przesunięcia pozostanie minimalna i bank centralny nie będzie w stanie w ogóle zwiększyć podaży pieniądza. Polityka pieniężna okaże się wówczas całkowicie nieskuteczna.

Rysunek 6.2. Skuteczność polityki pieniężnej przy sztywnym i płynnym kursie walutowym w modelu Mundella-Fleminga



Źródło: opracowanie własne.

Prawa część rysunku 6.2 ilustruje skutki ekspansji pieniężnej w systemie płynnego kursu. W punkcie wyjścia gospodarka znajduje się w równowadze (punkt *A*). Wzrost bazy monetarnej i podaży pieniądza przesuwa krzywą *LM* w prawo, co skutkuje wzrostem inwestycji oraz produkcji z punktu *A* do punktu *B*. Jednocześnie pojawia się deficyt w bilansie płatniczym i kurs walutowy ulega deprecjacji. W rezultacie wzrasta eksport, maleje import i rośnie produkcja krajowa. Deprecjacja

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

oraz spowodowany nią wzrost popytu zagranicznego w połączeniu z przesunięciem popytu krajowego na dobra krajowe wpływa na przesunięcie krzywych *IS* i *BP* w prawo, podczas gdy gospodarka przesuwa się do nowego punktu równowagi *C*. Ekspansja pieniężna powoduje więc trwały wzrost produkcji i dochodu *Y*. Zauważmy, że wielkość przyrostu produkcji – czyli odległość od punktu *A* do *C* – wynika ze skali deprecjacji waluty krajowej, a ta z kolei zależy od stopnia mobilności przepływów kapitału. Im większa okaże się mobilność kapitału, tym bardziej płaska stanie się krzywa *BP*, a jednocześnie tym głębsza będzie deprecjacja i tym silniejszy nastąpi efekt ekspansji pieniężnej na produkcję i dochód. W tych warunkach polityka pieniężna pozostaje bardzo skuteczna. Tabela 6.3. podsumowuje warunki skuteczności polityki pieniężnej w ramach różnych reżimów kursowych i przy różnym poziomie mobilności kapitału.

Tabela 6.3. Skuteczność polityki pieniężnej przy sztywnym i płynnym kursie walutowym

Kurs sztywny		Kurs płynny	
Pełna mobilność kapitału	Brak mobilności kapitału	Pełna mobilność kapitału	Brak mobilności kapitału
Brak autonomii w polityce pieniężnej, podaż pieniądza endogeniczna, polityka pieniężna jest nieskuteczna.	Częściowa autonomia w polityce pieniężnej, podaż pieniądza egzogeniczna, polityka pieniężna oddziałuje tylko na rachunek bieżący, ograniczona i przejściowa skuteczność polityki pieniężnej.	Autonomia w polityce pieniężnej, podaż pieniądza egzogeniczna, polityka pieniężna staje się super-skuteczna poprzez deprecjację kursu.	Autonomia w polityce pieniężnej, podaż pieniądza egzogeniczna, deprecjacja ograniczona tylko poprzez rachunek bieżący, polityka pieniężna pozostaje umiarkowanie skuteczna.

Źródło: na podstawie Acocella (2005, s. 274–277).

Należy zaznaczyć, że analiza skuteczności polityki pieniężnej w ramach różnych systemów kursowych prowadzona była przy założeniu, że bank centralny nie podejmuje żadnych działań mających na celu niwelowanie pieniężnych skutków napływu i odpływu kapitału. Tymczasem bank centralny, gdy zmuszony jest sprzedawać (lub kupować) waluty obce w celu obrony poziomu kursu, ma możliwość prowadzenia równolegle operacji *repo* (lub *reverse repo*). Dzięki temu może utrzymywać podaż pieniądza rezerwowego na niezmienionym poziomie, mimo zmian w zakresie oficjalnych aktywów rezerwowych. Mechanizm sterylizacji oraz interwencji sterylizowanych zostanie omówiony w dalszej części rozdziału.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

6.3.2. System kursowy a skuteczność polityki fiskalnej

Obecnie omówimy skuteczność ekspansywnej polityki fiskalnej przy pobudzaniu produkcji i zatrudnienia w warunkach różnych systemów kursowych. W systemie **kursu stałego** wzrost wydatków rządowych, który nie jest finansowany dodrukiem pieniądza przez bank centralny (baza monetarna pozostaje niezmieniona), pociąga za sobą wzrost rynkowych stóp procentowych, ponieważ wzrostowi popytu transakcyjnego na pieniądz z tytułu dodatkowych wydatków rządowych nie towarzyszy wzrost podaży pieniądza[10]. Skutkiem wzrostu stóp procentowych będzie spadek inwestycji[11]. Gdy spadek inwestycji okaże się mniejszy niż początkowy wzrost wydatków rządowych, całkowity popyt zagregowany w gospodarce zwiększy się, lecz kiedy spadek inwestycji stanie się większy niż początkowy wzrost wydatków rządowych, popyt zagregowany spadnie. W tej sytuacji nie możemy *ex ante* stwierdzić, że wzrost wydatków rządowych spowoduje *per saldo* wzrost produkcji i zatrudnienia. Rezultat netto w odniesieniu do całej gospodarki zależy od tego, czy przeważy (pozytywny) efekt wzrostu wydatków rządowych, czy (negatywny) efekt wzrostu stóp procentowych. Ogólnie rzecz biorąc, im niższa jest wrażliwość inwestycji na wzrost stóp procentowych, tym skuteczniejsza będzie polityka fiskalna.

Lecz ekspansywna polityka fiskalna wywiera również dwojaki wpływ na bilans płatniczy. Z jednej strony wzrost wydatków rządowych oznacza bowiem także zwiększenie wydatków na import, co – przy danym poziomie eksportu – wywołuje pogorszenie salda obrotów bieżących i presję na osłabienie kursu waluty krajowej. Jest to efekt handlowy. Z drugiej zaś strony wzrost stóp procentowych powoduje napływ kapitału zagranicznego, poprawę salda obrotów finansowych i presję na umocnienie kursu waluty krajowej. Jest to efekt kapitałowy. Efekt netto w zakresie bilansu płatniczego będzie zależał od tego, która z tych dwóch zmian okaże się silniejsza. Gdy transgraniczna mobilność kapitału jest wysoka, efekt kapitałowy stanie się silniejszy od efektu handlowego i pojawi się presja na umocnienie kursu waluty krajowej. Bank centralny będzie wówczas zmuszony bronić stałego kursu i zacznie skupować waluty obce w zamian za walutę krajową, co spowoduje wzrost

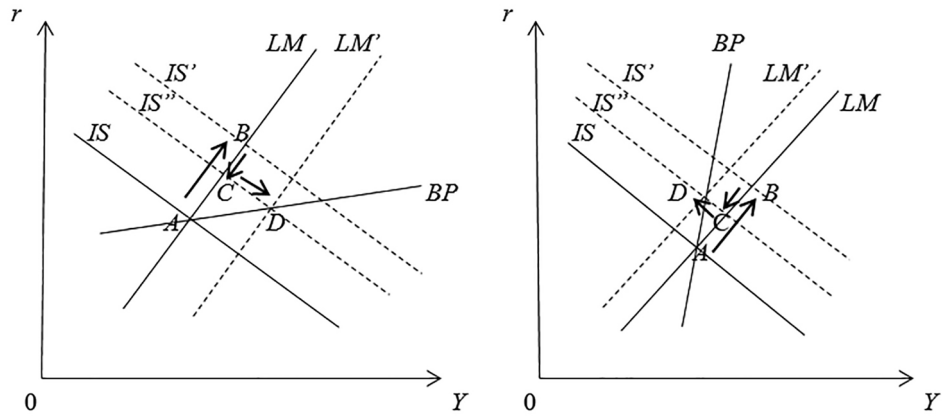
[10] Również bank centralny może podnieść oficjalne stopy procentowe, jeśli uzna, że wzrost wydatków rządowych zagraża realizacji celu inflacyjnego.

[11] Skutkiem wzrostu stóp procentowych będzie także spadek konsumpcji spowodowany wzrostem kosztów kredytu i zwiększoną atrakcyjnością lokat bankowych. W dalszej analizie efekt konsumpcyjny pomijamy.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

bazy monetarnej i spadek stóp procentowych. W tej sytuacji zmiany w zakresie bilansu płatniczego wzmacniają skuteczność polityki fiskalnej. Gdy mobilność kapitału pozostaje jednak ograniczona, efekt handlowy staje się silniejszy od efektu kapitałowego i pojawia się presja na osłabienie kursu. Bank centralny jest wtedy zmuszony sprzedawać waluty obce, co wywołuje spadek bazy monetarnej i wzrost stóp procentowych. W tej sytuacji zmiany w zakresie bilansu płatniczego ograniczają skuteczność polityki fiskalnej.

Rysunek 6.3. Skuteczność polityki fiskalnej przy sztywnym kursie walutowym w modelu Mundella-Fleminga w warunkach silnej i słabej mobilności kapitału



Źródło: opracowanie własne.

Rysunek 6.3 przedstawia skuteczność ekspansywnej polityki fiskalnej w warunkach stałego kursu walutowego w ramach modelu Mundella-Fleminga. Lewa strona rysunku ukazuje sytuację, w której kapitał reaguje silnie na wzrost stóp procentowych (duża mobilność kapitału) i dzięki temu efekt kapitałowy staje się silniejszy niż efekt handlowy (krzywa *BP* jest względnie płaska). Początkowym punktem równowagi jest *A*. Wzrost wydatków rządowych przesuwa początkowo krzywą *IS* w prawo do *IS'*, a gospodarka przemieszcza się najpierw do punktu *B*. Lecz wywołany zwiększeniem wydatków wzrost stóp procentowych powoduje spadek inwestycji i w efekcie krzywa *IS* przesuwa się w lewo do *IS"*. Tak długo, jak nowa krzywa *IS"* znajduje się powyżej początkowej krzywej *IS*, autonomiczny wzrost wydatków rządowych pozostaje jednak większy niż spadek inwestycji (takie założenie przyjęto na rysunku 6.3). Punkt równowagi *C* leży natomiast powyżej krzywej *BP*, co oznacza

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

nadwyżkę płatniczą (napływ kapitału jest większy niż odpływ kapitału z tytułu pogorszenia salda rachunku bieżącego). W systemie kursu stałego nadwyżka płatnicza oznacza wzrost podaży pieniądza krajowego, a zatem krzywa *LM* przesuwa się w prawo do *LM'*. Wzrost podaży pieniądza pociąga za sobą obniżenie stóp procentowych, odpływ kapitału za granicę i spadek nadwyżki. Gospodarka przesuwa się do nowego punktu równowagi *D*, w którym nadwyżka płatnicza zanika. Polityka fiskalna okazuje się w tych warunkach skuteczna, ponieważ poziom produkcji i dochodu w punkcie *D* jest zdecydowanie wyższy niż w początkowym punkcie *A*.

Prawa strona rysunku 6.3 ukazuje zaś sytuację, w której kapitał słabo reaguje na wzrost stóp procentowych (ograniczona mobilność kapitału), a w rezultacie efekt handlowy staje się silniejszy niż efekt kapitałowy (krzywa *BP* jest względnie stroma). Początkowym punktem równowagi jest *A*. Wzrost wydatków rządowych przesuwa krzywą *IS* w prawo do *IS'*, a gospodarka przemieszcza się początkowo do punktu *B*. Lecz – podobnie jak poprzednio – wzrost stóp procentowych wywołuje spadek inwestycji, a w efekcie krzywa *IS* przesuwa się w lewo do *IS"*. Tak długo, jak nowa krzywa *IS"* znajduje się powyżej początkowej krzywej *IS*, autonomiczny wzrost wydatków rządowych pozostanie jednak większy niż spadek inwestycji. Punkt równowagi *C* leży natomiast poniżej krzywej *BP*, co oznacza pojawienie się deficytu płatniczego (napływ kapitału jest mniejszy niż odpływ kapitału z tytułu pogorszenia salda rachunku bieżącego). W systemie kursu stałego deficyt płatniczy oznacza spadek podaży pieniądza krajowego, a zatem krzywa *LM* przesuwa się tym razem w lewo do *LM'*. Spadek podaży pieniądza pociąga za sobą wzrost stóp procentowych, napływ kapitału zagranicznego i obniżenie deficytu. Gospodarka przesuwa się do nowego punktu równowagi *D*, w którym deficyt płatniczy zanika. Polityka fiskalna okazuje się w tych warunkach mało skuteczna, ponieważ poziom produkcji i dochodu w punkcie *D* jest tylko nieznacznie wyższy niż w początkowym punkcie *A*.

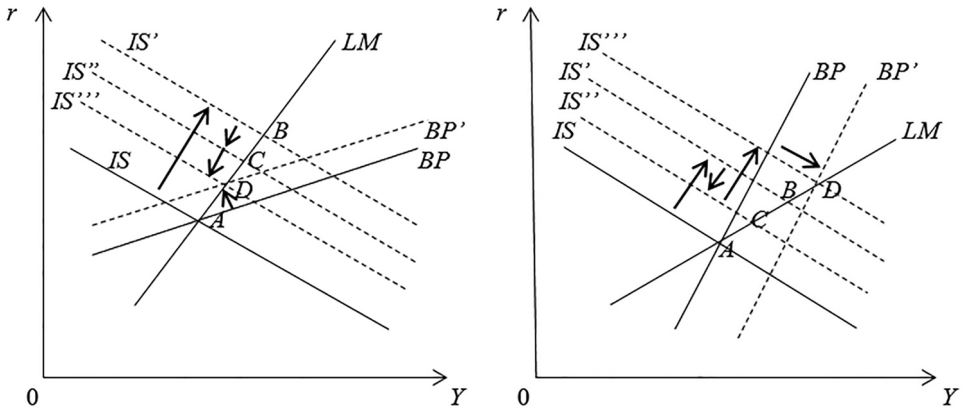
Pamiętajmy jednak, że w obu przypadkach zakładaliśmy, że spadek inwestycji wywołany wzrostem stóp procentowych jest mniejszy niż początkowy wzrost wydatków publicznych. Tak nie musi jednak być. Gdy spadek inwestycji okaże się większy niż wzrost wydatków rządowych, krzywa *IS* przesunie się ostatecznie w lewo od początkowego położenia, a nowy punkt równowagi będzie cechował się niższym poziomem dochodu i produkcji. Ekspansywna polityka fiskalna stanie się w takiej sytuacji przeciwskuteczna.

W systemie **kursu płynnego** wzrost wydatków rządowych, który nie jest finansowany dodrukiem pieniądza przez bank centralny, wywiera początkowo taki sam

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wpływ na popyt zagregowany jak w przypadku stałego kursu. Efekt netto w odniesieniu do gospodarki będzie pozytywny pod warunkiem, że autonomiczny wzrost wydatków rządowych okaże się większy niż spadek inwestycji z tytułu wzrostu stóp procentowych. Również zmiany dokonujące się w zakresie bilansu płatniczego będą miały podobny charakter. Wzrost importu spowodowany wzrostem wydatków rządowych wywoła presję na osłabienie kursu waluty, a napływ kapitału spowodowany wzrostem stóp procentowych doprowadzi do powstania presji na umocnienie kursu waluty. Różnica polega na tym, że w tym przypadku kurs ulega zmianie i równowaga płatnicza jest przywracana dzięki zmianom kursu, a nie za sprawą zmian podaży pieniądza rezerwowego. Gdy przeważy efekt handlowy, nastąpi deprecjacja kursu, która wywoła dodatkowy wzrost produkcji, dochodu i zatrudnienia (wzrasta eksport, spada import, rośnie popyt zagraniczny, a popyt krajowy przesuwa się na produkcję krajową – tzw. *demand shift*). Polityka fiskalna jest wówczas skuteczna. Gdy przeważy jednak efekt kapitałowy, dojdzie do aprecjacji kursu, która spowoduje obniżenie popytu na produkcję krajową, a polityka fiskalna okaże się w tej sytuacji nieskuteczna. Ostateczny rezultat będzie więc zależał od relatywnej siły efektów cząstkowych: reakcji inwestycji na wzrost stóp procentowych i wrażliwości kapitału na wzrost stóp procentowych. Im słabsza reakcja inwestycji oraz im mniejsza wrażliwość kapitału na wzrost stóp procentowych, tym silniejsze staje się osłabienie kursu i tym skuteczniejsza jest polityka fiskalna w pobudzaniu produkcji i zatrudnienia.

Rysunek 6.4. Skuteczność ekspansywnej polityki fiskalnej w systemie płynnego kursu walutowego w warunkach silnej i słabej mobilności kapitału



Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 6.4 ukazuje skuteczność ekspansywnej polityki fiskalnej w warunkach płynnego kursu walutowego w ramach modelu Mundella-Fleminga (zob. Acocella 2005). Lewa strona ilustracji przedstawia sytuację, w której kapitał reaguje silnie na wzrost stóp procentowych (duża mobilność kapitału) i dzięki temu efekt kapitałowy jest silniejszy niż efekt handlowy (krzywa *BP* jest względnie płaska). Początkowym punktem równowagi jest *A*. Wzrost wydatków rządowych przesuwa początkowo krzywą *IS* w prawo do *IS'*, a gospodarka przemieszcza się najpierw do punktu *B*. Lecz wywołany zwiększeniem wydatków wzrost stóp procentowych powoduje spadek inwestycji i w efekcie krzywa *IS* przesuwa się w lewo do *IS"*. Tak długo, jak nowa krzywa *IS"* znajduje się powyżej początkowej krzywej *IS*, autonomiczny wzrost wydatków rządowych pozostaje jednak wciąż większy niż spadek inwestycji (taką sytuację przedstawiono na rysunku 6.4). Punkt równowagi *C* leży natomiast powyżej krzywej *BP*, co oznacza nadwyżkę płatniczą (napływ kapitału jest większy niż odpływ kapitału z tytułu pogorszenia salda rachunku bieżącego). W systemie kursu płynnego nadwyżka płatnicza przekłada się na aprecjację waluty krajowej, co pociąga za sobą spadek eksportu netto i zmniejszenie produkcji krajowej. Krzywa *IS* przesuwa się dalej w lewo do położenia *IS'"*. W tych warunkach zachowanie równowagi płatniczej wymaga wyższego poziomu stóp procentowych w przypadku każdego danego poziomu dochodu, tak aby dodatkowy napływ kapitału zrównoważył pogorszenie salda rachunku bieżącego na skutek aprecjacji. W efekcie krzywa *BP* przesuwa się w lewo. Przecięcie się krzywej *BP'* i krzywej *IS'"* w punkcie *D* wyznacza nowy punkt równowagi, w którym nadwyżka płatnicza zanika. Polityka fiskalna okazuje się w tych warunkach skuteczna, ponieważ poziom produkcji i dochodu w punkcie *D* jest wyższy niż w początkowym punkcie *A*. Skuteczność pozostaje jednak umiarkowana, gdyż efekt pobudzający impulsu fiskalnego jest częściowo zneutralizowany efektem aprecjacji kursu waluty spowodowanej silnym napływem kapitału z zagranicy w reakcji na wzrost stóp procentowych.

Prawa strona rysunku 6.4 przedstawia sytuację, w której mobilność kapitału jest ograniczona – kapitał słabo reaguje na wzrost stóp procentowych. W rezultacie efekt handlowy jest silniejszy niż efekt kapitałowy (krzywa *BP* jest stroma). Początkowym punktem równowagi jest *A*. Wzrost wydatków rządowych (przy niezmienionej podaży pieniądza) wywołuje wzrost dochodu i zwiększenie stóp procentowych – krzywa *IS* przesuwa się najpierw w prawo do *IS'*, a następnie do *IS"* na skutek spadku inwestycji. Gospodarka przesuwa się z punktu *A* do punktu *B*, a następnie do punktu *C*. Ponieważ efekt handlowy przeważa nad efektem kapitałowym, w bilansie płatniczym pojawia się deficyt (punkt *C* znajduje się na prawo

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

od linii *BP*) i w efekcie waluta krajowa ulega deprecjacji. Ta z kolei wywołuje wzrost eksportu netto i zwiększenie dochodu – krzywa *IS* przesuwa się w prawo do *IS'''*. Jednocześnie deprecjacja kursu powoduje zmianę położenia krzywej *BP* – przy danym poziomie stóp procentowych zachowanie równowagi w bilansie płatniczym wymaga wyższego poziomu dochodu (tak aby wydatki na import uległy zwiększeniu). W rezultacie krzywa *BP* przesuwa się w prawo do *BP'*, a gospodarka przemieszcza się z punktu *C* ostatecznie do punktu równowagi *D*. Polityka fiskalna okazuje się więc skuteczna, ponieważ poziom produkcji i dochodu w punkcie *D* jest znacznie wyższy niż w początkowym punkcie *A*. Również skuteczność pozostaje wysoka, gdyż efekt pobudzający impulsu fiskalnego jest dodatkowo wzmocniony efektem deprecjacji kursu, będącej skutkiem pogorszeniem rachunku bieżącego.

Jak widać, w systemie kursu zmiennego ostateczny efekt zastosowania ekspansywnej polityki fiskalnej zależy od stopnia mobilności kapitału. Przy niskiej mobilności kapitału, wystąpi tendencja do deprecjacji waluty krajowej i w konsekwencji skuteczność polityki fiskalnej okaże się wyższa. Rysunek 6.4 pokazuje, że im niższa jest mobilność kapitału, tym wyżej położona będzie krzywa *IS* i tym większe stanie się przesuniecie krzywej *BP*. Z kolei przy wysokiej mobilności kapitału wystąpi tendencja do aprecjacji waluty krajowej i w konsekwencji skuteczność polityki fiskalnej będzie ograniczona. Rysunek 6.4 pokazuje, że im wyższa okaże się mobilność kapitału, tym niżej będzie położona krzywa *IS* i tym mniejsze stanie się przesuniecie krzywej *BP*. W skrajnym przypadku doskonałej mobilności kapitału krzywa *BP* ma kształt poziomy, a polityka fiskalna staje się całkowicie nieskuteczna. Tabela 6.4 podsumowuje omówione warunki skuteczności polityki fiskalnej w ramach różnych reżimów kursowych i przy różnym poziomie mobilności kapitału.

Tabela 6.4. Skuteczność polityki fiskalnej przy sztywnym i płynnym kursie walutowym

Kurs sztywny		Kurs płynny	
Pełna mobilność kapitału	Brak mobilności kapitału	Pełna mobilność kapitału	Brak mobilności kapitału
Polityka fiskalna bardzo skuteczna, napływ kapitału wywołuje wzrost podaży pieniądza.	Polityka fiskalna częściowo skuteczna, wzrost deficytu obrotów bieżących powoduje spadek rezerw walutowych i spadek podaży pieniądza.	Polityka fiskalna mało skuteczna lub całkowicie nieskuteczna, napływ kapitału powoduje aprecjację waluty.	Polityka fiskalna skuteczna, wzrost deficytu obrotów bieżących powoduje deprecjację waluty.

Źródło: na podstawie Acocella (2005, s. 298–302).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

6.4. Wymienialność waluty krajowej

6.4.1. Przyczyny występowania ograniczeń w zakresie wymienialności walut krajowych na waluty obce

Wymienialność walutowa jest powszechnie definiowana jako prawnie zagwarantowana swoboda wymiany waluty narodowej na inne waluty lub na złoto. Nie wszystkie waluty narodowe są jednak wymienialne. W niektórych krajach stosuje się administracyjne ograniczenia w zakresie wymienialności waluty krajowej na waluty obce. W warunkach istnienia tych ograniczeń rezydenci danego kraju (przedsiębiorstwa, osoby fizyczne) nie mogą swobodnie kupować waluty obcej za walutę krajową. Ograniczenia takie mogą dotyczyć swobody zakupu walut obcych zarówno na potrzeby dokonania przez podmioty krajowe importu dóbr czy usług, jak i na potrzeby zakupu przez nie aktywów zagranicznych, udzielenia pożyczki zagranicznej lub dokonania transferu bezzwrotnego za granicę. Ograniczenia w zakresie wymienialności były stosowane powszechnie od II wojny światowej do końca lat 80. W okresie ostatnich trzydziestu lat wiele krajów zdecydowało się jednak wycofać te restrykcje lub radykalnie zmniejszyć zakres ich stosowania. W rezultacie większość walut narodowych jest obecnie albo w pełni wymienialna, albo podlega jedynie niewielkim ograniczeniom wymienialności.

Istnieje kilka powodów, dla których kraje stosują ograniczenia w zakresie wymienialności walut. Pierwszym jest chęć utrzymania stabilnego, często zaniżonego kursu waluty przy jednoczesnym ograniczeniu ryzyka odpływu kapitału i utraty rezerw walutowych. Motyw ten stał się szczególnie popularny wśród państw rozwijających się, gdzie rezerwy walutowe były bardzo skromne i rządy starały się – poprzez ograniczenie wymienialności – kontrolować wydatki importowe. Kolejną przyczynę –powiązaną jednocześnie z poprzednim czynnikiem – może stanowić dążenie do utrzymania równowagi płatniczej w warunkach mało konkurencyjnej, administracyjnie zarządzanej gospodarki. W tym przypadku ze względu na ogólne ograniczenia w funkcjonowaniu mechanizmu rynkowego, powstaje nadwyżkowy popyt na waluty obce, przy danym – zaniżonym – poziomie kursu walutowego. Jeszcze innym powodem stało się dążenie do utrzymania możliwie dużego zakresu autonomii w polityce pieniężnej. Ograniczenie zakresu wymienialności pozwala na zmniejszenie skali wahań poziomu rezerw walutowych, a to zwiększa kontrolę banku centralnego nad podażą pieniądza krajowego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W drugiej połowie XX wieku w wielu krajach popularne było stosowanie nakazowo-rozdzielczych metod zarządzania gospodarką narodową w ramach systemów centralnego planowania. W takich warunkach, przy administracyjnie ustalanych cenach, powstawała tendencja do niedoborów rynkowych, w tym także niedoborów walut obcych. Utrzymanie równowagi płatniczej wymagało w takiej sytuacji administracyjnego ograniczenia wymienialności waluty krajowej i zastąpienia rynku walutowego systemem administracyjnego rozdzielnictwa walut obcych. Obrodziło to jednak w szereg negatywnych konsekwencji. Po pierwsze, w takim systemie cena waluty obcej była z reguły poważnie zaniżona, co prowadziło do nieefektywnego i nieoptymalnego wykorzystywania zasobów walutowych kraju. W szczególności zaniżony kurs zaburzał konkurencyjność krajowego sektora eksportowego i zachęcał do nadmiernego importu. Po drugie, rosły koszty transakcji międzynarodowych w związku z koniecznością uzyskiwania zezwoleń na zakup walut obcych. Po trzecie, ze względu na to, że ograniczenia w zakresie wymienialności dotyczyły głównie przepływów finansowych, kraje uboższe i rozwijające się nie mogły korzystać z zasobów kapitałowych państw rozwiniętych w celu przyspieszania własnego rozwoju. Wreszcie, za nieuchronną konsekwencję takiego stanu rzeczy przychodzi uznać powstanie i rozwój „czarnego" rynku walutowego, który dodatkowo zniekształcał sygnały rynkowe i obniżał efektywność alokacji ograniczonych zasobów walut obcych.

Dlatego też już bezpośrednio po II wojnie światowej podjęto starania mające na celu stopniowe przywrócenie wymienialności walut. Jednym z głównych celów utworzonego w 1944 roku Międzynarodowego Funduszu Walutowego (*International Monetary Fund*) stało się właśnie wspieranie tego typu działań (zob. ramka 6.4). Początki nie były łatwe. W pierwszych latach po wojnie w pełni wymienialne waluty posiadało jedynie kilka państw (m.in. USA). Większość krajów borykała się bowiem z trudnościami w odzyskaniu równowagi płatniczej zarówno ze względu na długi zaciągnięte w czasie wojny, jak i chroniczny niedobór dolarów amerykańskich spowodowany strukturalnym deficytem w handlu z tym krajem. Dopiero w 1958 roku w większości państw Europy Zachodniej zniesiono ograniczenia w zakresie wymienialności walut na potrzeby transakcji bieżących bilansu płatniczego, obejmujących eksport i import dóbr i usług, dochody czynników produkcji oraz transfery bieżące. W dalszej kolejności liberalizowano wymienialność w zakresie obrotów kapitałowych bilansu płatniczego. Znoszenie ograniczeń przebiegało najszybciej w krajach rozwiniętych. Na początku lat 90. pełna wymienialność obowiązywała już praktycznie we wszystkich państwach rozwiniętych. Ich śladem poszła też duża

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

liczba krajów rozwijających się, w tym zwłaszcza tzw. kraje wschodzące (*emerging economies*). Obecnie ponad połowa państw świata posiada w pełni lub prawie w pełni wymienialne waluty.

Ramka 6.4. Wymienialność walut w statucie MFW

Zapisy dotyczące zobowiązań państw członkowskich do uchylania ograniczeń w zakresie wymienialności walut zostały zawarte w kilku artykułach statutu MFW. Obowiązek do zapewnienia wymienialności odnosi się jednak tylko do transakcji bieżących. Art. I w punkcie b) stanowi, że jednym z celów MFW jest:

wspieranie działań na rzecz ustanowienia wielostronnego systemu płatności w odniesieniu do bieżących transakcji międzynarodowych między państwami członkowskimi i usunięcia restrykcji walutowych, które hamują wzrost handlu światowego (to assist in the establishment of a multilateral system of payments in respect of current transactions between members and in the elimination of foreign exchange restrictions which hamper the growth of world trade).

W punkcie c) art. IV dotyczącym zobowiązań odnoszących się do reżimów kursowych, państwa członkowskie zobowiązują się

powstrzymać od wpływania na kursy walutowe lub na międzynarodowy system walutowy, tak aby nie zakłócać efektywnych dostosowań w bilansach płatniczych i nie dążyć do uzyskania nieuczciwej przewagi konkurencyjnej nad innymi państwami członkowskimi (avoid manipulating exchange rates or the international monetary system in order to prevent effective balance of payments adjustment or to gain an unfair competitive advantage over other members).

W sekcji 2 art. VIII odnoszącego się do zobowiązań ogólnych państw członkowskich czytamy, że, uwzględniając zapisy sekcji 3(b) art. VII oraz sekcji 2 art. XIV, państwa członkowskie zobowiązują się nie nakładać, bez zgody Funduszu, „ograniczeń na płatności i transfer środków w związku z dokonywaniem bieżących transakcji międzynarodowych" (no member shall, without the approval of the Fund, impose restrictions on the making of payments and transfers for current international transactions).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Z kolei w sekcji 3 art. VIII znalazł się następujący zapis:

żadne państwo członkowskie nie zaangażuje się ani nie zezwoli jakiejkolwiek własnej rządowej agencji fiskalnej na stosowanie jakichkolwiek praktyk dyskryminacyjnych w odniesieniu do walut obcych… (no member shall engage in, or permit any of its fiscal agencies referred to in Article V, Section 1 to engage in, any discriminatory currency arrangements or multiple currency practices…)[12].

Jednocześnie zapisy art. XIV pozwalają państwom członkowskim utrzymywać przejściowe (*transitory*) restrykcje w zakresie wymiany walut i płatności bieżących pod warunkiem notyfikowania ich Funduszowi oraz dokonywania corocznego przeglądu tych restrykcji wspólnie z Funduszem, z uwagi na postanowienia art. I nakładające obowiązek znoszenia tych ograniczeń. Wyjątek dopuszczony art. XIV dotyczy, między innymi, państw, które stosowały określone ograniczenia w zakresie wymienialności w momencie wstąpienia do MFW oraz państw przezywających poważne trudności z utrzymaniem równowagi w bilansie płatniczym. Ten rodzaj ograniczonej wymienialności nosi nazwę „wymienialności wewnętrznej" [13].

Kluczowe znaczenie w przypadku wymienialności mają jednak zapisy sekcji 2 art. VIII, które nakładają na państwa członkowskie zobowiązanie utrzymania wymienialności walut krajowej w ramach wszystkich transakcji bieżących bilansu płatniczego. Dlatego też ten rodzaj wymienialności nazywany jest często „wymienialnością w znaczeniu art. VIII MFW" (*Article VIII convertibility*).

6.4.2. Zakres wymienialności

Jak wspomniano wyżej, wymienialność waluty krajowej może dotyczyć wszystkich transakcji zagranicznych lub tylko niektórych ich rodzajów. Pod tym względem najczęściej rozróżnia się wymienialność na rachunku obrotów bieżących bilansu płatniczego i wymienialność na rachunku obrotów finansowych. W pierwszym przypadku waluta krajowa jest swobodnie wymienialna na waluty obce tylko na potrzeby przeprowadzenia transakcji bieżących, takich jak import dóbr i usług, transfery bieżące oraz przepływ dochodów w zakresie czynników produkcji. Ten

[12] https://www.imf.org/external/pubs/ft/aa/#intro (20.06.2015).

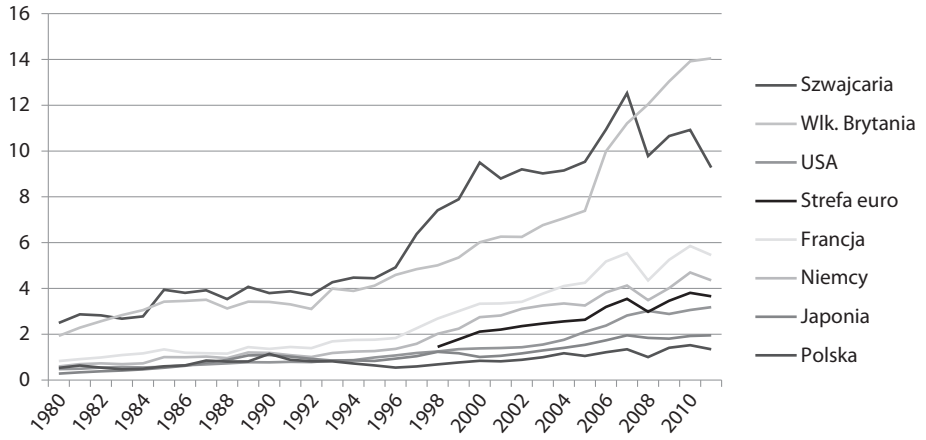
[13] Polska korzystała z wyłączeń dopuszczonych zapisami art. XIV w latach 1990–1995, ograniczając m.in. swobodę wymienialności w zakresie transakcji transferowych oraz tych dokonywanych przez cudzoziemców (zob. Rosati 1998, s. 242).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

rodzaj wymienialności, zalecany przez MFW, stosuje ogromna większość krajów świata. W drugim przypadku waluta krajowa jest swobodnie wymienialna na waluty obce również na potrzeby przeprowadzenia transakcji kapitałowych i finansowych, takich jak nabywanie aktywów za granicą w ramach inwestycji bezpośrednich lub portfelowych, udzielanie zagranicznych kredytów i pożyczek czy lokowanie środków w bankach zagranicznych. Mówi się wówczas o swobodzie transgranicznych przepływów kapitału (*cross-border capital mobility*). Ten rodzaj wymienialności stosują wszystkie państwa rozwinięte, a także większość tzw. krajów wschodzących.

Zakres wymienialności ma istotne znaczenie w przypadku kształtowania się kursu walutowego. Gdy waluta krajowa pozostaje wymienialna tylko na potrzeby przeprowadzania transakcji bieżących, kurs walutowy określony jest głównie przez strumienie płatności eksportowych i importowych. Ponieważ wielkości tych strumieni są stosunkowo stabilne w czasie (zmiany poziomu eksportu i importu dóbr i usług są wprawdzie większe niż zmiany poziomu PKB, lecz rzadko przekraczają 10–15% w skali roku), to również kurs wykazuje się stosunkowo dużą stabilnością. Gdy wymienialność dotyczy jednak także transakcji finansowych, wahania kursu mogą być znacznie silniejsze, ponieważ przepływy kapitałowe wykazują tendencje do dużych i często gwałtownych zmian, a ich skala jest wielokrotnie większa niż skala transakcji bieżących.

Wykres 6.6. Stopień „otwarcia finansowego" w wybranych krajach w latach 1980–2011 – stosunek sumy aktywów zagranicznych brutto i zobowiązań zagranicznych do PKB (dane w USD)



Źródło: na podstawie Philip Lane, http://www.philiplane.org/EWN.html (18.10.2014).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Poszerzanie zakresu wymienialności walut wśród większości krajów przyczyniało się do wzrostu skali międzynarodowych przepływów kapitałowych i pogłębiania integracji finansowej między poszczególnymi państwami. Wiarygodną miarą stopnia „finansowego otwarcia" danego kraju jest stosunek sumy aktywów zagranicznych brutto i sumy zobowiązań zagranicznych do wielkości PKB. Relacja ta wzrosła wyraźnie po 1980 roku. Wykres 6.6, wykorzystujący bazę danych zgromadzonych przez amerykańskiego ekonomistę Philipa Lane'a, ilustruje proces postępującego otwarcia finansowego w wybranych krajach. Szczególne przyspieszenie procesu otwarcia miało miejsce w latach 1992–2007. W najbardziej otwartych krajach, pełniących funkcję międzynarodowych centrów finansowych, łączna suma zagranicznych należności i zobowiązań była ponad dziesięciokrotnie większa niż PKB tych państw (zob. również Lane i Millesi-Ferretti 2007, Obstfeld 2010).

6.4.3. Czy kapitał powinien swobodnie przepływać między krajami?

W latach 1950–1980, gdy w świecie dominowały systemy kursów stałych, poszczególne kraje niechętnie godziły się na swobodę przepływów kapitałowych i utrzymywały liczne ograniczenia w zakresie wymienialności transakcji kapitałowych. Obawiano się głównie skutków w postaci masowego odpływu kapitału prywatnego. Taki odpływ, wywołany zwłaszcza czynnikami politycznymi bądź spekulacyjnymi, mógłby zagrażać stabilności kursu walutowego, hamować rozwój handlu i grozić kryzysem walutowym w przypadku wyczerpania się rezerw walutowych. W rezultacie transgraniczne przepływy kapitału ograniczały się głównie do pożyczek oficjalnych (międzyrządowych) i, w mniejszym stopniu, do zagranicznych inwestycji bezpośrednich, z kolei pozostałe przepływy były ściśle kontrolowane.

Ograniczenia w zakresie wymienialności transakcji kapitałowych miały jednak swoje istotne wady. Przede wszystkim nie pozwalały na optymalną alokację zasobów kapitałowych w skali międzynarodowej. Przedsiębiorstwa międzynarodowe z krajów rozwiniętych powstrzymywały się od inwestowania w krajach o niewymienialnych walutach przede wszystkim ze względu na brak możliwości repatriowania zysków i zainwestowanego kapitału. Ucierpiały na tym głównie kraje rozwijające się, które potrzebowały zagranicznych inwestycji, ponieważ ich własne zasoby krajowych oszczędności były niewielkie, a to ograniczało możliwości inwestycyjne i hamowało wzrost gospodarczy. Brak mobilności kapitału nie pozwalał także na transfer i dyfuzję nowoczesnych technologii. Jednocześnie nadwyżki kapitału w krajach zamożnych były przeznaczane na inwestycje w przedsięwzięcia o niskich stopach przychodu.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Co więcej, w ramach systemu stałych kursów i braku mobilności kapitału, poszczególne kraje miały ograniczone możliwości przywracania zakłóconej równowagi płatniczej. Na przykład utrzymujący się deficyt handlowy – w sytuacji braku napływu kapitału – prowadził do wyczerpywania się rezerw walutowych i ostatecznie zmuszał do przeprowadzenia dewaluacji. Alternatywą stawało się wówczas silne zaostrzenie polityki pieniężnej w celu zahamowania popytu, co pociągało jednak za sobą z reguły przewlekłą recesję i duży wzrost bezrobocia.

W ramach analizy makroekonomicznej przepływy kapitału były tradycyjnie traktowane jako symetryczne odpowiedniki salda rachunku obrotów bieżących. Przykładowo deficyt bieżący wynikał z nadwyżki inwestycji krajowych nad oszczędnościami państwa, co musiało zostać sfinansowane napływem oszczędności zagranicznych w postaci kredytów, pożyczek oraz inwestycji bezpośrednich. Również odwrotnie, nadwyżka oszczędności nad inwestycjami prowadziła do powstania nadwyżki bieżącej i odpływu kapitału za granicę. Neoklasyczna teoria ekonomii przewidywała, że w skali międzynarodowej kapitał zasadniczo płynie „w dół" (*downhill*), czyli od krajów zamożnych o dużych zasobach kapitału do krajów ubogich o niewielkich zasobach kapitału. Wynikało to z tego, że w krajach zamożnych relacja zasobu pracy do zasobu kapitału jest niska, a to implikuje niskie stopy przychodu z kapitału, z kolei w krajach ubogich relacja ta pozostaje wysoka, co implikuje wysokie stopy przychodu z kapitału. W efekcie kapitał przepływał do krajów o wysokich stopach zwrotu, a te dzięki temu mogły rozwijać się szybciej. Mechanizm ten stał się podstawą teorii realnej konwergencji (Barro 1991, Barro i Sala-i-Martin 1992), która przewidywała, że kraje uboższe rozwijają się szybciej, wobec czego są w stanie „doganiać" kraje bogate pod względem poziomu produkcji na mieszkańca. W związku z tym ujemne salda rachunków bieżących w krajach uboższych były z reguły traktowane jako przejaw procesu „doganiania" (*catching-up* – zob. Blanchard i Giavazzi 2002, Rosati 2004)[14].

Dlatego też w latach 80. najważniejsze międzynarodowe organizacje gospodarcze, takie jak MFW, Bank Światowy i OECD, zaangażowały się w działania mające na celu usunięcie barier ograniczających wymienialność walut krajowych

[14] W praktyce okazało się jednak, że napływ kapitału z krajów rozwiniętych do rozwijających się był ograniczony, a największe przepływy miały miejsce między krajami rozwiniętymi, co w sposób oczywisty kłóciło się z teorią. Zjawisko to – tzw. paradoks Lucasa – próbowano wyjaśniać różnicami w poziomie rozwoju kapitału ludzkiego (na korzyść krajów rozwiniętych), korzyściami pozarynkowymi z tytułu rozprzestrzeniania się wiedzy w skali międzynarodowej i niedoskonałościami na rynku kapitałowym w krajach rozwijających się wraz ze zwiększonym ryzykiem politycznym (Lucas 1990).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

na potrzeby transakcji kapitałowych. Liberalizacja przepływów kapitałowych stała się nieodzownym składnikiem programów reform gospodarczych zalecanych krajom rozwijającym się w ramach tzw. konsensu waszyngtońskiego (*Washington consensus*). Termin ten, wprowadzony do obiegu przez brytyjskiego ekonomistę Johna Williamsona, obejmował pakiet dziesięciu reform, mających na celu liberalizację i deregulację mikroekonomiczną, stabilizację makroekonomiczną oraz prywatyzację państwowych przedsiębiorstw w gospodarkach rozwijających się (zob. ramka 6.5). Konsens waszyngtoński odegrał też istotną rolę w procesie transformacji rynkowej post-socjalistycznych krajów Europy Środkowej i Wschodniej.

Ramka 6.5. Konsens waszyngtoński

Nazwa konsensu waszyngtońskiego (*Washington consensus*) pochodzi od miasta Waszyngton, stolicy USA i siedziby dwóch najważniejszych międzynarodowych organizacji finansowych – Międzynarodowego Funduszu Walutowego i Banku Światowego[15]. W swej oryginalnej wersji konsens został sformułowany w 1989 roku przez Johna Williamsona w postaci dziesięciu ogólnych rekomendacji polityki gospodarczej dla krajów stabilizujących swoje gospodarki i wprowadzających reformy prorynkowe (Williamson 1989):

1. **Zdyscyplinowana polityka fiskalno-budżetowa**, a zwłaszcza unikanie dużego deficytu w relacji do PKB;
2. **Zmiana struktury wydatków publicznych** poprzez obniżenie udziału subsydiów na rzecz wzrostu udziału usług pro-wzrostowych, takich jak podstawowa edukacja, podstawowa opieka zdrowotna oraz infrastruktura;
3. **Reforma podatkowa**, w tym przede wszystkim poszerzenie bazy podatkowej i przyjęcie umiarkowanych krańcowych stawek podatkowych;
4. **Polityka stóp procentowych** uwzględniająca relacje rynkowe i zapewniająca umiarkowanie dodatni poziom stóp realnych;
5. **Konkurencyjne kursy walutowe**;
6. **Liberalizacja handlu zagranicznego**, w tym zwłaszcza zniesienie ograniczeń ilościowych w imporcie i ograniczenie się w polityce handlowej do stosowania taryf celnych o niskich i w miarę jednolitych stawkach;

[15] Bank Światowy (*World Bank*) formalnie nosi nazwę Międzynarodowego Banku Odbudowy i Rozwoju (*International Bank for Reconstruction and Development* – IBRD).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

7. **Liberalizacja napływu bezpośrednich inwestycji zagranicznych**;
8. **Prywatyzacja przedsiębiorstw państwowych**;
9. **Deregulacja** poprzez zniesienie przepisów ograniczających konkurencję i dostęp do rynków, z wyjątkiem takich, które są uzasadnione względami bezpieczeństwa, ochrony konsumenta i ochrony środowiska naturalnego oraz koniecznością nadzoru nad instytucjami finansowymi;
10. Ochrona prawna **praw własności**.

Konsens odzwierciedlał dominujące w latach 80. poglądy ekonomistów tzw. głównego nurtu (*mainstream*) na temat stabilizacji makroekonomicznej i zakresu pożądanych reform rynkowych w krajach rozwijających się, które dążyły do poprawy konkurencyjności i trwałego przyspieszenia wzrostu gospodarczego. Rekomendacje konsensu były powszechnie stosowane w tzw. programach dostosowawczych, wdrażanych razem z pomocą finansową udzielaną przez MFW krajom przeżywającym trudności gospodarcze w postaci wysokiej inflacji, wysokiego bezrobocia, deficytu w bilansie płatniczym i nadmiernego zadłużenia.

Koszty społeczne związane ze stosowaniem programów reform w oparciu o konsens waszyngtoński stały się przyczyną spadku ich popularności Najostrzejsza krytyka płynęła ze strony ekonomistów reprezentujących poglądy anty- lub alterglobalistyczne oraz ze strony polityków reprezentujących partie lewicowe i populistyczne. Na przykład liberalizacja handlu zagranicznego była postrzegana jako wyrafinowany sposób penetracji rynku krajów rozwijających się przez międzynarodowe korporacje. Prywatyzacja była z kolei traktowana jako wyprzedaż za bezcen cennych aktywów krajowych, a deregulacja miała stanowić przyzwolenie dla nadużyć i ograniczania praw pracowniczych. Krytycy (Noam Chomsky, Susan George czy Naomi Klein) traktują generalnie konsens waszyngtoński jako przejaw ekonomicznego neoliberalizmu oraz instrument pogłębiający globalizację, która generalnie prowadzi do wyzysku pracowników w krajach uboższych i dalszego pogłębienia nierówności dochodowych i majątkowych na świecie.

Bardziej umiarkowani krytycy wskazują, że konsens waszyngtoński nie obejmuje wszystkich niezbędnych reform, takich jak na przykład działania nakierowane na wzmocnienie potencjału innowacyjnego gospodarki, aktywna polityka rynku pracy, wsparcie dla małych i średnich przedsiębiorstw, wzmocnienie pionu sądownictwa gospodarczego czy konieczność socjalnej ochrony najuboższych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Pod wpływem tych działań w latach 90. doszło do szybkiej liberalizacji przepływów kapitałowych. Ich efekty okazały się jednak zróżnicowane. Kryzysy walutowe w Meksyku oraz w krajach Azji Południowo-Wschodniej w latach 1996–1998 dały wyobrażenie o skali ryzyka związanego ze swobodą przepływów kapitału. Pokazały również, że systematyczny napływ kapitału, zwłaszcza w formie krótkoterminowych pożyczek bankowych, powoduje wzrost inflacji, spekulacyjny wzrost cen aktywów i wzrost zadłużenia. Po przekroczeniu pewnego poziomu zadłużenia ryzyko niewypłacalności kraju gwałtownie rośnie i inwestorzy zagraniczni zaczynają masowo wycofywać swój kapitał. Nagłe wstrzymanie napływu kapitału (tzw. *sudden stop*), a następnie jego odpływ (*flow reversal*) wywołują silną presję na osłabienie kursu waluty. Ponieważ rezerwy walutowe banku centralnego są zawsze ograniczone, prędzej czy później dochodzi do załamania kursu i gospodarka wpada w kryzys walutowy. Co więcej, taki kryzys nierzadko prowadzi także do przewlekłej recesji. Gwałtowne zatrzymanie lub odwrócenie przepływu kapitału może dokonać się bardzo szybko, lecz odpowiednie dostosowanie w sferze realnej (spadek importu i wzrost eksportu w celu domknięcia luki w bilansie obrotów bieżących) wymaga dłuższego czasu. Powodem takiego stanu rzeczy jest istnienie nominalnych sztywności (*nominal rigidities*) płac i cen, które dostosowują się powoli do zmienionej sytuacji (Obstfeld i Rogoff 2005, Mendoza 2010).

Kryzysy końca lat 90., a zwłaszcza światowy kryzys finansowy mający miejsce w latach 2008–2010, uświadomiły także, że napływ kapitału zagranicznego tylko częściowo finansuje inwestycje produkcyjne. Duża część napływu przeznaczana jest bowiem na inwestycje nieprodukcyjne, spekulacyjne – głównie w formie nieruchomości i aktywów finansowych. Gdy ceny tych aktywów przestają rosnąć i zaczynają spadać, obsługa zaciągniętych na ich zakup kredytów załamie się, a gospodarka wpadnie w kryzys zadłużenia.

Ryzyko wystąpienia kryzysu walutowego oraz kryzysu zadłużenia jest szczególnie wysokie w warunkach stałego kursu walutowego. Płynny kurs zmniejsza zagrożenie kryzysem, lecz nie eliminuje go całkowicie. Ryzyko kryzysu występuje bowiem zawsze, gdy kraj posiada własną walutę, a kapitał pozostaje mobilny w skali międzynarodowej.

Pod wpływem tych doświadczeń Międzynarodowy Fundusz Walutowy złagodził w ostatnich latach swoje stanowisko w sprawie liberalizacji przepływów kapitałowych. Uznano, że korzyści płynące z napływu mają charakter bardziej potencjalny niż automatyczny, a ryzyko związane z przepływami jest znacznie większe niż dotychczas sądzono (Darvas i in. 2014). Uznano także, że w sytuacji zagrożenia

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

dużą nierównowagą w zakresie bilansu płatniczego poszczególne kraje mogą ograniczać swobodę przepływów kapitału przy pomocy instrumentów kontroli administracyjnej (zakazy, nakazy) lub fiskalnej (opodatkowanie przepływów). Oznacza to w praktyce odejście od jednego z głównych filarów konsensu waszyngtońskiego.

6.5. Rynek walutowy

6.5.1. Podaż i popyt na rynku walutowym

Jak wspomnieliśmy wcześniej, kurs walutowy jest ceną jednostki waluty obcej. Może być ona ustalona administracyjnie przez władzę gospodarczą – mamy wówczas do czynienia z systemem kursu stałego (sztywnego). W ramach takiego systemu wszystkie transakcje kupna i sprzedaży walut odbywają się po z góry ustalonym kursie, a bank centralny jest odpowiedzialny za dostarczanie i skup walut obcych na rynku walutowym. Podaż i popyt na waluty obce dostosowują się do stałej ceny – czyli kursu, a różnica między podażą a popytem w każdym momencie jest wyrównywana przez zmianę oficjalnych rezerw walutowych danego kraju. Stały kurs jest stosowany przez te państwa, które przede wszystkim dążą do stabilizowania warunków wymiany międzynarodowej i ograniczenia niepewności związanej z wpływem wahań kursu na działalność gospodarczą.

W wielu innych krajach kurs kształtuje się jednak mniej lub bardziej swobodnie na rynku walutowym pod wpływem zmian w zakresie wielkości podaży i popytu. Firmy, banki oraz gospodarstwa domowe kupują i sprzedają waluty obce, aby dokonywać transakcji międzynarodowych. Na przykład źródłem podaży walut obcych w Polsce są dochody uzyskane przez polskich eksporterów ze sprzedaży dóbr i usług za granicą, dochody uzyskane przez polskie firmy mające swoje filie i oddziały w innych krajach oraz przez polskich obywateli pracujących za granicą, a także transfery płynące do Polski z budżetu Unii Europejskiej. Oprócz tego źródłem podaży walut obcych stają się także inwestycje finansowe dokonywane przez inwestorów zagranicznych w naszym kraju. Gdy zagraniczny inwestor chce kupić polskie obligacje skarbowe, akcje polskich przedsiębiorstw czy też po prostu założyć terminową lokatę złotową w polskim banku, musi najpierw zamienić posiadane środki w walutach obcych na polskie złote. Wskutek tego rośnie podaż walut obcych na krajowym rynku walutowym. Banki krajowe dokonują również wymiany walut obcych na złote w ramach zaciągania pożyczek za granicą oraz

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

transakcji terminowych i pochodnych. Wreszcie, źródłem podaży walut są także zagraniczni turyści odwiedzający nasz piękny kraj.

Z kolei popyt na waluty obce w Polsce zgłaszają przede wszystkim importerzy, którzy nabywają waluty w celu dokonania zakupu dóbr i usług za granicą, a także zagraniczne firmy i obywatele innych państw, którzy uzyskują dochody w Polsce i następnie transferują je do swoich krajów. Waluty obce są nabywane też przez polskich turystów wybierających się za granicę, a także przez rodzime firmy dokonujące inwestycji w innych krajach. Również banki krajowe dokonują wymiany złotych na waluty obce w ramach spłaty pożyczek zagranicznych, udzielania kredytów walutowych (na przykład kredytów hipotecznych na zakup nieruchomości) lub nabywania innych aktywów, takich jak zagraniczne papiery wartościowe lub lokaty w bankach zagranicznych.

Jak widać, celem kupna lub sprzedaży walut obcych może być dokonanie zarówno transgranicznej transakcji importu lub eksportu dóbr i usług, jak i transgranicznej transakcji zakupu lub sprzedaży aktywów (finansowych lub rzeczowych). Rozróżnienie między transakcjami walutowymi związanymi z finansowaniem przepływów handlowych i transakcjami walutowymi dotyczącymi finansowania zakupu i sprzedaży aktywów stało się podstawą do formułowania różnych, alternatywnych teorii kształtowania się kursu walutowego. Tradycyjne podejście handlowe, nazywane też podejściem strumieniowym, uzależniało kształtowanie się kursu od relacji wielkości strumieni eksportu oraz importu dóbr i usług. Z kolei ujęcie, w którym główną rolę odgrywają operacje na aktywach, uzależnia kształtowanie się kursu waluty od popytu na aktywa finansowe i rzeczowe (*asset approach*). W ramach tych transakcji waluta obca funkcjonuje po prostu jako pieniądz zagraniczny, umożliwiający dokonanie danej transakcji. Najczęściej ich celem staje się nabycie aktywu zagranicznego, na przykład obligacji wyemitowanej przez rząd innego kraju, w celu uzyskania określonego dochodu. Na dochód ten składają się odsetki od danej obligacji (tzw. kupon) oraz różnica między kosztem zakupu obligacji a ceną nominalną, wypłacaną w momencie jej umorzenia. Również sama waluta obca może być traktowana jako szczególny rodzaj aktywu finansowego, który przynosi dochód. Składają się na niego odsetki od bankowej lokaty walutowej oraz dodatnia różnica między kosztem zakupu waluty obcej a ceną jej sprzedaży.

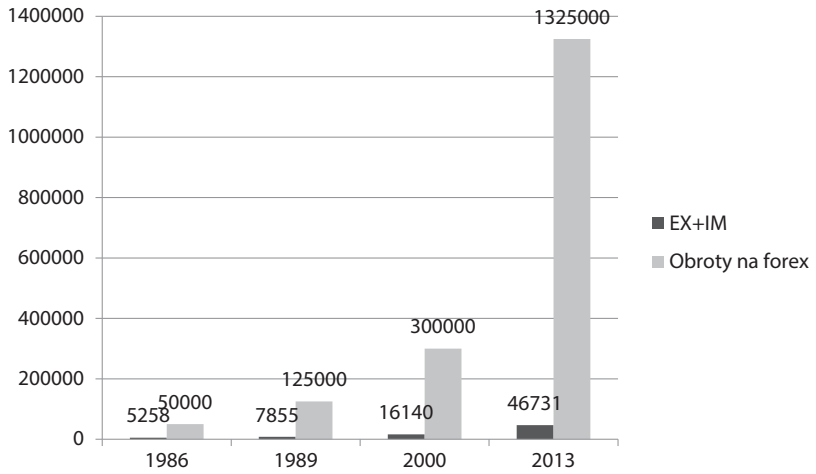
Zgodnie z parytetem stóp procentowych, inwestorzy kupują i sprzedają walutę obcą, uwzględniając relatywne stopy procentowe oraz oczekiwane zmiany kursu walut. W przypadku danej różnicy między krajową i zagraniczną stopą procentową inwestorzy nabywają określoną walutę obcą, gdy oczekują, że nastąpi umocnienie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jej kursu względem waluty krajowej, a sprzedają ją, gdy spodziewają się jej osłabienia. Ogromna większość tych transakcji dokonywana jest na rynku pieniądza zagranicznego, czyli na rynku walutowym (*foreign exchange market – forex*). Jego uczestnikami są przede wszystkim banki, które kupują i sprzedają waluty obce zarówno na rachunek własny, jak i na zlecenie oraz na rachunek swoich klientów.

W ostatnich latach, dzięki rozwojowi elektronicznych systemów rejestrowania i dokonywania transakcji (platformy cyfrowe), rynek forex stał się w dużym stopniu rynkiem zautomatyzowanym. Jednocześnie w obrotach na foreksie szybko wzrósł udział finansowych transakcji walutowych, niezwiązanych z transakcjami eksportu i importu, lecz mających charakter spekulacyjny. Wykres 6.7 ukazuje ewolucję struktury obrotów na rynkach walutowych w świecie – między 1986 a 2013 rokiem udział obrotów związanych z transakcjami handlu zagranicznego w całości obrotów na foreksie spadł z 10,5% do zaledwie 3,5%. Łączne obroty z tego rynku w 2013 roku były 30 razy większe niż suma globalnego eksportu oraz importu dóbr i usług, a także 18 razy większe niż światowy PKB.

Wykres 6.7. Obroty na rynkach walutowych na tle obrotów handlu zagranicznego w latach 1986–2013 (dane w mld USD)



Źródło: opracowanie własne na podstawie danych *Triennial Central Bank Survey of foreign exchange and derivatives market activity in 2013*, Bank for International Settlements, April 2013; WTO Statistics Database oraz Krugman i Obstfeld (1991, s. 322–323).

Najważniejszą walutą na rynkach walutowych pozostaje niezmiennie dolar amerykański, który występuje w ponad 70% wszystkich transakcji wymiany walut.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Na drugim miejscu znajduje się euro (33%), a na trzecim japoński jen (23%). Wśród pozostałych najważniejszych walut międzynarodowych znalazły się funt brytyjski (11,8%), dolar australijski (8,6%), frank szwajcarski (5,2%) i dolar kanadyjski (4,6%)[16].

Oprócz transakcji bieżących (*spot transactions*) na rynku forex zawierane są także transakcje terminowe i transakcje na instrumenty pochodne. Służą one głównie zabezpieczeniu eksporterów i importerów przed ryzykiem zmiany kursu walutowego, ale często ich celem staje się również spekulacja rynkowa. Kursy walut kwotowane w tych transakcjach noszą odpowiednio nazwę kursu *spot* (kursu bieżącego) i kursu terminowego. Różnica między kursem *spot* a kursem terminowym polega na tym, że w przypadku drugiego z nich rynki oczekują zmiany kursu w przyszłości. Wśród transakcji terminowych wykorzystujących kursy terminowe najczęściej występują transakcje *forward*, kontrakty *futures* oraz opcje walutowe.

6.5.2. Bilans płatniczy

Jak wspomniano wcześniej, w systemie kursu płynnego kurs walutowy kształtuje się swobodnie na rynku walutowym pod wpływem zmian wielkości podaży i popytu. Z kolei wielkości podaży i popytu na waluty obce są określone przez wielkość transakcji międzynarodowych, czyli transakcji między danym krajem i zagranicą dokonywanych przez firmy, banki, gospodarstwa domowe i sektor rządowy w określonym czasie. Zbiorczym zestawieniem tych transakcji jest bilans płatniczy danego kraju. Przyjrzymy się teraz tym transakcjom, aby lepiej zrozumieć ich wpływ na popyt i podaż walut obcych, a w konsekwencji także na kurs walutowy.

Wszystkie transakcje są ujęte w trzech głównych kategoriach bilansu płatniczego, czyli na trzech tzw. rachunkach – rachunku obrotów bieżących, rachunku obrotów kapitałowych i rachunku obrotów finansowych. Rachunek obrotów bieżących – w skrócie rachunek bieżący (*current account*) – obejmuje wszystkie bieżące płatności otrzymywane przez rezydentów danego kraju z zagranicy z tytułu eksportu dóbr i usług, dochodów z pracy świadczonej za granicą i od kapitału zainwestowanego za granicą oraz innych nieodpłatnych transferów finansowych z zagranicy. Zalicza się tu również wszystkie bieżące płatności dokonane na rzecz

[16] Suma udziałów wynosi w tym przypadku 200%, ponieważ każda transakcja wymiany dotyczy dwóch walut.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zagranicy, takie jak płatności za import dóbr i usług, dochody z pracy świadczonej w kraju i od kapitału zainwestowanego w kraju, oraz innych nieodpłatnych transferów bieżących. Wszystkie transakcje bieżące są dokonywane „tu i teraz", w związku z czym nie prowadzą do powstawania zobowiązań bądź należności na przyszłość.

Rachunek kapitałowy (*capital account*) obejmuje nieodpłatne transfery kapitałowe między krajem i zagranicą, czyli przepływy kapitału nie mające odpowiednika (*counterpart*) w pozostałych strumieniach płatniczych, lecz wpływające na zmiany poziomu oficjalnych aktywów rezerwowych. Do transakcji tych należą na przykład jednostronne i nieodpłatne granty na cele inwestycyjne, dokonywane – przykładowo – w ramach europejskich funduszy strukturalnych, a także umorzenia długu zagranicznego.

Wreszcie rachunek finansowy (*financial account*) ukazuje zmiany w poziomie aktywów i pasywów zagranicznych dokonane przez podmioty krajowe. Do aktywów tych zalicza się inwestycje bezpośrednie, inwestycje portfelowe, pożyczki i kredyty, finansowe instrumenty pochodne, a także oficjalne aktywa rezerwowe, czyli rezerwy walutowe kraju znajdujące się w posiadaniu banku centralnego. Spadek zasobu aktywów zagranicznych oznacza napływ kapitału do kraju, podczas gdy spadek zasobu pasywów zagranicznych świadczy o odpływie kapitału za granicę.

Zauważmy, że każda transakcja zakupu lub sprzedaży zarejestrowana w bilansie płatniczym musi być w jakiś sposób sfinansowana, co oznacza, że każdej transakcji między krajem i zagranicą towarzyszy płatność płynąca w odwrotnym kierunku. Dlatego też suma sald poszczególnych rachunków – po uwzględnieniu zmian, jakie dokonują się w oficjalnych rezerwach – powinna być równa zeru, zgodnie z następującą tożsamością:

$$BCUR + BCAP + BFIN = 0, \tag{6.3}$$

gdzie *BCUR* to saldo rachunku bieżącego, *BCAP* oznacza saldo rachunku kapitałowego, a *BFIN* jest saldem rachunku finansowego. W praktyce powyższa tożsamość rzadko zostaje jednak spełniona. Ze względu na trudności techniczne w pomiarze poszczególnych transakcji i wątpliwości klasyfikacyjne wiele transakcji jest – przynajmniej przejściowo – ujmowanych w kategorii „salda błędów i opuszczeń" (*net errors and omissions*). Dopiero po wyjaśnieniu wątpliwości co do charakteru i natury danej transakcji jest ona ostatecznie reklasyfikowana i „trafia" do właściwej części bilansu płatniczego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 6.5 przedstawia bilans płatniczy Polski za 2014 rok[17]. Interpretacja poszczególnych pozycji *rachunku bieżącego* jest prosta – reprezentują one strumienie płatności z tytułu eksportu i importu dóbr i usług oraz dochodów[18]. Ujemne saldo rachunku bieżącego w wysokości –8,3 mld EUR stanowiło około 1,8% PKB. Deficyt bieżący spowodowany był głównie wysokim ujemnym saldem dochodów pierwotnych –13,2 mld EUR). W obrotach handlowych zarejestrowano niewielki deficyt (–3,3 mld EUR), a bilans usług wykazywał nadwyżkę (8,6 mld EUR). Napływ netto kapitału w ramach rachunku kapitałowego wyniósł 10,0 mld EUR i sfinansował dużą część deficytu bieżącego oraz deficytu na rachunku finansowym (–3,2 mld EUR). Pozostała część niedoboru (1,5 mld EUR) została pokryta częściowo odpływem rezerw oficjalnych (–0,5 mld EUR). W końcu 2014 roku znaczna część transakcji za 2014 rok pozostawała nadal nierozliczona (około 5,0 mld EUR).

Tabela 6.5. Bilans płatniczy Polski za 2014 rok (dane w mln EUR)

Kategoria	2014
A. Rachunek bieżący (*current account*)	**–8303**
A.1. Saldo obrotów towarowych (*balance on goods*)	–3255
A.1.1. Eksport (*goods: exports*)	158619
A.1.2. Import (*goods: imports*)	161874
A.2. Saldo usług (*balance on services*)	8572
A.2.1. Przychody (*services: credit*)	36271
A.2.2. Rozchody (*services: debit*)	27699
A.3. Saldo dochodów pierwotnych (*balance on primary income*)	–13225
A.3.1. Przychody (*primary income: credit*)	11406
A.3.2. Rozchody (*primary income: debit*)	24631
A.4. Saldo dochodów wtórnych (*balance on secondary income*)	–395

[17] Do 2012 roku bilans płatniczy w Polsce sporządzany był na bazie płatności. Począwszy od 2013 roku NBP, zgodnie z zaleceniami Międzynarodowego Funduszu Walutowego, stosuje zmienioną metodologię sporządzania bilansu płatniczego (metodologia BPM5), odnoszącą się w części finansowej do zmian stanu aktywów i pasywów zagranicznych oraz uwzględniającą nieco inną klasyfikację transakcji bieżących. Z tego powodu dane za lata 2013 i 2014 nie są w pełni porównywalne z wcześniejszymi wynikami. Szczegółowy opis zmian można znaleźć m.in. w publikacji *Międzynarodowa pozycja inwestycyjna Polski w 2013 roku*, NBP, Warszawa 2014.

[18] Saldo dochodów pierwotnych stanowi różnicę między przychodami pierwotnymi, obejmującymi wynagrodzenia pracowników polskich zatrudnionych za granicą na okres krótszy niż rok oraz dochody polskich przedsiębiorców z inwestycji dokonanych za granicą, a rozchodami pierwotnymi, do których zalicza się wynagrodzenia nierezydentów zatrudnionych w Polsce na okres krótszy niż rok oraz dochody zagranicznych przedsiębiorców z inwestycji dokonanych w Polsce.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Kategoria	2014
A.4.1. Przychody (*secondary income: credit*)	5942
A.4.2. Rozchody (*secondary income: debit*)	6337
B. Rachunek kapitałowy (*capital account*)	**10034**
B.1. Przychody (*capital account: credit*)	10809
B.2. Rozchody (*capital account: debit*)	775
C. Rachunek finansowy (*financial account*)	**–3234**
C.1. Inwestycje bezpośrednie – aktywa (*direct investment – assets*)	4606
C.1.1. Akcje i inne formy udziałów kapitałowych (*equity and investment fund shares*)	3168
C.1.2. Instrumenty dłużne (*debt instruments*)	1438
C.2. Inwestycje bezpośrednie – pasywa (*direct investment – liabilities*)	12824
C.2.1. Akcje i inne formy udziałów kapitałowych (*equity and investment fund shares*)	7792
C.2.2. Instrumenty dłużne (*debt instruments*)	5032
C.3. Inwestycje portfelowe – aktywa (*portfolio investment – assets*)	4194
C.3.1. Udziałowe papiery wartościowe (*equity securities*)	1780
C.3.2. Dłużne papiery wartościowe (*debt securities*)	2414
C.4. Inwestycje portfelowe – pasywa (*portfolio investment – liabilities*)	2661
C.4.1. Udziałowe papiery wartościowe (*equity securities*)	2297
C.4.2. Dłużne papiery wartościowe (*debt securities*)	364
C.5. Pozostałe inwestycje – aktywa (*other investment – assets*)	3100
C.5.1. NBP (*monetary authorities*)	1
C.5.2. Sektor rządowy (*general government*)	–5
C.5.3. Monetarne Instytucje Finansowe – z wyłączeniem NBP (*MFIs*)	484
C.5.4. Pozostałe sektory (*other sectors*)	2620
C.6. Pozostałe inwestycje – pasywa (*other investment – liabilities*)	–98
C.6.1. NBP (*monetary authorities*)	–1168
C.6.2. Sektor rządowy (*general government*)	1955
C.6.3. Monetarne Instytucje Finansowe – z wyłączeniem NBP (*MFIs*)	976
C.6.4. Pozostałe sektory (*other sectors*)	–1665
C.7. Pochodne instrumenty finansowe (*financial derivatives*)	–16
D. Oficjalne aktywa rezerwowe (*official reserve assets*)	**465**
E. Saldo błędów i opuszczeń (*net errors and omissions*)	**–4965**

Źródło: NBP.

Interpretacja poszczególnych pozycji *rachunku finansowego* jest mniej intuicyjna. W przeciwieństwie do rachunku bieżącego i rachunku kapitałowego, gdzie rejestrowane są strumienie płatności, w rachunku finansowym odnotowuje się zmiany

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w zagranicznych aktywach i pasywach polskich rezydentów. Zgodnie z nową metodologią zapisu, saldo rachunku finansowego w wysokości –3,3 mld EUR odzwierciedla więc *napływ* netto kapitału do Polski. Znak „minus" może okazać się mylący z tego powodu, że wartość ta przedstawia łączne saldo zmian w aktywach i pasywach zagranicznych. Trzeba bowiem pamiętać, że spadek stanu aktywów oznacza napływ kapitału z zagranicy, podobnie jak wzrost stanu pasywów, z kolei wzrost stanu aktywów świadczy o odpływie kapitału, podobnie jak ma to miejsce wówczas, gdy następuje spadek stanu pasywów. W tabeli 6.6 ukazano przykłady najbardziej typowych transakcji finansowych i kapitałowych wraz z ich interpretacją dokonaną na potrzeby bilansu płatniczego.

Tabela 6.6. Wybrane przykładowe transakcje finansowe i sposób ich zapisu w bilansie płatniczym

Nr	Opis transakcji	Rachunek bilansu płatniczego i pozycja w bilansie płatniczym[a]	Wpływ na aktywa/pasywa	Wpływ na bilans płatniczy (napływ/odpływ kapitału z kraju)
1	zakup nieruchomości za granicą przez polskiego obywatela	RF, C.1.1	wzrost stanu aktywów	odpływ kapitału
2	zakup firmy polskiej przez firmę zagraniczną	RF, C.2.1	wzrost stanu pasywów	napływ kapitału
3	sprzedaż przez polski bank własnych obligacji za granicą	RF, C.4.1	wzrost stanu pasywów	napływ kapitału
4	pożyczka udzielona przez firmę zagraniczną polskiej filii tej firmy	RF, C.6.4	wzrost stanu pasywów	napływ kapitału
5	zakup obligacji Skarbu Państwa przez bank zagraniczny	RF, C.4.2	wzrost stanu pasywów	napływ kapitału
6	sprzedaż akcji firmy zagranicznej przez firmę polską	RF, C.1.1	spadek stanu aktywów	napływ kapitału
7	reinwestowanie w kraju zysków firmy polskiej stanowiącej własność firmy zagranicznej	RF, C.1.1	wzrost stanu pasywów	napływ kapitału
8	spłata kredytu zagranicznego przez polski bank	RF, C.6.3	spadek stanu pasywów	odpływ kapitału
9	kredyt organizacji międzynarodowej udzielony polskiemu rządowi	RF, C.6.2	wzrost stanu pasywów	napływ kapitału
10	napływ środków z unijnego Funduszu Spójności	RK, B.1	wzrost stanu aktywów	napływ kapitału

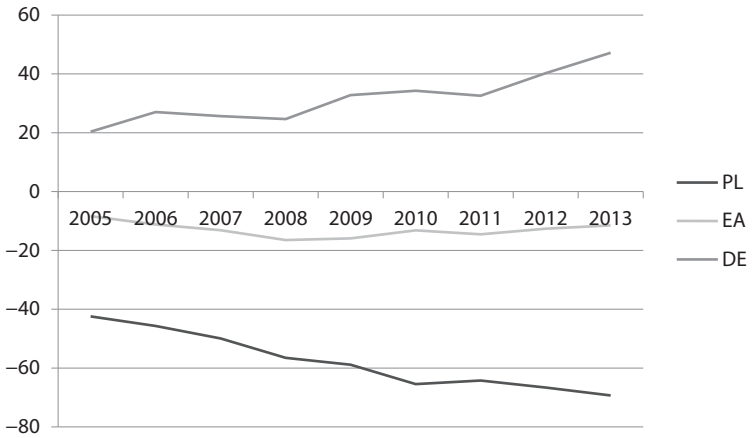
a) RF – rachunek finansowy, RK – rachunek kapitałowy.

Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Gdy kraj wykazuje systematyczny deficyt obrotów bieżących, oznacza to, że jego rezydenci więcej wydają niż zarabiają, co powoduje systematyczne narastanie zobowiązań finansowych wobec zagranicy. Mogą one przyjmować postać napływu inwestycji bezpośrednich, zwiększania zasobu krajowych papierów wartościowych w posiadaniu zagranicznych inwestorów lub bezpośredniego wzrostu zadłużenia zagranicznego rezydentów krajowych z tytułu zaciągniętych za granicą kredytów i pożyczek. Gdy skumulowana wartość tych zobowiązań wobec zagranicy (zadłużenie zagraniczne) jest większa niż skumulowana wartość posiadanych aktywów zagranicznych, mówi się, że kraj posiada ujemną międzynarodową pozycję inwestycyjną netto. Jest ona zestawieniem, które przedstawia stan zagranicznych aktywów i pasywów podmiotów krajowych na koniec danego roku. Różnica między wielkością aktywów i pasywów zagranicznych stanowi więc międzynarodową pozycję inwestycyjną netto (*net foreign asset position* – NFAP) oraz informuje o tym, czy dany kraj jest wierzycielem, czy też dłużnikiem netto w stosunku do zagranicy (NBP 2014).

Wykres 6.8. Międzynarodowa pozycja inwestycyjna netto Polski, Niemiec i strefy euro w latach 2005–2013 (dane w % PKB)



Źródło: opracowanie własne na podstawie danych Eurostatu.

Jak pokazuje wykres 6.8, międzynarodowa pozycja inwestycyjna netto Polski była od wielu lat ujemna, co jest charakterystyczne w przypadku kraju, który szybko się rozwija, doganiając kraje rozwinięte, i który w związku z tym jest importerem netto kapitału. W latach 2005–2013 relacja NFAP do PKB Polski wzrosła z 42,4%

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

do 69,3% PKB. Z kolei NFAP Niemiec była dodatnia i wzrosła z 20,3% do 47,2% PKB. NFAP całej strefy euro pozostawała zaś nieznacznie ujemna i utrzymywała się w tym samym czasie w granicach 10–15% PKB. Należy zauważyć, że wielkość NFAP zależy nie tylko od zmian w poziomie aktywów i pasywów zagranicznych, ale także od zmian ich cen oraz zmian kursów walutowych, które określają aktualną wartość rynkową poszczególnych aktywów i pasywów.

6.5.3. Równowaga na rynku walutowym

W warunkach pełnej swobody transgranicznego przepływu kapitału rynki walutowe w poszczególnych krajach są ze sobą wzajemnie zintegrowane. Parytet stóp procentowych sprawia wówczas, że stopy przychodu z inwestycji w podobne instrumenty finansowe denominowane w różnych walutach wyrównują się w skali międzynarodowej po uwzględnieniu oczekiwanych zmian kursu (zob. rozdział 5.10.1). Przypomnijmy, że równanie niezabezpieczonego parytetu stóp procentowych (UIP) przyjmuje następującą postać (równanie 5.72):

$$i = i^* + \frac{e^p - e}{e}. \tag{5.67}$$

Wówczas kurs e, w przypadku którego równanie parytetu zostaje spełnione, jest kursem równowagi.

Przedstawimy obecnie działanie parytetu stóp procentowych na podstawie porównania oprocentowania depozytów terminowych o jednakowym terminie zapadalności w euro i w PLN. Załóżmy, że inwestor zagraniczny posiada wolne środki, na przykład o wartości 10 mln EUR, które chce zainwestować na okres jednego roku. Najprościej jest ulokować pieniądze na lokacie terminowej w banku. Przyjmijmy więc, że dany inwestor ma do wyboru lokatę terminową w euro, która przyniesie mu odsetki w wysokości 2% rocznie, lub lokatę terminową w złotych, która zapewni mu oprocentowanie w wysokości 6% rocznie (ponieważ stopy procentowe w Polsce są wyższe niż w strefie euro)[19, 20]. Terminowe lokaty bankowe nie

[19] W ostatnich latach stopy procentowe zarówno w strefie euro, jak i w Polsce były znacznie niższe. Na przykład w końcu czerwca 2015 roku średnie oprocentowanie rocznego terminowego depozytu w euro w bankach strefy euro wynosiło 0,6% i w ciągu roku zmniejszyło się o 0,1 punktu procentowego, z kolei w przypadku rocznego terminowego depozytu w złotych w bankach w Polsce oprocentowanie wynosiło 2,1% i w ciągu roku obniżyło się o 0,7 punktu procentowego (na podstawie danych EBC i NBP).

[20] W warunkach pełnej wymienialności obu walut zarówno depozyty w euro, jak i depozyty w złotych mogą być bez problemów założone w tym samym banku z siedzibą w dowolnym kraju UE.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

są jednak szczególnie atrakcyjnym instrumentem finansowym, ponieważ nie mają one charakteru płynnego (nie mogą być przedmiotem obrotu na rynku finansowym), a rozwiązanie takiej lokaty przed terminem zapadalności wiąże się z reguły z dodatkowym kosztem. Dlatego też inwestorzy wolą kupować papiery wartościowe, które można upłynnić w każdej chwili na rynku wtórnym. Jakie powinno być oprocentowanie tych papierów wartościowych denominowanych w polskich złotych, aby opłacało się w nie zainwestować, uwzględniając ryzyko zmiany kursu złotego wobec euro?

Zgodnie z równaniem parytetu, kapitał będzie napływał (odpływał) z zagranicy do kraju aż do momentu, gdy stopy przychodu z obu rodzajów inwestycji, po uwzględnieniu oczekiwanej zmiany kursu, wyrównają się. Wynikająca z parytetu stóp procentowych równowaga na rynku walutowym została przedstawiona na rysunku 6.5, gdzie na osi poziomej odłożono oczekiwaną stopę przychodu z instrumentu denominowanego w złotych, a na osi pionowej kurs bieżący (*spot*) PLN/EUR e^s. Zależność między kursem bieżącym a oczekiwaną stopą przychodu ilustruje krzywa opadająca *D*. Przy danym oczekiwanym kursie e^p za rok, im wyższy stanie się bieżący kurs walutowy e^s, tym niższa może być stopa procentowa na instrumentach złotowych ze względu na to, że tym niższa będzie również oczekiwana deprecjacja kursu złotego wobec euro. Załóżmy, że kursem oczekiwanym za dany rok e^s jest kurs na poziomie $e^p = 4{,}20$ PLN/EUR. Jeśli bieżący kurs wynosi $e^s = 4{,}00$ PLN/EUR, oznacza to, że rynki finansowe oczekują deprecjacji złotego o 5% ($\frac{4{,}20-4{,}00}{4{,}00}=\frac{0{,}20}{4{,}00}=0{,}05$) w ciągu roku. Równanie parytetu wymaga, aby stopa procentowa w Polsce była równa stopie procentowej w strefie euro powiększonej o stopę oczekiwanej deprecjacji złotego, czyli:

$$i = i^* + \frac{e^p - e^s}{e^s} = 0{,}02 + 0{,}05 = 0{,}07.$$

Punktem równowagi na rynku walutowym stanie się punkt przecięcia funkcji parytetu *D* i linii pionowej wyznaczającej stopę procentową w Polsce na poziomie 7% – jest to punkt *A*. Wówczas kursem równowagi będzie kurs $e = 4{,}00$ PLN/EUR.

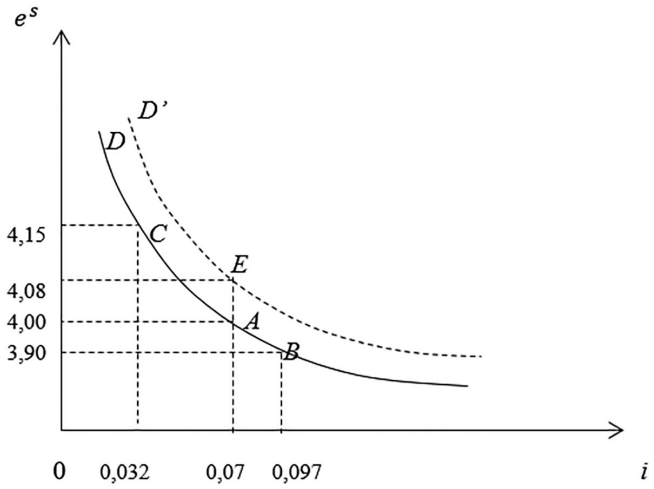
Gdy stopa procentowa w Polsce stanie się niższa niż 7%, lewa strona równania parytetu będzie niższa od prawej, a kapitał będzie odpływał z Polski za granicę. Stały odpływ kapitału oznacza spadek podaży pieniądza, co spowoduje, że rynkowe stopy procentowe wzrosną do poziomu 7%. Gdy stopa procentowa w Polsce

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

stanie się z kolei wyższa niż 7%, kapitał będzie napływał z zagranicy. Stały napływ kapitału oznacza wzrost podaży pieniądza, co sprawi, że rynkowe stopy procentowe w Polsce spadną do poziomu 7%.

Załóżmy, że z jakiegoś powodu dochodzi do umocnienia się kursu bieżącego z 4,00 PLN/EUR do, na przykład, 3,90 PLN/EUR. Powoduje to wzrost wymaganej parytetem stopy procentowej równowagi w złotych do 9,7%, ponieważ oczekiwana deprecjacja złotego w ciągu roku zwiększa się z 5% do 7,7%. Punktem równowagi stanie się teraz punkt *B*. Z kolei osłabienie kursu bieżącego, na przykład do 4,15 PLN/EUR, obniża wymaganą parytetem stopę procentową równowagi w złotych do 3,2%, ponieważ spada oczekiwana deprecjacja złotego z 5% do 1,2%. Równowaga ustali się w tym przypadku w punkcie *C*. Ta sama zależność sprawia, że wzrost krajowej stopy procentowej przy danym oczekiwanym kursie walutowym za dany rok oznacza umocnienie bieżącego kursu złotego wobec euro, a spadek krajowej stopy procentowej świadczy o osłabieniu bieżącego kursu.

Rysunek 6.5. Równowaga na rynku walutowym dla danej oczekiwanej zmiany kursu



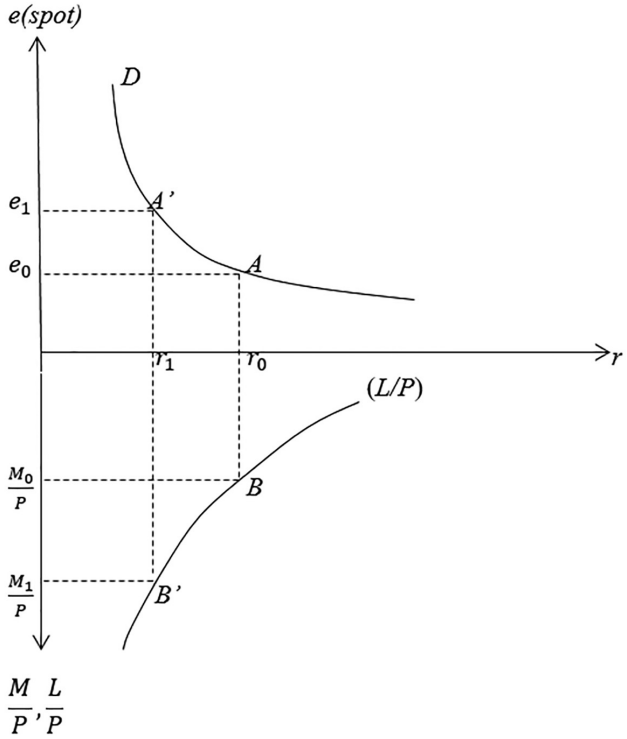
Źródło: opracowanie własne na podstawie Krugman i Obstfeld (1992).

Punkt *A* przestanie być punktem równowagi, gdy dochodzi do zmiany zagranicznej stopy procentowej. Na przykład, gdy stopa zagraniczna i^* wzrośnie z 2% do 4%, to przy danym oczekiwanym kursie walutowym za rok krzywa *D* przesunie się do położenia *D'* i nastąpi natychmiastowa deprecjacja złotego wobec euro z 4,00 PLN/EUR do 4,08 PLN/EUR. Oczekiwana w ciągu roku deprecjacja złotego

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

spadnie z 5% do około 3% ($\frac{4,20-4,08}{4,08}=\frac{0,12}{4,08}\simeq 0,03$) i nowym punktem równowagi stanie się punkt *E*, spełniający warunek parytetu stóp procentowych. Dokładnie taki sam przebieg ma dostosowanie w przypadku spadku oczekiwanego kursu PLN/EUR.

Rysunek 6.6. Równowaga na rynku pieniądza krajowego i na rynku walutowym



Źródło: opracowanie własne na podstawie Krugman i Obstfeld (1992).

Pamiętamy, że, zgodnie z twierdzeniem o „trójkącie niemożności" (zob. rozdział 5.10.2), w warunkach pełnej swobody przepływów kapitału krajowa stopa procentowa i kurs walutowy są ze sobą powiązane i nie mogą być ustalane niezależnie od siebie. Ponieważ bank centralny reguluje podaż pieniądza w gospodarce za pomocą krótkoterminowej stopy procentowej, zmiany podaży pieniądza wywierają wpływ – za pośrednictwem stóp procentowych – na poziom bieżącego kuru walutowego. Związek ten ukazano na rysunku 6.6, którego górna część jest

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

repliką rysunku 6.5, a dolna przedstawia zależność między popytem na pieniądz realny (L/P) a realną stopą procentową r w krótkim okresie czasu. Jeśli realna stopa oprocentowania depozytów złotowych wynosi r_0, to punkt B wyznacza równowagę między podażą i popytem na rynku pieniądza krajowego, podczas gdy punkt A opisuje walutowy kurs równowagi na rynku pieniądza zagranicznego.

Wzrost realnej podaży pieniądza krajowego (PLN) z $\left(\frac{M_0}{P}\right)$ do $\left(\frac{M_1}{P}\right)$ wywołuje spadek realnych stóp procentowych z r_0 do r_1 i natychmiastową deprecjację złotego z e_0 do e_1. Deprecjacja kursu bieżącego jest spowodowana – tak jak poprzednio – spadkiem oczekiwanej skali deprecjacji złotego wobec euro (spadek oczekiwanej aprecjacji euro wobec złotego) w perspektywie najbliższego roku. Nowym punktem równowagi na rynku pieniądza zagranicznego staje się punkt A', z kolei na rynku pieniądza krajowego punkt B'.

6.5.4. Krótkookresowa zmienność kursu walutowego – zjawisko „przestrzelenia" (*exchange rate overshooting*)*

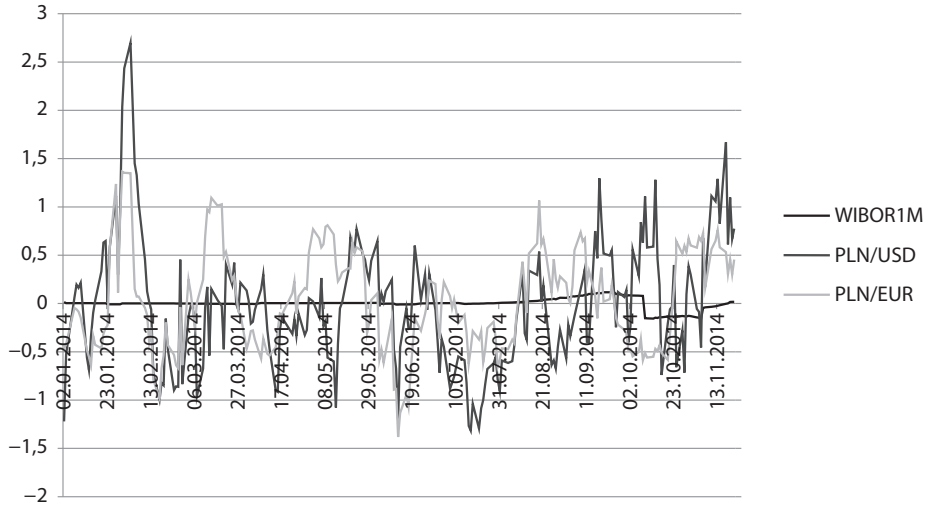
Twierdzenie o „trójkącie niemożności" i rysunek 6.6 pokazują, że w gospodarce otwartej na przepływy kapitałowe istnieje powiązanie między krajową stopą procentową i kursem walutowym. Równanie parytetu stóp procentowych (5.67) wskazuje, że warunkiem równowagi płatniczej jest równość krajowej stopy procentowej i zagranicznej stopy procentowej po uwzględnieniu przewidywanej zmiany kursu złotego. Jeśli nastąpi więc zmiana krajowych lub zagranicznych stóp procentowych, pojawią się przepływy kapitału, które – w systemie płynnego kursu – spowodują odpowiednie zmiany bieżącego kursu walutowego, a to przy danym oczekiwanym kursie wywoła odpowiednie zmiany *różnicy* między kursem bieżącym a kursem oczekiwanym. Gdy krajowa stopa procentowa będzie wyższa (lub niższa) od poziomu zapewniającego równowagę, wystąpi napływ (lub odpływ) kapitału, który spowoduje aprecjację (lub deprecjację) waluty krajowej. Z równania parytetu wynika, że zmiana kursu powinna być zasadniczo równa powstałej różnicy między krajową i zagraniczną stopą procentową.

Obserwacje krótkookresowych wahań kursu nie potwierdzają jednak tej prawidłowości. Okazuje się, że zmiany kursu w krótkim okresie są znacznie silniejsze niż zmiany stóp procentowych. Zwłaszcza w warunkach wysokiej mobilności kapitału nawet niewielkie zmiany stóp procentowych powodują znaczne wahania

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kursu walutowego. Potwierdzają to, między innymi, polskie statystyki. Na przykład w okresie od 1 stycznia do 31 grudnia 2014 roku różnica miedzy najwyższym i najniższym kursem złotego wobec euro (4,188 i 4,288 PLN/EUR) wyniosła 2,4% średniego kursu, podczas gdy różnica miedzy najwyższą i najniższą wartością dziennych notowań WIBOR 3M osiągnęła wartość zaledwie 0,78 punktu procentowego. Wykres 6.9 pokazuje, że zmienność kursu PLN/EUR i PLN/USD była w 2014 roku wielokrotnie wyższa niż zmienność rynkowej stopy 3-miesięcznego WIBOR-u.

Wykres 6.9. Zmienność kursu PLN/EUR i PLN/USD na tle zmienności rynkowej stopy 3-miesięcznego WIBOR w okresie od 1 stycznia do 15 listopada 2014 roku (dane w %[a])



a) Wartości odchyleń dziennych notowań od 90-dniowych średnich ruchomych.

Źródło: opracowanie własne na podstawie danych NBP (kursy walutowe) i portalu money.pl (stawki WIBOR).

Zjawisko to wyjaśnił w 1976 roku niemiecki ekonomista Rudiger Dornbusch. Wykazał on, że w gospodarce o płynnym kursie walutowym i sztywnych cenach zmiany kursu wykazują w krótkim okresie tendencję do „przestrzeliwania" (*overshooting*) w reakcji na zmiany stóp procentowych. Zjawisko „przestrzelenia" oznacza, że w krótkim okresie reakcja kursu na impuls pieniężny jest silniejsza niż reakcja długookresowa. Główną przyczyną takiego stanu rzeczy jest sztywność cen (*price rigidity*), która powoduje, że zmiany stóp procentowych wpływają z opóźnieniem na zmiany cen, a jednocześnie mają natychmiastowy wpływ na przepływy kapitału oraz na kurs walutowy (Dornbusch 1976). Teoria Dornbuscha wyjaśnia, dlaczego

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

niewielkie zmiany w polityce pieniężnej wywołują w krótkim okresie znacznie silniejsze zmiany kursu walutowego.

Załóżmy, że mamy do czynienia z gospodarką, w której ceny są sztywne (*sticky prices*) w krótkim okresie, lecz w długim pozostają giętkie (*flexible*), i gdzie obowiązuje system płynnego kursu. Załóżmy następnie, że w gospodarce tej bank centralny zwiększa *trwale* podaż pieniądza o 5%. Zgodnie z twierdzeniem o neutralności pieniądza, w długim okresie doprowadzi to do wzrostu poziomu cen o około 5%, a zgodnie z parytetem siły nabywczej wywoła w konsekwencji (relatywną) deprecjację kursu waluty krajowej także o około 5%. Lecz w krótkim okresie, bezpośrednio po decyzji banku centralnego, siła krajowej waluty spadnie o więcej niż 5%. Stanie się tak dlatego, że wzrost nominalnej podaży pieniądza w warunkach sztywnych cen oznacza wzrost realnej podaży pieniądza, co powoduje spadek rynkowych stóp procentowych. To z kolei narusza równanie parytetu stóp, czego skutkiem jest natychmiastowy odpływ kapitału za granicę oraz nagła deprecjacja waluty krajowej (skala nagłej deprecjacji będzie tym większa, im większa jest transgraniczna mobilność kapitału). Wraz z upływem czasu poziom cen jednak stopniowo wzrasta. Przekłada się to na systematyczny spadek realnej podaży pieniądza oraz stopniowy wzrost stóp procentowych. Spadek *różnicy* między krajową i zagraniczną stopą procentową zmniejsza *oczekiwaną* aprecjację. Ostatecznie stopy krajowe powracają do poziomu stóp zagranicznych, a kurs walutowy stabilizuje się (zob. ramka 6.6).

Ramka 6.6. Mechanizm „przestrzelenia" kursu walutowego według Dornbuscha

Niemiecki ekonomista Rudiger Dornbusch, profesor w Massachussetts Institute of Technology (MIT), wyjaśnił, dlaczego kurs walutowy reaguje szybciej na zmiany oczekiwań niż stopy procentowe. Rozważmy następujący prosty przykład liczbowy. Przyjmijmy, że w punkcie wyjścia rynkowe stopy procentowe w Polsce oraz w strefie euro są sobie równe i wynoszą 5%, a oczekiwany za rok kurs walutowy jest równy kursowi bieżącemu i wynosi $e^p = e = 4{,}00$ PLN/EUR. Równanie parytetu zostaje zatem spełnione. Załóżmy następnie, że polski bank centralny zwiększa podaż pieniądza na przykład o 5% i w rezultacie krajowe stopy procentowe spadają do 4%. W warunkach wysokiej elastyczności przepływu kapitału względem stóp procentowych i przy założeniu, że giętkość kursu względem przepływu kapitału jest równa jedności, natychmiastowa deprecjacja kursu wywołana odpływem kapitału wyniesie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

więcej niż 5%, a zatem więcej niż oczekiwana deprecjacja w długim okresie wywołana oczekiwanym wzrostem cen. Natychmiastowe osłabienie kursu będzie trwało aż do momentu, kiedy skala *oczekiwanej aprecjacji* zapewni spełnienie równania parytetu dla nowego poziomu krajowej stopy procentowej[21]. Stopniowy wzrost cen obniża realną wielkość podaży pieniądza i wywołuje stopniowy wzrost stóp procentowych. Spadek różnicy między krajową i zagraniczną stopą procentową zmniejsza oczekiwaną aprecjację. Ostatecznie stopy krajowe powracają do poziomu stóp zagranicznych, a kurs walutowy stabilizuje się. Cały ten proces dostosowania kursu można podzielić na etapy tak, jak przedstawiono to w tabeli 6.7.

Tabela 6.7. Etapy procesu dostosowania kursu walutowego do impulsu polityki pieniężnej

Etapy dostosowania	Parytet stóp procentowych (bez pokrycia) $i = i^* + \frac{e^p - e}{e}$	Wartość liczbowa parytetu
1. Stan początkowej równowagi	$0{,}05 = 0{,}05 + \frac{4{,}00 - 4{,}00}{4{,}00}$	$0{,}05 = 0{,}05$
2. Wzrost podaży pieniądza i spadek stóp procentowych: naruszenie parytetu	$0{,}04 < 0{,}05 + \frac{4{,}00 - 4{,}00}{4{,}00}$	$0{,}04 < 0{,}05$
3. Pojawienie się oczekiwanej deprecjacji i „przestrzelenie" kursu bieżącego	$0{,}04 > 0{,}05 + \frac{4{,}2 - 4{,}5}{4{,}5}$	$0{,}04 > -0{,}0167$
4. Przywrócenie parytetu i pojawienie się oczekiwanej aprecjacji	$0{,}04 = 0{,}05 + \frac{4{,}2 - 4{,}242}{4{,}242}$	$0{,}04 = 0{,}04$
5. Stabilizacja kursu	$0{,}05 = 0{,}05 + \frac{4{,}2 - 4{,}2}{4{,}2}$	$0{,}05 = 0{,}05$

Źródło: opracowanie własne.

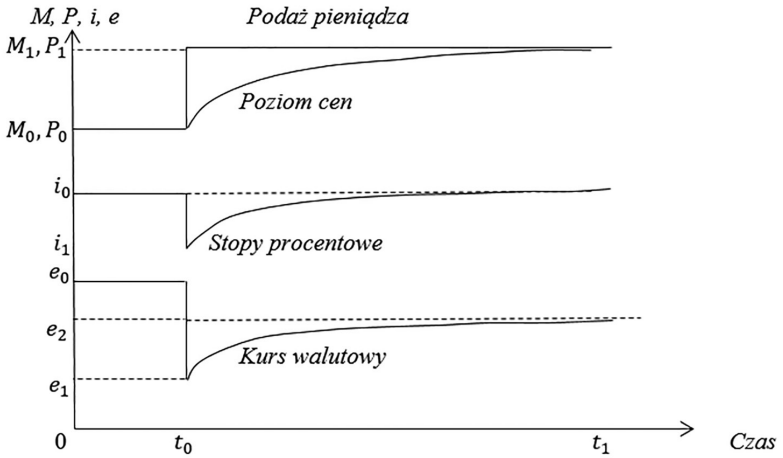
Reakcje stóp procentowych, cen oraz kursu walutowego na jednorazowe i trwałe zwiększenie podaży pieniądza można też przedstawić przy pomocy rysunku 6.7. Pokazuje on, że w warunkach braku natychmiastowego dostosowania cen kurs walutowy ulega w krótkim okresie silniejszej deprecjacji (z e_0 do e_1) niż docelowo w długim okresie (e_2). Wzrost podaży pieniądza z M_0 do M_1 w czasie t_0 wywołuje natychmiastowy spadek stóp procentowych z i_0 do i_1 oraz deprecjację kursu z e_0 do e_1. Z kolei poziom cen dostosowuje się stopniowo – z P_0 w czasie t_0 do P_1 w czasie t_1. W miarę dostosowywania się cen wzrastają też stopniowo stopy procentowe

[21] Zgodnie z równaniem parytetu, dodajemy do siebie procentową zmianę kursu i zmianę stóp procentowych w punktach procentowych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(z powrotem do poziomu i_0)[22], a kurs walutowy ulega stopniowej aprecjacji (do poziomu e_2). Gdyby ceny były w pełni giętkie i dostosowały się natychmiast do zwiększonej podaży pieniądza, kurs walutowy od razu obniżyłby się do swojej długookresowej wartości e_2 i nie wystąpiłoby zjawisko „przestrzelenia".

Rysunek 6.7. „Przestrzelenie" kursu walutowego w reakcji na impuls polityki pieniężnej według Dornbuscha



Uwaga: kurs walutowy wyrażony w EUR/PLN.

Źródło: Dornbusch (1976), Krugman i Obstfeld (1990, s. 374–375), Bénassy-Queré i in. (2011, s. 298–299).

6.6. Czynniki określające kurs walutowy

6.6.1. Kurs walutowy i strumienie handlu: podejście tradycyjne (strumieniowe) i warunek Marshalla-Lernera-Robinson

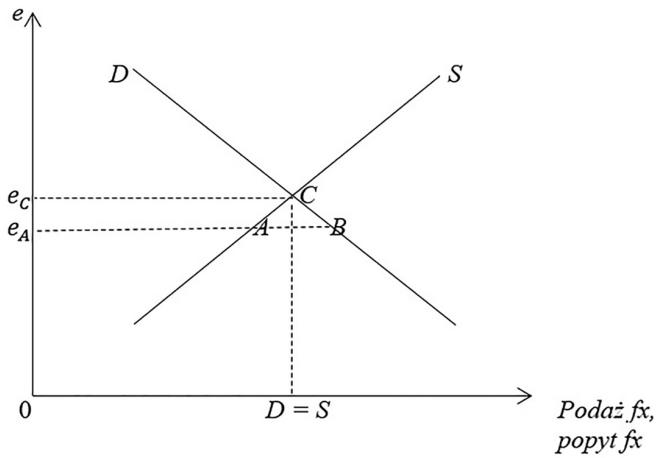
Przez niemal cały XX wiek waluty narodowe większości krajów nie podlegały swobodnej wymianie, a transgraniczna mobilność kapitałów była poważnie

[22] Proces dostosowywania się nominalnych stóp procentowych do zmian oczekiwanego poziomu cen nazywany jest często efektem Fishera, od nazwiska amerykańskiego ekonomisty Irvinga Fishera, który szczegółowo opisał to zjawisko. W długim okresie, przy innych założeniach stałych, wzrost oczekiwanej inflacji w danym kraju wywołuje stopniowy wzrost nominalnych stóp procentowych, wobec czego realna stopa procentowa nie ulega zmianie (Fisher 1930).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ograniczona. W takich warunkach poziom kursu walutowego zależał w znacznym stopniu od strumieni płatności bieżących, w tym przede wszystkim płatności związanych z handlem zagranicznym. Relacja między wielkością eksportu oraz importu dóbr i usług określała więc relacje między podażą a popytem na waluty obce. W najprostszym ujęciu popyt na waluty obce był w przeważającej mierze określony przez wielkość importu, a podaż walut obcych przez wielkość eksportu danego kraju. Kurs walutowy zależał więc głównie od relacji występujących na rachunku bieżącym bilansu płatniczego, podczas gdy rachunek kapitałowy i finansowy miał zdecydowanie mniejsze znaczenie.

Rysunek 6.8. Kurs walutowy jako funkcja podaży i popytu na walutę obcą (*fx*)



Źródło: opracowanie własne.

Na rysunku 6.8 ukazano relację między kursem walutowym *e* a podażą (*S*) i popytem (*D*) na rynku walutowym. Krzywa *S* jest rosnąca, ponieważ wzrost kursu prowadzi do wzrostu eksportu ze względu na to, że rośnie wówczas jego opłacalność, krzywa *D* jest zaś opadająca, ponieważ wzrost kursu prowadzi do spadku importu gdyż rośnie wówczas jego cena. W systemie kursu płynnego egzogeniczne zmiany podaży i popytu na walutę obcą wywołują zmiany kursu walutowego. Z kolei w systemie kursu stałego to ustalony poziom kursu określa wielkość podaży i popytu na waluty obce. Na przykład przy kursie e_A popyt jest większy od podaży, co oznacza deficyt w bilansie płatniczym w wysokości równej odcinkowi *AB*. Dewaluacja kursu do poziomu e_C przywraca równowagę między podażą i popytem na walutę obcą ($D = S$).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ponieważ wielkości eksportu oraz importu – jako zmienne realne – zmieniają się stosunkowo powoli, w krótkim okresie położenie krzywych D i S nie ulega większym zmianom. W rezultacie kurs równowagi był względnie stabilny. Tłumaczy to popularność stosowania stałych kursów walutowych w tamtym okresie.

Związek między poziomem kursu a poziomem eksportu i importu stał się przesłanką wykorzystania polityki kursowej w celu uzyskania pożądanych zmian w bilansie płatniczym. Kraje dokonywały dewaluacji kursu swojej waluty, aby pobudzić eksport i ograniczyć lub zlikwidować deficyt płatniczy bądź też uzyskać nadwyżkę[23]. Nie zawsze działania te okazywały się jednak skuteczne. Dewaluacja wpływa na wzrost kosztów importu ponoszonych przez odbiorców krajowych i przekłada się na wzrost eksportu dzięki obniżeniu cen eksportowych, lecz ostateczny wynik w ramach bilansu płatniczego zależy zarówno od elastyczności popytu zagranicznego na eksport kraju, jak i od elastyczności popytu krajowego na import z zagranicy. Zgodnie ze znanym twierdzeniem Marshalla-Lernera-Robinson (MLR), dewaluacja kursu przynosi poprawę bilansu płatniczego, gdy suma elastyczności popytu na eksport i import jest większa od jedności. Możemy więc zapisać, że poprawa bilansu płatniczego zachodzi wówczas, aby spełniony jest następujący warunek:

$$\eta_E^D + \eta_M^D > 1, \tag{6.4}$$

gdzie η_E^D, η_M^D oznaczają kolejno elastyczność cenową popytu zagranicznego na eksport z danego kraju oraz elastyczność cenową popytu danego kraju na import. Wyprowadzenie warunku nierówności (6.4) przedstawione jest w ramce 6.7.

Warunek Marshalla-Lernera-Robinson jest jednym z najbardziej znanych twierdzeń ekonomicznych odnoszących się do kursu walutowego[24]. Wyjaśnimy go najpierw za pomocą prostej ilustracji zachowania się popytu zagranicznego i sprzedaży eksportowej w reakcji na zmiany cen eksportowych. Na rysunku 6.9 krzywa popytu *a* przyjmuje kształt standardowy – im niższa staje się cena, tym wyższy towarzyszy jej popyt. W początkowym punkcie A wartość sprzedaży eksportowej wynosi $Y_0 = p_0 y_0$. Załóżmy teraz, że na skutek dewaluacji kursu waluty krajowej następuje spadek zagranicznych cen eksportowych do poziomu p_1, co pociąga za sobą wzrost popytu i sprzedaży do poziomu y_1. Wartość sprzedaży eksportowej

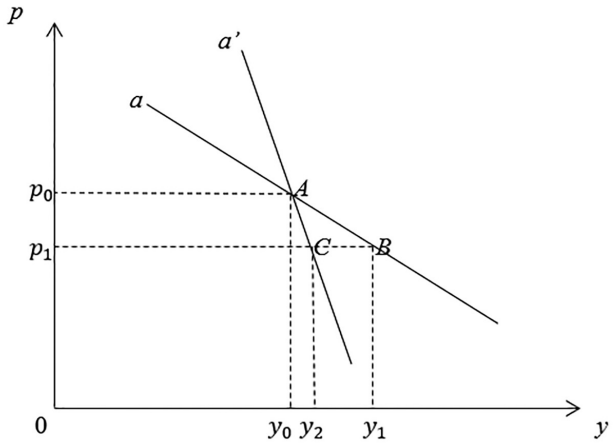
[23] Niekiedy, choć znacznie rzadziej, dokonywano też rewaluacji kursu.

[24] Warunek Marshalla-Lernera-Robinson (nazywany też po prostu warunkiem Marshalla-Lernera) został stopniowo rozwinięty niezależnie przez troje ekonomistów: Alfreda Marshalla, Abbę Lernera i Joan Robinson, w latach 20.i 30. ubiegłego wieku (Marshall 1924, Lerner 1949, Robinson 1937).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w punkcie B wynosi $Y_1 = p_1 y_1$. Nasze zadanie polega na określeniu, kiedy sprzedaż eksportowa wyrażona w *walucie obcej* wzrośnie pod wpływem dewaluacji (zakładamy, że krajowa podaż towarów eksportowanych jest doskonale elastyczna).

Rysunek 6.9. Zmiany wartości sprzedaży eksportowej w reakcji na zmiany cen



Źródło: opracowanie własne.

Wzrost wartości sprzedaży eksportowej wymaga, aby $Y_1 > Y_0$. Inaczej mówiąc, pole prostokąta $By_1 0p_1$ powinno być większe od pola prostokąta $Ay_0 0p_0$. Warunek ten możemy zapisać następująco (subskrypty w celu uproszczenia pomijamy):

$$(p - dp)(y + dy) > py, \tag{6.5}$$

gdzie p, y stanowią początkowe wartości cen i popytu, a dp, dy to zmiany cen i popytu wywołane dewaluacją kursu. Wykonując mnożenie po lewej stronie, a następnie dzieląc obustronnie przez py i porządkując, otrzymujemy:

$$\frac{dy}{y} - \frac{dp}{p} - \left(\frac{dp}{p}\right)\left(\frac{dy}{y}\right) > 0. \tag{6.6}$$

W przypadku niskich wartości względnych zmian cen i popytu możemy przyjąć, że $(dp/p)(dy/p) \cong 0$. Wówczas nierówność (6.6) przyjmuje postać:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

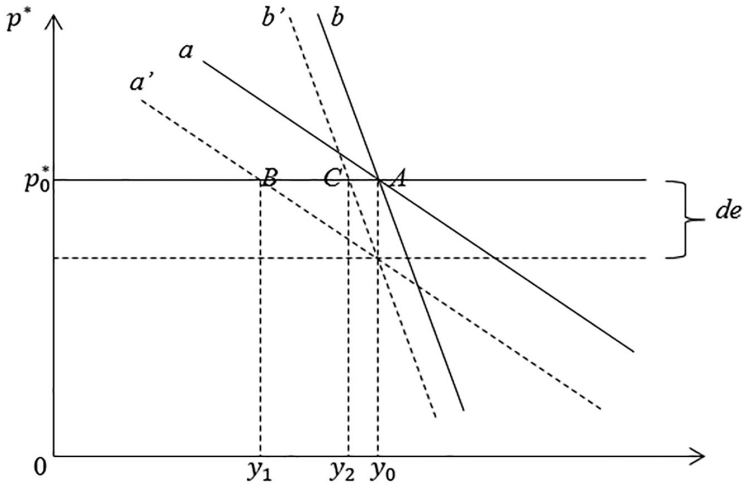
$$\frac{\frac{dy}{y}}{\frac{dp}{p}} > 1,$$

czyli:

$$\eta^D > 1, \tag{6.7}$$

gdzie η^D jest elastycznością cenową popytu na eksport. Nierówność (6.7) wskazuje, że wzrost *wartości* sprzedaży w walucie może nastąpić wówczas, gdy elastyczność popytu jest większa od jedności. Można więc uznać ten wynik za zgodny z intuicją. Im bardziej płaska stanie się bowiem krzywa popytu przedstawiona na rysunku 6.9, tym większa będzie wartość sprzedaży po obniżce cen, z kolei im bardziej stroma staje się krzywa popytu, tym mniejsza jest wartość sprzedaży po obniżce cen. Na przykład, gdy popyt okazuje się mało elastyczny (krzywa *a'*), wzrost ilości sprzedaży jest niewielki (z y_1 do y_2), co przy danej stopie spadku cen (przy danej stopie dewaluacji) oznacza spadek wartości sprzedaży (ponieważ pole C$y_2$0p_1 jest mniejsze od pola A$y_0$0p_0).

Rysunek 6.10. Wpływ dewaluacji kursu na import kraju



Źródło: opracowanie własne.

Zauważmy, że łączny wpływ dewaluacji wielkości sprzedaży eksportowej stanowi sumę oddziaływania dwóch efektów – efektu ilościowego (*volume effect*),

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

który wpływa na zwiększenie wartości eksportu, oraz efektu cenowego (*price effect*), który wywołuje spadek wartości eksportu. Gdy popyt na eksport jest elastyczny, efekt ilościowy przeważa nad efektem cenowym i wartość eksportu rośnie. Gdy popyt staje się natomiast nieelastyczny, efekt cenowy przeważna nad efektem ilościowym i wartość eksportu maleje. Graniczną wartością jest elastyczność popytu równa jedności.

W celu zachowania kompletności analizy trzeba jeszcze zbadać skutki dewaluacji w zakresie importu. Wyjaśnienie w tym przypadku staje się prostsze niż w przypadku eksportu. Na rysunku 6.10 przedstawiono krajowy popyt na import jako funkcję ceny zagranicznej p^* przeliczonej na walutę krajową po kursie e. Zakładamy, że zagraniczna podaż towarów importowanych jest doskonale elastyczna. Dewaluacja kursu o wielkość de, przy niezmienionej cenie zagranicznej p_0^*, podnosi cenę krajową importu o de. Krzywa popytu przesuwa się w dół o wielkość dewaluacji – z położenia a do położenia a' – i popyt spada z y_0 do y_1. Wartość importu w walucie obcej maleje, ponieważ przy tej samej zagranicznej cenie zakupu p_0^* popyt na import jest niższy ($y_0 > y_1$). Spadek wartości importu staje się równy różnicy $p_0^* y_0$ i $p_0^* y_1$, czyli między powierzchnią prostokąta A $y_0 0 p_0^*$ a powierzchnią prostokąta B $y_1 0 p_0^*$.

Podobnie jak w przypadku eksportu, im bardziej stroma jest krzywa popytu na import, czyli im mniej elastyczny staje się popyt względem ceny zakupu, tym mniejszy następuje spadek popytu oraz spadek wartości importu. Na przykład, gdy krzywa popytu ma kształt b, dewaluacja kursu o tę samą wielkość de przesuwa krzywą popytu do położenia b', co odpowiada spadkowi popytu do poziomu y_2. Spadek wartości importu jest więc mniejszy niż poprzednio (bo $y_1 < y_2$). Zauważmy, że w przypadku importu występuje tylko efekt ilościowy, ponieważ ceny zagraniczne (przy przyjętych założeniach) nie ulegają zmianie.

Ramka 6.7. Warunek Marshalla-Lernera-Robinson

Załóżmy, że transgraniczne przepływy kapitału podlegają ograniczeniom. Wówczas saldo bilansu płatniczego B można przedstawić jako różnicę między eksportem X a importem M dóbr i usług:

$$B = X - eM, \tag{6.8}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gdzie e jest *stałym* kursem waluty obcej wyrażonym w walucie krajowej (dewaluacja kursu oznacza wzrost e). Załóżmy też w celu uproszczenia wywodu, że bilans płatniczy jest w punkcie wyjścia zrównoważony, czyli $B = 0$. W jakich warunkach dewaluacja kursu przyniesie poprawę salda bilansu płatniczego? Różniczkując wzór (6.8), otrzymujemy:

$$\frac{\delta B}{\delta e} = \frac{\delta X}{\delta e} - \left(\frac{\delta M}{\delta e} e + \frac{\delta e}{\delta e} M \right) = \frac{\delta X}{\delta e} - \frac{\delta M}{\delta e} e - M .$$

Dzieląc obustronnie przez X i podstawiając $X = Me$, uzyskujemy:

$$\frac{\delta B}{\delta e} \frac{1}{X} = \frac{\delta X}{\delta e} \frac{1}{X} - \frac{1}{M} \frac{\delta M}{\delta e} - \frac{1}{e} .$$

Następnie mnożąc przez e, otrzymujemy równanie:

$$\frac{\delta B}{\delta e} \frac{e}{X} = \frac{\delta X}{\delta e} \frac{e}{X} - \frac{e}{M} \frac{\delta M}{\delta e} - 1. \tag{6.9}$$

Zdefiniujmy elastyczność popytu na eksport i elastyczność popytu na import względem kursu walutowego odpowiednio jako[25]:

$$\eta_X^D = \frac{\delta X}{\delta e} \frac{e}{X}$$

oraz

$$\eta_M^D = \frac{-\delta M}{\delta e} \frac{e}{M}. \tag{6.10}$$

Podstawiając (6.10) do (6.9), uzyskujemy:

$$\frac{\delta B}{\delta e} \frac{e}{X} = \eta_X^D + \eta_M^D - 1 \ . \tag{6.11}$$

Pamiętając, że, aby dewaluacja wywołała poprawę bilansu płatniczego, wyrażenie $\left(\frac{\delta B}{\delta e} \right)$ musi być większa od zera, otrzymujemy ostatecznie warunek (6.4):

[25] Pamiętajmy, że dewaluacja, czyli wzrost kursu, powoduje *wzrost* eksportu, ponieważ cena zagraniczna *spada*, oraz *spadek* importu, gdyż cena krajowa *rośnie*. To wyjaśnia znaki zastosowane w równaniu (6.10).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$\eta_X^D + \eta_M^D > 1. \tag{6.4}$$

Należy przypomnieć, że warunek MLR w postaci nierówności (6.4) zakłada, że w punkcie wyjścia bilans handlowy jest zrównoważony oraz że podaż eksportowa i podaż z importu są doskonale elastyczne. Uchylenie założenia mówiącego o tym, że $X = Me$, oraz uwzględnienie skończonych elastyczności podaży ($\eta_E^S, \eta_M^S < \infty$) przekształca warunek MLR we wzór o następującej postaci (zob. Rosati 1990):

$$\frac{\eta_M^S\left(\eta_E^D + 1\right)}{\eta_E^D - \eta_E^S} + b\frac{\eta_M^D\left(\eta_M^S + 1\right)}{\eta_M^D + \eta_M^S} > 0, \tag{6.12}$$

gdzie $b = X/Me$. Warunek (6.12) jest znacznie bardziej skomplikowany niż (6.4) i nie pozwala na wyciąganie prostych wniosków. W takiej sytuacji warunek MLR traci sporo ze swej początkowej atrakcyjności.

Zgodnie z tradycyjnym podejściem zmiany kursu waluty określone są przez relacje strumieni eksportu oraz importu. Przy ograniczonych przepływach kapitałowych i finansowych, nie związanych z handlem zagranicznym, kurs był zasadniczo funkcją salda obrotów bieżących. Pozwalało to wykorzystywać kurs walutowy jako instrument polityki przywracania równowagi płatniczej. Dzięki usztywnieniu kursu na określonym poziomie opisana zależność ulegała odwróceniu – w sytuacji, gdy kurs był administracyjnie ustalony, to on określał wielkości eksportu oraz importu. Dewaluacja kursu pozostawała w tych warunkach instrumentem pobudzania eksportu i ograniczania importu, a rewaluacja kursu funkcjonowała odwrotnie. System taki sprawdzał się dobrze w świecie stałych kursów i niewymienialnych walut (choć skutki dewaluacji były z reguły krótkotrwałe i zanikały w miarę wzrostu inflacji wywołanej dewaluacją). Stopniowe znoszenie ograniczeń w przepływach kapitału i postępująca integracja rynków finansowych, utrzymująca się od końca lat 70., powodowały jednak, że „strumieniowe" podejście do kwestii kształtowania się kursów walutowych stawało się coraz mniej wystarczające w celu wyjaśnienia zmian kursu i w efekcie traciło na znaczeniu.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

6.6.2. Kurs walutowy i rynek aktywów: model monetarny i model równowagi portfela

Alternatywą w przypadku podejścia „strumieniowego" (*trade flows approach*) jest podejście monetarne (*monetary approach*). Zgodnie z jego założeniami zmiany w podaży i popycie na waluty zależą od zmian w zakresie podaży i popytu na aktywa finansowe. Oznacza to, że zmiany w podaży i popycie na waluty rozpatrywane są z punktu widzenia zmian aktywów (*asset view*). Aktywa finansowe obejmują krajowe i zagraniczne papiery wartościowe oraz pieniądz krajowy i zagraniczny. Podejście monetarne zasadniczo zakłada, że kapitał jest doskonale mobilny w skali międzynarodowej oraz że nie występują koszty transakcyjne. W ramach tego ogólnego ujęcia można wyróżnić kilka odrębnych typów modeli kształtowania się kursu walutowego. Różnią się one, po pierwsze, założeniem co do stopnia substytucyjności między aktywami krajowymi i zagranicznymi, oraz, po drugie, założeniem co do stopnia giętkości cen.

Gdy krajowe i zagraniczne obligacje są doskonałymi substytutami, nie występuje dodatkowa premia za ryzyko. Wówczas inwestorzy kierują się przy zakupie obligacji wyłącznie oczekiwanymi stopami zwrotu, przy czym stopy zwrotu na takich samych obligacjach krajowych i zagranicznych są z definicji jednakowe. Wynika z tego, że kurs walutowy określony jest przez równanie niezabezpieczonego parytetu stóp procentowych (UIP), opisane już w rozdziale 5. Przyjmując za punkt wyjścia równanie parytetu (5.67), podejście monetarne implikuje, że:

$$\frac{e^p - e}{e} = i - i^* . \tag{6.13}$$

Gdy w ramach takiego modelu przyjmiemy także założenie o doskonale giętkich cenach oraz założenie o doskonałej substytucyjności towarów krajowych i zagranicznych, otrzymamy najbardziej „klasyczną" wersję monetarnego modelu kursu walutowego. Z kolei, gdy krajowe i zagraniczne obligacje nie są doskonałymi substytutami, premie za ryzyko na poszczególnych aktywach różnią się między sobą i inwestorzy przy zakupie aktywów kierują się zarówno oczekiwaną stopą zwrotu, jak i skalą ryzyka. Taki model nosi nazwę modelu równowagi portfela (*portfolio-balance model*).

Podstawowymi równaniami klasycznego modelu monetarnego są równania podaży i popytu na pieniądz w kraju i za granicą (zob. rozdział 5.1.3). Oznaczając

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ceny, dochód oraz stopę procentową kolejno jako *P, Y* i *R*, otrzymujemy (w warunkach równowagi):

$$M = L(P, Y, R) \tag{6.14}$$

$$M^* = L(P^*, Y^*, R^*). \tag{6.14a}$$

Wygodnie jest zapisać te równania w postaci logarytmicznej[26]. Łącząc oba powyższe wzory i zakładając, że $M = L$, otrzymujemy równanie relatywnego popytu na pieniądz w kraju względem popytu na pieniądz za granicą:

$$(m - m^*) = (p - p^*) + \varphi(y - y^*) - \lambda(i - i^*), \tag{6.15}$$

gdzie małe litery oznaczają logarytm naturalny odpowiedniej zmiennej, φ to dochodowa elastyczność popytu na pieniądz, a λ oznacza elastyczność popytu na pieniądz względem stopy procentowej. Założenie o doskonałej substytucyjności obligacji krajowych i zagranicznych oznacza, że spełniony jest niezabezpieczony parytet stóp procentowych (UIP), w związku z czym inwestorzy podejmują decyzje o zakupie aktywów, kierując się jedynie różnicami w stopach dochodowości i oczekiwaną zmianą kursu $E(\Delta e)$. Z kolei założenie o doskonałej substytucyjności towarów krajowych i zagranicznych oznacza, że spełniony jest warunek parytetu siły nabywczej (PPP). Po uwzględnieniu tych zależności otrzymujemy następujące monetarne równanie kursu walutowego (zob. wyprowadzenie równania w ramce 6.8):

$$e = \left(m - m^*\right) - \varphi\left(y - y^*\right) + \lambda\left[E\left(\Delta p\right) - E\left(\Delta p^*\right)\right]. \tag{6.16}$$

Prawa strona równania (6.16) pokazuje, że kurs walutowy *e* (w postaci logarytmicznej) jest funkcją relatywnych wielkości podaży i popytu na pieniądz w kraju i za granicą. Wzrost krajowej podaży pieniądza (m) wpływa na proporcjonalną deprecjację kursu waluty krajowej, podczas gdy wzrost zagranicznej podaży pieniądza (m^*) wywołuje proporcjonalną aprecjację. Podobnie wzrost dochodu krajowego (y) lub spadek oczekiwanej stopy inflacji krajowej $E(\Delta p)$ zwiększa popyt na pieniądz krajowy i pociąga za sobą aprecjację waluty krajowej.

[26] Postać logarytmiczna pozwala sprowadzić równania do formy liniowej, co upraszcza dalsze przekształcenia i ułatwia interpretację wyników.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 6.8. Monetarny model kursu walutowego[27]

Punktem wyjścia są równania popytu na pieniądz w kraju (6.14) i za granicą (6.14a). Założenie o doskonałej substytucyjności obligacji krajowych i zagranicznych oznacza, że spełniony jest niezabezpieczony parytet stóp procentowych (UIP):

$$i = i^* + E(\Delta e), \tag{6.17}$$

gdzie $E(\Delta e)$ jest logarytmem naturalnym oczekiwanej stopy dewaluacji kursu. Podstawiając (6.17) do równania monetarnego (6.15), otrzymujemy w przypadku relacji cen krajowych i zagranicznych następujący wzór:

$$\left(p - p^*\right) = \left(m - m^*\right) - \varphi\left(y - y^*\right) + \lambda E(\Delta e). \tag{6.18}$$

Założenie o doskonałej substytucyjności towarów krajowych i zagranicznych oznacza, że spełniony jest warunek parytetu siły nabywczej (PPP):

$$e = p - p^*, \tag{6.19}$$

gdzie e jest logarytmem bieżącego kursu walutowego. Z równania (6.19) wynika, że oczekiwana deprecjacja jest równa różnicy w oczekiwanych stopach inflacji w kraju i za granicą:

$$E(\Delta e) = E(\Delta p) - E\left(\Delta p^*\right). \tag{6.20}$$

Łącząc wzory (6.19), (6.18) i (6.20) otrzymujemy monetarne równanie kursu walutowego (6.16). Gdy ponadto założy się, że oczekiwania są formułowane w sposób racjonalny oraz że wzrost produkcji jest zdeterminowany egzogenicznie, wówczas oczekiwana stopa inflacji będzie równa oczekiwanej stopie wzrostu podaży pieniądza. Monetarne równanie kursu można wówczas zapisać w postaci uwzględniającej także *tempo wzrostu* podaży pieniądza π:

$$e = \left(m - m^*\right) - \varphi\left(y - y^*\right) + \lambda\left(\pi - \pi^*\right). \tag{6.21}$$

[27] Na podstawie Frankel (1983).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Monetarne modele kursu walutowego (6.16) i (6.21) zakładają w pełni giętkie ceny. Odmianą modelu monetarnego, w którym uwzględnia się sztywne ceny w krótkim okresie, jest model „przestrzelenia" kursu omówiony powyżej (punkt 6.5.4). Mimo swej zasadniczej poprawności (przy przyjętych założeniach) empiryczna weryfikacja obu typów modeli w wielu przypadkach przynosiła jednak wyniki niezgodne z ustaleniami teoretycznymi (dotyczyło to m.in. pamiętnego epizodu osłabienia dolara amerykańskiego wobec marki niemieckiej w latach 1978–1979 mimo szybszego wzrostu podaży pieniądza w Niemczech niż w USA). Jednym z możliwych uzasadnień w przypadku tego paradoksu jest uwzględnienie zróżnicowanego ryzyka na aktywach, co oznacza, że założenie o doskonałej substytucyjności krajowych i zagranicznych aktywów finansowych nie jest w rzeczywistości spełnione. Na tym gruncie rozwinęła się alternatywna teoria równowagi portfela (*portfolio-balance approach*), która wyjaśnia kształtowanie się kursu walutowego za pomocą różnic w poziomie ryzyka obarczającego różne rodzaje aktywów finansowych.

Ramka 6.9. Kurs walutowy w modelu równowagi portfela*

W modelu równowagi portfela zakłada się, podobnie jak w modelu monetarnym, że kapitał jest doskonale mobilny w skali międzynarodowej, lecz jednocześnie uchyla się założenie o doskonałej substytucyjności obligacji krajowych i zagranicznych. W warunkach funkcjonowania dwóch walut i zmiennego kursu walutowego inwestorzy wybierają taką relację zasobów obligacji krajowych i zagranicznych, która jest funkcją oczekiwanej względnej stopy zwrotu:

$$\frac{B}{\Phi F} = \psi\left[i - i^{*} - E(\Delta e)\right], \tag{6.22}$$

gdzie *B, F* oznaczają odpowiednio zasoby krajowych i zagranicznych obligacji, *Φ* jest kursem walutowym, a ψ to pewna funkcja, którą w celu uproszczenia dalszych przekształceń zdefiniujemy jako funkcję wykładniczą o postaci $exp\left[\alpha + \beta\left(i - i^{*} - E(\Delta e)\right)\right]$. Równanie (6.22) pokazuje, że wzrost krajowej stopy procentowej względem stopy zagranicznej – lub spadek oczekiwanej deprecjacji waluty krajowej – skłania inwestorów do zwiększania zasobu obligacji krajowych *B* kosztem zasobu obligacji zagranicznych *F*. Załóżmy, że oczekiwania są racjonalne i statyczne, co implikuje, że $E(\Delta e) = 0$. Logarytmując równanie (6.22)

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

otrzymujemy w ramach modelu równowagi portfela następujące najprostsze równanie kursu walutowego:

$$e = -\alpha + \beta\left(i - i^{*}\right) + b - f, \tag{6.23}$$

gdzie $e \equiv \ln \Phi$, $b \equiv \ln B$, $f \equiv \ln F$[28]. Zgodnie ze wzorem (6.23) kurs walutowy określony jest wyłącznie przez relacje dotyczące zasobów obligacji krajowych i zagranicznych oraz różnice w stopach procentowych. Ekonometryczna weryfikacja równania (6.23) nie potwierdza jednak w sposób jednoznaczny, że podejście oparte na równowadze portfela aktywów wyjaśnia w pełni mechanizm kształtowania się kursu walutowego (zob. np. Frankel 1983).

Uchylenie założenia o doskonałej substytucyjności aktywów krajowych i zagranicznych niewątpliwie przybliża portfelową teorię kursu walutowego do rzeczywistości. W istocie krajowe i zagraniczne obligacje różnią się pod wieloma względami (na przykład z uwagi na ryzyko kursowe, ryzyko płynności, ryzyko niewypłacalności, ryzyko wymienialności czy też stawki opodatkowania zysków). W ramce 6.9 omówiono koncepcję równowagi portfela, biorąc pod uwagę istnienie ryzyka kursowego. Punktem wyjścia w przypadku modeli równowagi portfela jest – tak jak w czystych modelach monetarnych – standardowe równanie popytu na pieniądz (6.14). Ponadto w ujęciu opartym na równowadze portfela zakłada się, że o ile aktywa krajowe i zagraniczne nie stanowią doskonałych substytutów, o tyle inwestorzy krajowi są bardziej skłonni do nabywania krajowych obligacji niż inwestorzy zagraniczni. Wówczas, gdy kraj notuje nadwyżkę obrotów bieżących w relacjach z zagranicą, napływa do niego kapitał (zagraniczne papiery wartościowe), co zwiększa majątek znajdujący się w posiadaniu inwestorów krajowych. Ten napływ bogactwa i wzrost majątku wywołuje zwiększenie zakupów obligacji krajowych w relacji do obligacji zagranicznych, co z kolei przekłada się na aprecjację kursu waluty krajowej wobec waluty zagranicznej. Również odwrotnie, gdy kraj notuje deficyt w obrotach bieżących, następuje odpływ kapitału i rośnie majątek pozostający w rękach inwestorów zagranicznych. Wzrost majątku powoduje zwiększenie

[28] Logarytmując równanie (6.22), otrzymujemy:

$$\ln B - \ln F - \ln \Phi = \ln[\exp\left(\alpha + \beta\left(i - i^{*} - E(\Delta e)\right)\right].$$

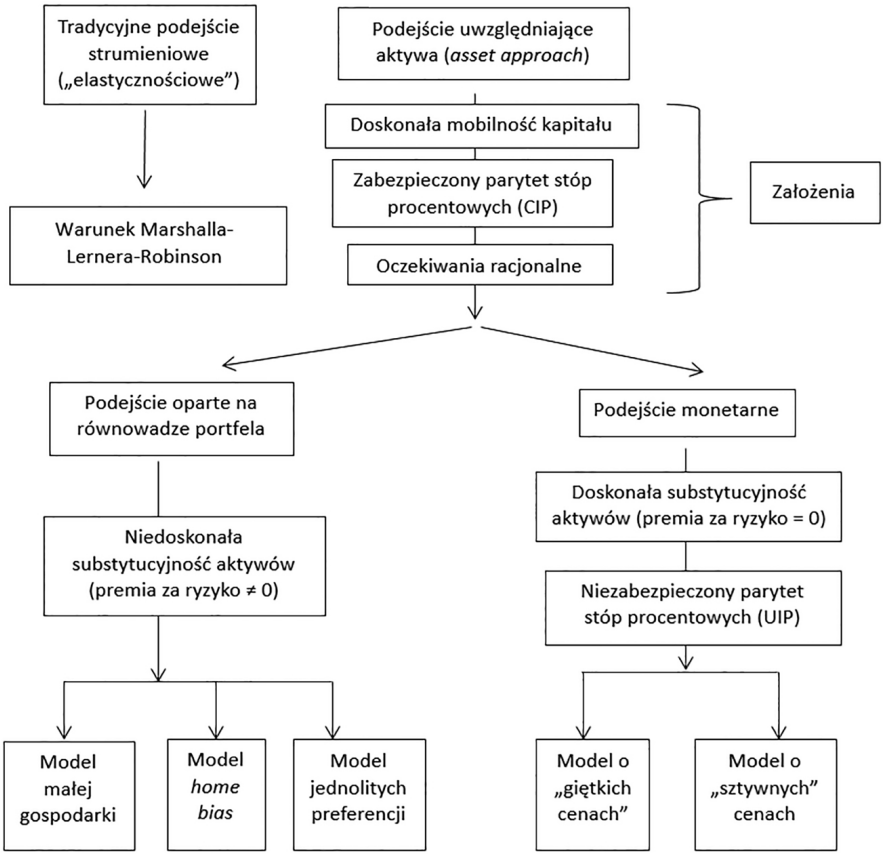
Ze względu na to, że logarytm naturalny ze stałej Eulera jest równy jedności, otrzymujemy równanie (6.23).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zakupów obligacji zagranicznych, co wywołuje deprecjację waluty krajowej wobec waluty zagranicznej.

Podsumowując dyskusję na temat mechanizmów kształtowania się kursu walutowego, trzeba zauważyć, że nie ma wśród ekonomistów zgody co do tego, jaki model kształtowania się kursu najlepiej odpowiada zaobserwowanym zjawiskom. O ile panuje powszechne przekonanie, że tradycyjne podejście „elastycznościowe", oparte na relacjach strumieni handlu, straciło swoją aktualność we współczesnych warunkach, kiedy o poziomie kursu decydują głównie masowe przepływy kapitału, o tyle nie ma pewności co do tego, które z modeli monetarnych pozostają najbardziej wiarygodne w opisie rzeczywistości. Rysunek 6.11 przedstawia klasyfikację różnych modeli kształtowania się kursu walutowego.

Rysunek 6.11. Klasyfikacja różnych modeli kursu walutowego



Źródło: Frankel (1983).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

6.7. Kurs walutowy w krótkookresowym modelu makroekonomicznym*

6.7.1. Kurs walutowy jako czynnik określający popyt w gospodarce otwartej

Omówimy obecnie kształtowanie się kursu walutowego w ramach krótkookresowego, keynesowskiego modelu makroekonomicznego. Punktem wyjścia będzie model IS-LM opisany szczegółowo w rozdziale 2. Jak pamiętamy, model ten ma charakter krótkookresowy, ponieważ obowiązuje w nim założenie stałych cen. Uzupełnimy go obecnie o kurs walutowy i zbadamy to, jak zmiany kursu wpływają na poziom produkcji i na równowagę ogólną. Przypomnijmy, że zagregowany popyt D na cały wytworzony produkt w gospodarce Y jest sumą konsumpcji (C), inwestycji (I), popytu sektora rządowego (G) oraz eksportu netto (X), który stanowi różnicę między eksportem EX i importem IM:

$$D \equiv C + I + G + (EX - IM) \equiv C + I + G + X. \tag{6.24}$$

Należy pamiętać, że równanie (6.24) jest tożsamością wynikającą z przyjętych definicji, a nie funkcją czy też warunkiem równowagi. Równowaga na rynku produktów oznacza, że popyt równa się podaży, a więc $D = Y$ (co w rzeczywistości może, ale nie musi zachodzić *ex ante*, choć zawsze zachodzi *ex post*).

Na każdy składnik popytu ma wpływ wiele innych czynników. W najprostszym (keynesowskim) ujęciu konsumpcja jest funkcją dochodu pozostającego do dyspozycji (dochód całkowity Y minus podatki T), inwestycje są funkcją realnej stopy procentowej r, z kolei eksport netto X stanowi różnicę między eksportem (EX) a importem (IM) dóbr i usług, przy czym eksport jest rosnącą, a import malejącą funkcją realnego kursu waluty obcej $e^R = e\frac{P^*}{P}$ (czyli wskaźnika kursu nominalnego e pomnożonego przez stosunek wskaźnika cen zagranicznych P^* oraz wskaźnika cen krajowych P). Ponadto import stanowi rosnącą funkcję dochodu Y. Wydatki rządowe G (po odjęciu podatków) są zmienną egzogeniczną (zależną od decyzji władzy gospodarczej). Po uwzględnieniu tych zależności oraz pominięciu podatkow, równanie zagregowanego popytu przyjmuje następującą postać:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$D = C(Y) + I(r) + G + X\left[\left(e\frac{P^*}{P}\right), Y\right]. \tag{6.25}$$

Tak więc – w postaci ogólnej – popyt D jest funkcją czterech zmiennych:

$$D = D\left[\left(e\frac{P^*}{P}\right), Y, r, G\right]. \tag{6.26}$$

Interesuje nas wpływ zmian poszczególnych zmiennych na efektywny popyt zagregowany D. Wzrost kursu e, czyli deprecjacja waluty krajowej, wpływa na spadek cen eksportowych za granicą i wzrost cen importu w kraju, co przesuwa popyt z dóbr zagranicznych na dobra krajowe zarówno w kraju, jak i za granicą (podobnie działa wzrost cen zagranicznych P^* i spadek cen krajowych P). Spadek kursu e, czyli aprecjacja waluty krajowej, działa odwrotnie. Z kolei wzrost dochodu Y (przy stałych podatkach T) wywołuje wzrost dochodu pozostającego do dyspozycji, co oznacza zwiększenie poziomu konsumpcji (choć o mniejszą wielkość niż wzrost Y, ponieważ część dochodu jest oszczędzana) oraz wzrost importu. Pierwszy efekt wpływa na wzrost D, podczas gdy drugi wywołuje spadek D (gdyż spada eksport netto X). Łączny efekt pozostanie jednak wzrostowy, ponieważ wydatki przeznaczane na całkowitą konsumpcję są zawsze większe niż wydatki na import konsumpcyjny.

6.7.2. Równowaga na rynku produktów i konstrukcja krzywej *DD*

Równowaga na rynku dóbr i usług wymaga, aby produkcja Y równała się popytowi D:

$$Y = D\left[\left(e\frac{P^*}{P}\right), Y, r, G\right]. \tag{6.27}$$

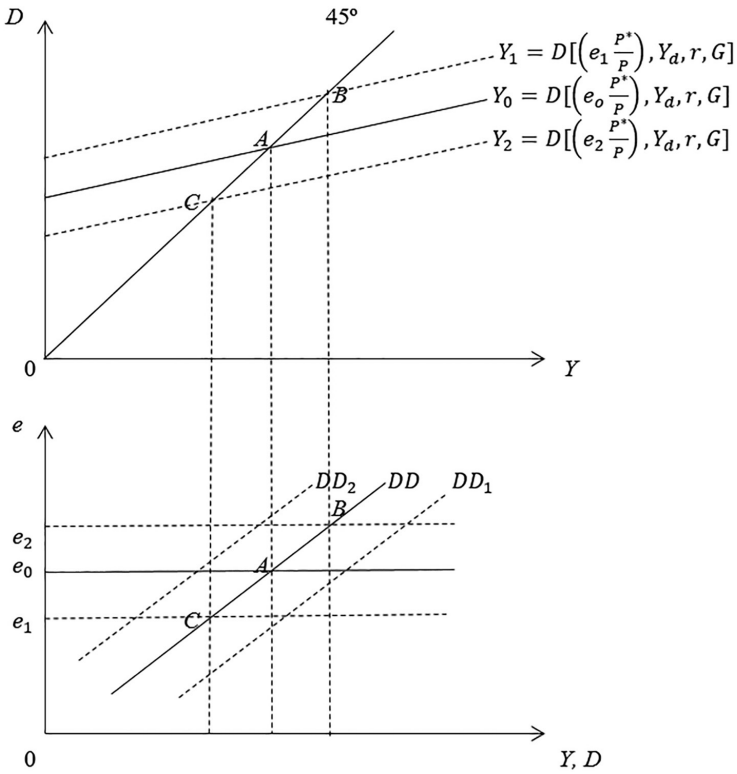
Wpływ zmian kursu na równowagę ogólną ukazany jest w górnej części rysunku 6.8. Gospodarka osiąga równowagę w punkcie A, gdzie zagregowana produkcja Y równa jest zagregowanemu popytowi D. Załóżmy, że w krótkim okresie ceny krajowe P i ceny zagraniczne P^* nie zmieniają się. Deprecjacja kursu waluty krajowej z e_0 do e_1 sprawia, że zagraniczne dobra i usługi stają się droższe w porównaniu z krajowymi[29].

[29] Załóżmy, że cena jednostki tego samego towaru wyprodukowanego za granicą wynosi 1 EUR, dobra wytworzonego w kraju 4 PLN, a kurs walutowy osiąga wartość 4 PLN/EUR. Deprecjacja kursu o 25%,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W związku z tym popyt przesuwa się na towary krajowe, co wywołuje wzrost eksportu netto X (eksport rośnie, a import spada) i przesuwa krzywą D w górę z położenia D_0 do D_1, wywołując stopniowy wzrost produkcji z Y_0 do Y_1. Punkt równowagi przesuwa się z A do B. Również odwrotnie, aprecjacja kursu walutowego z e_0 do e_2 wpływa na spadek cen towarów zagranicznych w porównaniu z krajowymi, popyt przesuwa się na towary zagraniczne, a krzywa D przemieszcza się w dół do położenia D_2, powodując stopniowy spadek produkcji do poziomu Y_2. Punkt równowagi przesuwa się z A do C.

Rysunek 6.12. Równowaga wewnętrzna i konstrukcja krzywej *DD*



Źródło: opracowanie własne.

a więc z 4 PLN/EUR do 5 PLN/EUR, sprawia, że jednostka towaru kupiona za granicą kosztuje nas teraz 5 PLN. Oznacza to, że popyt krajowy na towary sprowadzane z zagranicy maleje, podczas gdy wzrasta popyt na towary produkowane w kraju.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rzutujemy teraz punkty równowagi *A*, *B*, *C* dotyczące różnych wartości kursu walutowego na dolną część rysunku 6.12, gdzie na osi pionowej odłożono kurs walutowy. Linia łącząca punkty *A*, *B* i *C* wyznacza krzywą *DD,* przedstawiającą zależność równowagi na rynku produktów od kursu walutowego. Krzywa *DD* stanowi zbiór punktów reprezentujących takie wartości kursu walutowego *e* i produkcji *Y*, które zapewniają równowagę na rynku produktów (Krugman i Obstfeld 1991). Zauważmy, że wyższym poziomom produkcji *Y* odpowiadają wyższe poziomy kursu waluty zagranicznej, ponieważ wraz ze wzrostem produkcji krajowej popyt powinien przesuwać się na produkcję krajową kosztem produkcji zagranicznej, tak aby zapewniona była równowaga na rynku towarów. Zmiana kursu walutowego *e* pociąga za sobą odchylenie efektywnego popytu *D* od produkcji *Y*, co wywołuje stopniowe dostosowanie *Y* do nowego poziomu popytu, a dzięki temu następuje samoczynny powrót do równowagi. W szczególności zaś obniżenie kursu waluty krajowej (deprecjacja) z e_0 do e_1 powoduje wzrost eksportu netto *X*, co wpływa na stopniowe przesuwanie się krzywej *DD* w prawo i w dół.

Krzywa ta może przesuwać się pod wpływem zmian dokonujących się w zakresie innych czynników występujących w równaniu (6.26): wydatków rządowych *G,* realnej stopy procentowej *r,* podatków *T* oraz relacji cen zagranicznych i krajowych $\frac{P^*}{P}$. Na przykład wzrost wydatków rządowych *G*, obniżenie podatków *T* lub spadek stóp procentowych *r* – przy stałym kursie e_1 – przesuwa krzywą *DD* w prawo, do położenia DD_1. Z kolei spadek wydatków rządowych *G*, wzrost stóp procentowych lub wzrost relacji $\frac{P^*}{P}$ przesuwa krzywa *DD* w lewo, do położenia DD_2. Ogólnie rzecz biorąc, czynniki, które powiększają zagregowany popyt *D*, wpływają na przesunięcie krzywej *DD* w prawo, a czynniki, które obniżają zagregowany popyt *D*, powodują przemieszczenie się krzywej *DD* w lewo.

6.7.3. Równowaga na rynku pieniądza krajowego i na rynku pieniądza zagranicznego. Konstrukcja krzywej *AA*

W gospodarce otwartej funkcjonują rynki aktywów finansowych, do których zalicza się rynek pieniądza krajowego, rynek pieniądza zagranicznego oraz rynek papierów wartościowych. Aby równowaga ogólna została zachowana, wszystkie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

rynki cząstkowe muszą być w równowadze, w tym także rynek pieniądza zagranicznego, zwany też rynkiem walutowym. Równowaga na rynku pieniądza zagranicznego wymaga, aby podaż walut równała się popytowi przy danym poziomie kursu walutowego. Brak równowagi między podażą a popytem może powodować niekontrolowany odpływ kapitału i doprowadzić w konsekwencji do kryzysu walutowego. Może też wywołać niekontrolowany napływ kapitału i destabilizację finansową w postaci baniek spekulacyjnych. Równowaga między podażą i popytem zostanie zachowana wówczas, gdy oczekiwana dochodowość aktywów (lokat, depozytów, instrumentów pochodnych) w walucie krajowej oraz zagranicznej, z uwzględnieniem ryzyka, będzie taka sama.

W warunkach otwartej gospodarki, przy braku przeszkód w przepływie kapitału, występuje tendencja do wyrównywania się dochodowości aktywów w walucie krajowej i zagranicznej. Jak już wiemy, tendencję tę nazywa się parytetem stóp procentowych (*interest rate parity* – zob. rozdział 5.9.1). Przypomnijmy, że w przypadku danego oczekiwanego poziomu kursu walutowego w przyszłości e^p równanie parytetu stóp procentowych jest następujące[30]:

$$i = i^* + \frac{e^p - e}{e} = i^* + \frac{de}{e}.$$

Równanie to pokazuje, że rynek pieniądza zagranicznego pozostanie w równowadze (podaż walut będzie równa popytowi na waluty), gdy dochodowość lokat w walucie krajowej i zagranicznej, po uwzględnieniu oczekiwanej zmiany kursu, stanie się jednakowa.

Podobnie dzieje się, gdy rynek pieniądza krajowego znajduje się w równowadze, a podaż pieniądza krajowego M równa się popytowi L. Równowagę zapewnia odpowiedni poziom krajowej stopy procentowej i, zgodnie z równaniem (2.11b) opisanym w rozdziale 2:

$$M/P = L(i, y), \text{ przy czym } \delta L/\delta i < 0,\ \delta L/\delta y > 0.$$

Przyjmijmy w celu uproszczenia, że rynek aktywów finansowych obejmuje rynek pieniądza krajowego i rynek pieniądza zagranicznego[31]. Wówczas równowagę na obu rynkach można przedstawić na rysunku 6.13, który składa się z dwóch

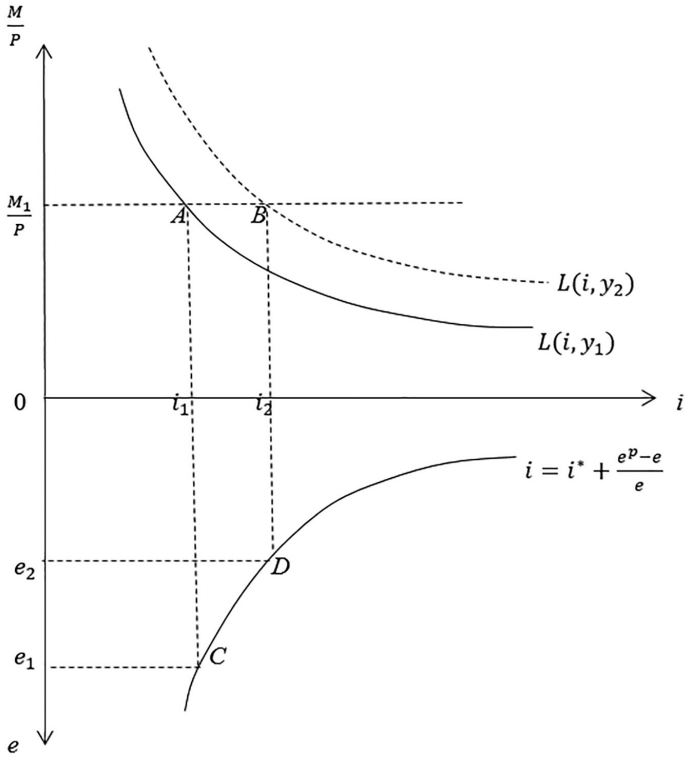
[30] Zob. równanie (5.67).

[31] Rozszerzenie analizy na rynek papierów wartościowych nie ma wpływu na wynik rozumowania.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

części. Górna część rysunku przedstawia równowagę na rynku pieniądza krajowego. Krzywa L stanowi zbiór takich wartości realnej podaży pieniądza i wartości stopy procentowej, które zapewniają równowagę między podażą pieniądza i popytem na pieniądz w przypadku danego poziomu produkcji y. Na przykład wobec danego poziomu produkcji y_1 i przy danej realnej podaży pieniądza $\left(\frac{M_1}{P}\right)$ rynek pieniądza krajowego znajdzie się w równowadze przy określonym poziomie stopy procentowej i_1. Krzywa popytu na pieniądz L jest opadająca, ponieważ spadkowi realnej podaży pieniądza musi towarzyszyć wzrost stóp procentowych, tak aby popyt na pieniądz spadał, dostosowując się do niższej podaży.

Rysunek 6.13. Równowaga na rynku pieniądza krajowego i pieniądza zagranicznego



Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Z kolei dolna część rysunku 6.13 ukazuje równowagę na rynku pieniądza zagranicznego. Krzywa parytetu (5.67) stanowi zbiór takich wartości stopy procentowej i kursu walutowego, które zapewniają równowagę między podażą pieniądza zagranicznego a popytem na ten pieniądz. Na przykład przy danym poziomie zagranicznych i krajowych stóp procentowych i_1 oraz i^*, a także przy danym oczekiwanym poziomie kursu walutowego w przyszłości e^p równowagę na rynku walutowym zapewnia poziom bieżącego kursu e_1. Krzywa parytetu jest także opadająca, ponieważ wzrost krajowej stopy procentowej przy danym i^* oraz e^p wymaga spadku e, czyli aprecjacji waluty krajowej, aby zachowana została równowaga.

Załóżmy, że rynek pieniądza krajowego znajduje się w równowadze w punkcie A, a rynek walutowy osiąga równowagę w punkcie C, i przyjmijmy, że następuje wzrost produkcji z y_1 do y_2. Wywołuje to wzrost popytu (transakcyjnego) na pieniądz krajowy, co wymaga przy danej podaży pieniądza wzrostu stopy procentowej do poziomu i_2. Nowym punktem równowagi na rynku pieniądza krajowego jest teraz punkt B. Lecz w tym miejscu zostaje naruszona równowaga na rynku walutowym – na skutek wzrostu krajowej stopy procentowej lewa strona równania parytetu (5.67) jest teraz większa od prawej. Przywrócenie równowagi przy danym oczekiwanym kursie w przyszłości e^p wymaga *bieżącej* aprecjacji waluty krajowej z e_1 do e_2 po to, aby zwiększona skala *oczekiwanej* deprecjacji $(e^p - e_2)$ doprowadziła do odpowiedniego wzrostu prawej strony równania parytetu (równowaga zostaje przywrócona w punkcie D). Wynika z tego, że utrzymanie równowagi na rynku pieniądza krajowego i pieniądza zagranicznego w warunkach wzrostu produkcji krajowej wymaga aprecjacji kursu, a w warunkach spadku produkcji krajowej konieczna jest deprecjacja kursu.

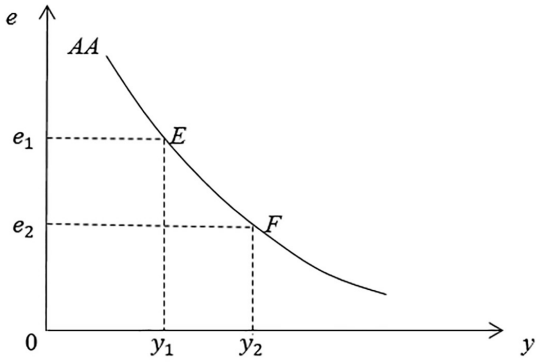
Oznacza to, że w przypadku każdego poziomu produkcji y istnieje tylko jeden poziom kursu e, który spełnia warunek parytetu stóp procentowych (przy danych M/P, i^* oraz e^p). Im wyższy staje się poziom y, tym niższy musi być kurs waluty obcej. Zależność tę ilustruje krzywa AA na rysunku 6.14. Stanowi ona zbiór punktów reprezentujących takie pary wartości produkcji y i kursu walutowego e, które zapewniają jednoczesną równowagę na rynkach pieniądza krajowego i zagranicznego.

Wyprowadzenie krzywej AA jest proste i wynika bezpośrednio z rysunku 6.13. Zauważmy mianowicie, że przy danej realnej podaży pieniądza krajowego $\left(\frac{M_1}{P}\right)$ oraz określonym poziomie produkcji y_1 równowagę na rynku pieniądza krajowego

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zapewnia stopa procentowa i_1, podczas gdy o równowadze na rynku pieniądza zagranicznego decyduje kurs e_1. Przejdźmy teraz do rysunku 6.14 i odłóżmy na osi poziomej y wielkość produkcji y_1, a na osi pionowej e poziom kursu e_1. Wyznaczony przez te współrzędne punkt E jest więc punktem równowagi zarówno na rynku pieniądza krajowego, jak i na rynku walutowym. Podobnie wyznaczymy punkt F, który odpowiada wielkości produkcji y_2 i kursowi walutowemu e_2. Punkt F musi znajdować się po prawej stronie i poniżej punktu E, ponieważ $y_2 > y_1$ oraz $e_2 < e_1$. Łącząc punkty E i F, otrzymujemy krzywą AA.

Rysunek 6.14. Zależność między produkcją i kursem walutowym. Krzywa *AA*



Źródło: opracowanie własne.

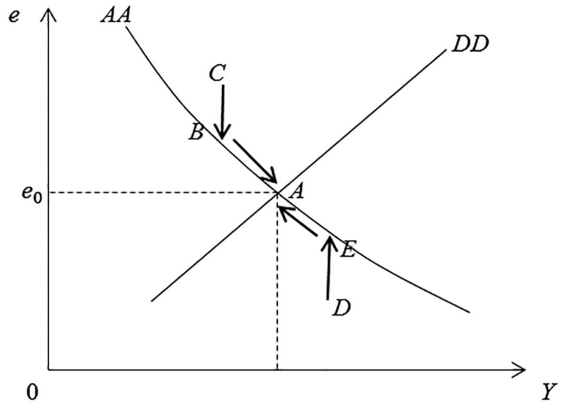
Krzywa AA może przesuwać się w lewo lub w prawo na skutek zmian każdej z pięciu zmiennych występujących w równaniu zagregowanego popytu (6.27), które określają położenie krzywej D. W szczególności zaś wzrost nominalnej podaży pieniądza M, przy stałych cenach, zwiększa realną podaż pieniądza i przesuwa krzywą AA w prawo, podczas gdy wzrost cen P zmniejsza realną podaż pieniądza i przesuwa krzywą AA w lewo. Z kolei wzrost oczekiwanego kursu walutowego e^p w przyszłości powoduje wzrost oczekiwanej deprecjacji i stóp procentowych, co decyduje o przesunięciu krzywej AA w prawo. Wzrost zagranicznych stóp procentowych i^* wywołuje deprecjację waluty krajowej i także przesuwa krzywą AA w prawo. Wreszcie, egzogeniczny wzrost popytu na pieniądz realny L przesuwa krzywą L w prawo, powodując tym samym wzrost i oraz spadek e, w związku z czym krzywa AA przemieszcza się w lewo.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

6.7.4. Równowaga ogólna i przywracanie równowagi

Krzywa *DD* odzwierciedla kombinacje produkcji i kursu walutowego zapewniające równowagę na rynku produktów, z kolei krzywa *AA* ukazuje kombinacje produkcji i kursu walutowego zapewniające równowagę na rynkach aktywów finansowych (na rynku pieniądza krajowego oraz na rynku pieniądza zagranicznego). Równowaga ogólna w gospodarce wymaga, aby wszystkie rynki cząstkowe znajdowały się w równowadze. Tę równowagę ogólną przedstawiono na rysunku 6.15, gdzie wykreślono krzywą *DD* i krzywą *AA* (nadal zakładamy stałość cen *P*, stałość zagranicznych stóp procentowych i^* oraz stałość oczekiwanego kursu walutowego w przyszłości e^p). Ponieważ krzywa *DD* jest monotonicznie rosnąca, a krzywa *AA* monotonicznie opadająca, istnieje tylko jeden punkt przecięcia się obu tych krzywych – punkt *A*. Wyznacza on jedyną kombinację wartości produkcji (Y_0) i kursu walutowego (e_0), zapewniającą *jednocześnie* równowagę na rynku produktów oraz na rynkach aktywów finansowych.

Rysunek 6.15. Krótkookresowa równowaga na rynku produktów oraz na rynku pieniądza



Źródło: opracowanie własne.

Możemy się łatwo przekonać, że punkt *A* jest rzeczywiście punktem trwałej równowagi i że w przypadku jej naruszenia mechanizm rynkowy zapewnia gospodarce powrót do punktu równowagi. Załóżmy, że w wyniku negatywnego szoku zewnętrznego kurs walutowy odchylił się w górę od poziomu równowagi e_0

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(deprecjacja waluty krajowej) i gospodarka znalazła się w punkcie *C*. Wysoki poziom *e* implikuje wzrost oczekiwanej aprecjacji przy danym e^p. To oznacza, przy danych stopach procentowych, że dochodowość aktywów krajowych jest wyższa niż w przypadku aktywów zagranicznych. Rośnie więc popyt na walutę krajową, kurs walutowy ulega aprecjacji (obniża się) i gospodarka przesuwa się do punktu *B* położonego na krzywej *AA*. Lecz w punkcie *B* występuje przewaga popytu nad produkcją krajową (ponieważ punkt *B* znajduje się powyżej krzywej *DD*). Następuje więc wzrost produkcji i gospodarka stopniowo powraca do punktu *A*.

Podobnie wygląda dostosowanie, gdy gospodarka z jakiegoś powodu znajdzie się w punkcie *D*. Oczekiwana deprecjacja wywołuje spadek popytu na walutę krajową i kurs osłabia się w punkcie *E*. Lecz jednocześnie w tym miejscu produkcja staje się większa od popytu (punkt *E* znajduje się poniżej krzywej *DD*). Następuje więc stopniowy spadek produkcji i gospodarka powraca do punktu *A*. Należy podkreślić, że rynek aktywów reaguje natychmiast, podczas gdy rynek produktów wolno i z opóźnieniem. Dlatego też zawsze najpierw dostosowuje się kurs walutowy i gospodarka wraca na krzywą *AA*, a dopiero potem stopniowo przesuwa się wzdłuż krzywej *AA* do punktu przecięcia się z krzywą *DD*.

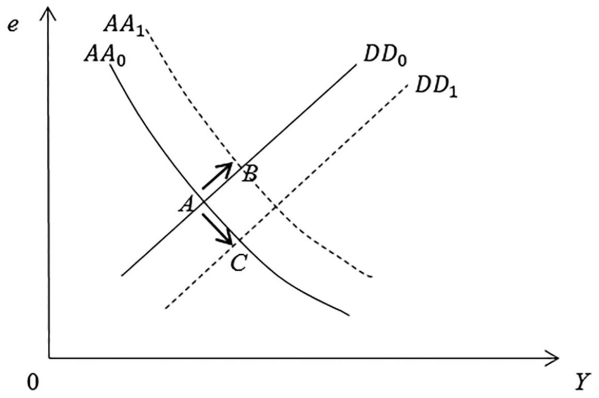
Zakłócenia równowagi mogą wynikać z określonych decyzji podejmowanych w zakresie polityki gospodarczej. Bardzo często mają one na celu zwiększenie poziomu produkcji, co w konsekwencji prowadzi do wzrostu zatrudnienia. Załóżmy, że władza gospodarcza chce pobudzić gospodarkę, aby uporać się z nadmiernym bezrobociem, i że ma do dyspozycji dwa instrumenty – politykę pieniężną i politykę fiskalną. Przyjmijmy także, że zmiany w polityce makroekonomicznej mają charakter przejściowy.

Wobec tych założeń wykorzystanie polityki pieniężnej do pobudzenia gospodarki oznacza przejściowy wzrost podaży pieniądza. Jak wynika z rysunku 6.13, przy stałych cenach wzrost podaży pieniądza prowadzi do spadku krajowych stóp procentowych, a to powoduje natychmiastową deprecjację waluty krajowej. W efekcie krzywa *AA* przesuwa się w prawo, podczas gdy położenie krzywej *DD* nie zmienia się. Rysunek 6.16 pokazuje, że nowym punktem krótkookresowej równowagi staje się punkt *B*, charakteryzujący się wyższym poziomem kursu i wyższym poziomem produkcji niż punkt *A*[32]. Wzrost produkcji wywołany jest w tym przypadku wzrostem popytu wynikającym z deprecjacji kursu walutowego.

[32] Ponieważ wzrost podaży pieniądza ma charakter przejściowy, oczekiwany w przyszłości kurs walutowy e^p nie zmienia się.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 6.16. Wpływ polityki pieniężnej i polityki fiskalnej na produkcję i kurs walutowy



Źródło: opracowanie własne.

Z kolei wykorzystanie polityki fiskalnej do pobudzenia gospodarki polega na przejściowym wzroście wydatków (lub przejściowej obniżce podatków). Zgodnie z równaniem (6.27) wzrost wydatków przesuwa krzywą *DD* w prawo, lecz nie wpływa na położenie krzywej *AA*. Nowym punktem równowagi staje się punkt *C*, charakteryzujący się wyższym poziomem produkcji przy niższym poziomie kursu niż punkt *A*. Dlaczego mamy w tym przypadku do czynienia z umocnieniem się kursu waluty krajowej? Wzrost wydatków rządowych – przy danej podaży pieniądza – powoduje wzrost stóp procentowych. Aby równanie parytetu było spełnione przy niezmienionym poziomie oczekiwanego kursu w przyszłości, obecny kurs musi się umocnić (czyli skala *oczekiwanej* deprecjacji musi się powiększyć). Wzrost produkcji wywołany jest w tym przypadku wzrostem zagregowanego popytu, spowodowanym zwiększonymi wydatkami rządowymi. Podobne efekty będzie miało obniżenie podatków.

Przesunięcia krzywych *DD* i *AA* mogą być też skutkiem szoków egzogenicznych (zewnętrznych lub wewnętrznych). Przykładem szoku zewnętrznego jest spadek popytu światowego na produkty krajowe, który wpływa na przesuwa się krzywej *DD* w lewo. W tej sytuacji kraj może albo zwiększyć wydatki rządowe i przywrócić pierwotne położenie *DD*, albo zwiększyć podaż pieniądza, przesuwając krzywą *AA* w prawo i przywracając poprzedni poziom *Y* przy wyższym kursie *e*. Przykładem szoku wewnętrznego jest egzogeniczny wzrost popytu krajowego na pieniądz, który przesuwa krzywą *AA* w lewo. W tej sytuacji kraj może albo zwiększyć

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

podaż pieniądza i przywrócić pierwotne położenie krzywej *AA*, albo zwiększyć wydatki rządowe, przesuwając krzywą *DD* w prawo oraz przywracając poprzedni poziom *Y* przy niższym kursie *e*. Wybór konkretnej polityki będzie zależał od źródła szoku (kraj czy zagranica), wpływu zmian kursu na inflację, poziomu długu publicznego i ograniczeń politycznych (zdolność rządu do podniesienia podatków i zwiększenia wydatków).

6.8. Kurs walutowy w średnim i długim okresie

6.8.1. Kurs równowagi długookresowej (*equilibrium exchange rate*)

Wszystkie transakcje międzynarodowe rejestrowane w bilansie płatniczym można podzielić na transakcje związane z zakupem i sprzedażą dóbr i usług (obroty bieżące) oraz transakcje związane z zakupem i sprzedażą aktywów finansowych (obroty finansowe). Te pierwsze odzwierciedlają potrzeby realnego sektora w gospodarce, drugie mają natomiast charakter spekulacyjny. Gdy międzynarodowe przepływy kapitałowe były ograniczone, kurs walutowy kształtował się głównie w odpowiedzi na zmiany zachodzące w realnej gospodarce. Wahania kursu walutowego były relatywnie niewielkie, ponieważ zmienne realne, takie jak produkcja, eksport czy import, zmieniały się stosunkowo powoli. U ich podstaw leżą bowiem zmiany technologiczne, inwestycje oraz wzrost wydajności pracy – czyli tzw. czynniki fundamentalne – które z natury dokonują się powoli i stopniowo. Gdy, poczynając od lat 80., wzrastała jednak rola transgranicznych przepływów kapitału, czynniki te zaczęły w coraz większym stopniu wpływać na kształtowanie się kursów walutowych. Ponieważ przepływy kapitałowe podlegają z kolei nieporównanie większym wahaniom niż zmienne realne, wzrosła również zmienność kursu walut. W efekcie aktualny poziom kursu coraz częściej odchylał się, niekiedy znacznie, od poziomu długookresowej równowagi, określonej przez potrzeby sektora realnego gospodarki.

Postawiło to przed polityką kursową trudne wyzwania. Gdy kraj decydował się zastosować system kursu stałego, należało ustalić, jaki poziom kursu zapewni równowagę podaży i popytu na walutę obcą w średnim/długim okresie, czyli będzie spójny z niską inflacją i pełnym zatrudnieniem. Z kolei w systemie kursu płynnego chodziło o sprawdzenie, czy wahania kursu rynkowego odzwierciedlają zmiany zachodzące w czynnikach fundamentalnych (np. wzrost konkurencyjności

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

dokonujący za sprawą wzrostu wydajności pracy), czy raczej są wynikiem spekulacyjnych przepływów kapitału. Dlatego też jednym z najważniejszych wyzwań, jakie stanęły przed polityką kursową, okazało się ustalenie takiego poziomu kursu, który zapewni równowagę płatniczą w średnim/długim okresie. Angielski ekonomista John Williamson nazwał ten kurs kursem równowagi fundamentalnej (*fundamental equilibrium exchange rate* – FEER). Kurs FEER jest definiowany jako kurs zapewniający trwałą równowagę bilansu płatniczego przy danej polityce gospodarczej[33].

6.8.2. Kurs parytetu siły nabywczej: rozwinięcie

Aby wyjaśnić, jak kształtuje się kurs równowagi w dłuższym okresie, trzeba wziąć pod uwagę czynniki kształtujące poziom relacji cen między krajem i zagranicą. O ile w ramach analizy krótkookresowej przyjmuje się, że poziom cen nie ulega zmianie (założenie o krótkookresowej sztywności cen), o tyle w długim okresie ceny zmieniają się w zależności od warunków podaży i popytu. Punktem wyjścia przy określeniu kursu równowagi w dłuższym okresie jest teoria parytetu siły nabywczej (PPP), zgodnie z którą kurs walutowy równy jest – w tendencji – relacji poziomów cen w kraju i za granicą (zob. punkt 6.1). Ponieważ obserwacje empiryczne nie zawsze potwierdzają jednak działanie PPP, należy przypuszczać, że w długim okresie na kurs muszą oddziaływać także inne czynniki.

Kluczowe znaczenie w przypadku teorii PPP ma tzw. prawo jednej ceny (*the law of one price*), czyli twierdzenie zakładające, że w warunkach wolnej konkurencji i braku barier w handlu, ceny tych samych towarów, przeliczone po danym kursie na wspólną walutę, wyrównują się w skali międzynarodowej. W związku z tym mechanizm arbitrażu rynkowego zapewnia, że:

$$P_i = eP_i^*.$$

Stąd zaś:

$$e = P_i / P_i^*, \tag{6.28}$$

gdzie P_i i P_i^* oznaczają cenę towaru i w kraju oraz za granicą, z kolei e jest kursem waluty zagranicznej wyrażonym w walucie krajowej. Gdy prawo jednej ceny

[33] Kurs FEER zapewnia więc utrzymanie pewnego poziomu salda obrotów bieżących (deficytu lub nadwyżki), któremu towarzyszy określona skala przepływów kapitału w okresie cyklu koniunkturalnego, i przy założeniu, że kraj prowadzi politykę gospodarczą, której celem jest utrzymanie wewnętrznej równowagi makroekonomicznej (Williamson 1983, Cline i Williamson 2008).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zastosujemy równocześnie do wielu produktów, wzór (6.28) przekształca się w równanie parytetu siły nabywczej:

$$e = P/P^{*}. \tag{6.29}$$

Równanie (6.29) wskazuje, że, zgodnie z teorią PPP, kurs walutowy równy jest stosunkowi *poziomów* cen, przy czym chodzi o ceny jednakowych, porównywalnych i reprezentatywnych „koszyków" dóbr oraz usług w kraju i za granicą. Na przykład, jeśli ten sam „koszyk" kosztuje 1000 EUR w Niemczech i 2000 PLN w Polsce, to kurs PPP wynosi 2000 PLN / 1000 EUR = 2 PLN/EUR. Gdy ceny w Polsce rosną w relacji do cen obowiązujących w Niemczech, siła nabywcza polskiej waluty spada i kurs złotego ulega deprecjacji, podczas gdy kurs euro ulega aprecjacji. W konsekwencji teoria PPP implikuje, że ceny tych samych dóbr na całym świecie pozostają sobie równe po przeliczeniu na jedną walutę[34].

Parytet PPP występuje w dwóch odmianach. W wersji bezwzględnej kurs walutowy równy jest relacji *poziomów* cen w kraju i za granicą (równanie 6.29), z kolei wersja względna (relatywna) zakłada, że relatywne *zmiany* kursu pozostają równe względnym *zmianom* relacji poziomów cen, czyli:

$$de / e = d\left(\frac{P}{P^{*}}\right) / \left(\frac{P}{P^{*}}\right). \tag{6.30}$$

Wersja względna okazuje się bardziej praktyczna w zastosowaniu, ponieważ pozwala posługiwać się zmianami procentowymi, co umożliwia dokonywanie porównań w skali międzynarodowej.

Jak pamiętamy z poprzedniego rozdziału, zmiany cen w długim okresie są określone przez zmiany relacji podaży i popytu na pieniądz na rynku pieniężnym. Połączenie teorii PPP z równaniem równowagi na rynku pieniądza pozwala uzyskać praktyczny instrument służący badaniu długookresowej zależności między kursem walutowym a podażą i popytem na pieniądz. Na tym właśnie polega tzw. monetarne podejście do kursu walutowego (*monetary approach to the exchange rate*). Rozwiązanie to zasadniczo abstrahuje od występowania krótkookresowych sztywności cen i zakłada pełną giętkość w dostosowaniach cenowych, co implikuje analizę długookresową.

[34] Różnica między PPP a prawem jednej ceny (PJC) polega na tym, że PJC odnosi się do pojedynczego produktu, podczas gdy PPP dotyczy całego szerokiego koszyka wszystkich dóbr.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Załóżmy, że w długim okresie spełniony jest warunek PPP opisany wzorem (6.28). Z równania równowagi na rynku pieniądza wynika, że poziom cen krajowych $P = M_s / L(r, y)$, a poziom cen za granicą $P^* = M_s^* / L^*\left(r^*, Y^*\right)$, gdzie M_s oznacza podaż pieniądza, L(.) jest funkcją popytu na pieniądz zależną od dwóch zmiennych – stopy procentowej r i dochodu realnego Y, z kolei znak * wskazuje na zagranicę. Łącząc równania cen ze wzorem (6.28) otrzymujemy podstawowe równanie podejścia monetarnego:

$$e = P/P^* = \left(M_s / M_s^* \right) \cdot \left[L\left(r^*, Y^*\right) / L(r, Y) \right]. \tag{6.31}$$

Powyższe równanie oznacza, że kurs walutowy jest w długim okresie wyznaczany przez relacje podaży i popytu na pieniądz realny w kraju i za granicą. Równanie (6.31) można przekształcić w sposób następujący (w przypadku względnie małych wartości r):

$$e = \left(M_s / M_s^* \right) \cdot \varphi\left[r - r^*, Y^r / Y \right], \tag{6.32}$$

gdzie $\varphi\left[r - r^*, Y^r / Y \right]$ stanowi względny popyt na pieniądz realny za granicą w relacji do poziomu popytu w kraju.

Z równania (6.32) wynikają następujące wnioski. Po pierwsze, wzrost podaży pieniądza w kraju (lub spadek podaży pieniądza za granicą) wywołuje proporcjonalną deprecjację waluty krajowej. Oznacza to, że w długim okresie wartość nominalnego kursu zmienia się wraz ze zmianami (względnymi) zachodzącymi w zakresie podaży pieniądza krajowego. Po drugie, wzrost relacji Y^* / Y zwiększa relatywnie popyt na pieniądz za granicą, co powoduje wzrost wyrażenia $\varphi\left[r - r^*, Y^r / Y \right]$ i pociąga za sobą deprecjację waluty krajowej. Po trzecie, wzrost różnicy między stopami procentowymi w kraju i za granicą relatywnie zwiększa także popyt na pieniądz za granicą, co przekłada się na wzrost wyrażenia $\varphi\left[r - r^*, Y^r / Y \right]$ i wywołuje deprecjację waluty krajowej. Również odwrotnie, relatywny spadek podaży pieniądza krajowego, spadek różnicy między stopami procentowymi i relatywny spadek dochodu krajowego powodują aprecjację waluty krajowej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ostatni z przedstawionych wniosków może wydawać się zaskakujący. Jak pamiętamy, zgodnie z równaniem parytetu stóp procentowych, wzrost krajowych stóp procentowych względem stóp zagranicznych wywoływał aprecjację waluty krajowej. Tymczasem z równania (6.32) wynika, że relatywny wzrost krajowej stopy procentowej pociąga za sobą deprecjację waluty krajowej. Opisany paradoks jest jednak pozorny i wynika z przyjęcia odmiennych założeń. Pozwala to wyjaśnić tzw. efekt Fishera oraz rozróżnienie między długim i krótkim okresem analizy (zob. ramka 6.10). W długim okresie, zgodnie z monetarnym podejściem do kursu walutowego, wzrost krajowej stopy procentowej względem stopy zagranicznej może mieć miejsce tylko wtedy, gdy oczekiwana inflacja krajowa jest wyższa niż oczekiwana inflacja za granicą, co z kolei zakłada szybszy wzrost podaży pieniądza w kraju niż za granicą. Szybszy wzrost cen krajowych powoduje więc deprecjację kursu, zgodnie z równaniem (6.29), oraz – równolegle – wzrost stóp procentowych. Tymczasem równanie parytetu stóp procentowych odnosi się do krótkiego okresu, w którym ceny pozostają niezmienione (sztywne). W tych warunkach w krótkim okresie krajowe stopy procentowe mogą wzrosnąć wówczas, gdy podaż pieniądza *spadnie*, ponieważ przy stałych cenach popyt na pieniądz będzie wyższy niż jego aktualna podaż. Gdyby ceny były natomiast giętkie – tak jak ma to miejsce w długim okresie – uległyby obniżeniu, a w konsekwencji nastąpiłby także spadek popytu na pieniądz. Wzrost stóp procentowych okazałby się w tej sytuacji niepotrzebny.

Ramka 6.10. Pozorny paradoks monetarnego podejścia do kursu walutowego. Efekt Fishera*

Pozorny paradoks polega na tym, że, zgodnie z równaniem parytetu stóp procentowych, relatywny wzrost krajowych stóp procentowych powoduje aprecjację waluty krajowej, podczas gdy, zgodnie z monetarnym podejściem do kursu walutowego (równanie 6.32), relatywny wzrost krajowej stopy procentowej pociąga za sobą deprecjację waluty krajowej. Tę pozorną sprzeczność można wyjaśnić przy pomocy tak zwanego efektu Fishera. Jak pamiętamy (punkt 6.4 w niniejszym rozdziale), efekt Fishera polega na tym, że w długim okresie, przy innych założeniach stałych, wzrost oczekiwanej inflacji w danym kraju wywołuje stopniowy wzrost nominalnych stóp procentowych. W efekcie ani poziom produkcji, ani realna stopa procentowa nie ulegają zmianie (Fisher 1930). Zapiszmy oczekiwaną stopę inflacji w ciągu roku jako:

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$\pi^e = \frac{\left(P^e - P\right)}{P}.$$

Jeśli spełniona jest względna wersja warunku PPP zgodnie z równaniem (6.30), to oczekiwana stopa deprecjacji waluty krajowej wobec waluty zagranicznej równa się różnicy oczekiwanych stóp inflacji w kraju i za granicą w danym okresie:

$$\frac{e^p - e}{e} = \pi^e - \pi^{e*}. \tag{6.33}$$

Przypomnijmy, że parytet stóp procentowych (bez pokrycia) ma następującą postać (równanie 5.67):

$$r = r^* + \frac{e^p - e}{e}.$$

Łącząc równanie (6.33) z równaniem parytetu, otrzymujemy:

$$r - r^* = \pi^e - \pi^{e*}. \tag{6.34}$$

Oznacza to, że jeśli spełniony jest warunek PPP, to w długim okresie oczekiwana stopa deprecjacji kursu równa się różnicy oczekiwanych stóp inflacji (równanie 6.34), a w konsekwencji różnica stóp procentowych musi być równa różnicy oczekiwanych stóp inflacji (równanie 6.33). Równanie (6.34) opisuje efekt Fishera oraz przedstawia długookresową zależność kursu walutowego i stóp procentowych (Fisher 1930). Zgodnie z tą regułą oraz po uwzględnieniu wcześniej omówionego związku między podażą pieniądza a oczekiwaną inflacją, względny wzrost podaży pieniądza krajowego pociąga za sobą wzrost oczekiwanej inflacji, co przekłada się natomiast zarówno na wzrost krajowych stóp procentowych, jak i na deprecjację kursu waluty krajowej.

Teoria PPP powstała w latach 20. (Cassel 1921), lecz przez wiele lat nie znajdowała potwierdzenia w badaniach empirycznych (Marston 1987, Krugman i Obstfel 1991). Dopiero w latach 90. wyniki nowszych doświadczeń wykazały, że istnieje pewna – choć niezbyt silna – tendencja do przemieszczania się kursu rynkowego w kierunku PPP. Rogoff (1996) oszacował, że dostosowanie to przebiega wolno,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

lecz jest wyraźne – w okresie 3–5 lat różnica między kursem realnym a jego wartością długookresową zmniejsza się przeciętnie o połowę.

Istnieje kilka czynników, które sprawiają, że dostosowanie kursu bieżącego do PPP okazuje się w najlepszym przypadku powolne i długotrwałe. Należą do nich: a) koszty transportu i inne bariery w handlu (cła, kontyngenty, odmienne standardy techniczne czy sanitarne), które powodują utrzymywanie się różnic między cenami na rynkach różnych krajów; b) monopole, oligopole oraz inne bariery wejścia i wyjścia funkcjonujące na rynkach; c) różne (nieidentyczne) „koszyki" towarów uwzględniane w badaniach porównawczych oraz d) duży udział towarów niewymiennych w indeksach cen.

Zwłaszcza ten ostatni czynnik odgrywał w swoim czasie niemałą rolę. Ekonomiści od dawna zauważali, że poziom cen w krajach biedniejszych, przeliczony na wspólną walutę po kursie rynkowym, jest niższy niż w krajach bogatych. Zjawisko to próbowano wyjaśniać na różne sposoby – dwa najbardziej znane z nich to teoria Kravisa-Lipseya oraz teoria Balassy-Samuelsona. Obie dowodziły, że o ile poziom cen towarów wymiennych (*tradables*) wyrównuje się między krajami, o tyle poziom cen towarów niewymiennych (*nontradables*), zwłaszcza usług, w krajach uboższych jest niższy niż w krajach bogatych. W rezultacie średni poziom cen całego „koszyka" dóbr i usług w krajach uboższych staje się niższy.

6.8.3. Efekt Balassy-Samuelsona

Handel zagraniczny w tendencji prowadzi do wyrównywania się cen towarów wymiennych w poszczególnych krajach. Ceny takich towarów jak samochody, telewizory, komputery, a także cement, ropa naftowa czy zboża po przeliczeniu na jednakową walutę pozostają do siebie zbliżone. Ewentualne różnice związane są z uiszczaniem cła, kosztami transportu lub rozbieżnościami w stawkach podatkowych. Z kolei ceny towarów niewymiennych, w tym zwłaszcza usług, są w krajach uboższych wyraźnie niższe niż w krajach zamożnych. W pracach opublikowanych w 1964 roku dwaj ekonomiści – Bela Balassa i Paul Samuelson – niezależnie od siebie wykazali, że zjawisko to jest spowodowane niższą wydajnością pracy w krajach rozwijających się w porównaniu z krajami rozwiniętymi. W szczególności zaobserwowali oni, że gdy płace jednostkowe wyrównują się w poszczególnych państwach, a międzynarodowe różnice dotyczące wydajności pracy w zakresie produkcji towarów niewymiennych są mniejsze niż w przypadku produkcji towarów wymiennych, towary niewymienne, zwłaszcza usługi, będą relatywnie droższe

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w krajach zamożniejszych (Balassa 1964, Samuelson 1964). W konsekwencji, gdy kraj uboższy rozwija się szybciej dzięki szybszym przyrostom wydajności pracy w zakresie produkcji towarów wymiennych, spowodowany tym szybszy wzrost płac w całej gospodarce wywołuje wyższy wzrost cen towarów niewymiennych, a w rezultacie także wyższą średnią inflację. Zjawisko to nosi nazwę efektu Balassy-Samuelsona (dalej B-S).

Punktem wyjścia dla Balassy i Samuelsona stało się spostrzeżenie, że tempo wzrostu wydajności pracy w sektorach wymiennych (ang. *tradeable*) jest wyższe niż w sektorach niewymiennych (ang. *non-tradeable*). Oznacza to, że jeśli płace w sektorach wymiennych podążają za wzrostem wydajności, tendencja do wyrównywania płac między sektorami w obrębie danej gospodarki spowoduje równie szybki wzrost płac w sektorze niewymiennym. Ponieważ wydajność pracy w tym sektorze rośnie jednak wolniej, wzrost płac musi wywołać wzrost cen w tym sektorze. Dlatego też ceny usług, które w większości nie podlegają wymianie międzynarodowej, rosną szybciej od cen pozostałych towarów.

Ramka 6.11. Efekt Balassy-Samuelsona

Mechanizm działania efektu B-S można wyjaśnić przy pomocy prostego dwusektorowego, jednoczynnikowego modelu gospodarki otwartej (Rosati 2001). Załóżmy, że pierwszy sektor wytwarza towary wymienne (*tradables*), a drugi towary niewymienne (*nontradables*). Przyjmijmy następnie, że produkcja jest funkcją jednego czynnika produkcji, czyli $Y_T = f(L_T)$, $Y_N = g(L_N)$, gdzie L_T i L_N stanowią zasoby pracy występujące odpowiednio w sektorze wymiennym i niewymiennym. Wówczas poziom cen krajowych P można przedstawić jako średnią ważoną poziomu cen towarów wymiennych P_T i niewymiennych P_N, przy czym α oznacza udział sektora niewymiennego w PKB:

$$P = (1-\alpha)P_T + \alpha P_N. \tag{6.35}$$

W warunkach konkurencji jednostkowe płace nominalne W są równe krańcowym produktom pracy pomnożonym przez ceny. Hipoteza B-S zakłada, że za sprawą pełnej mobilności siły roboczej płace wyrównują się w obu sektorach. Przyjmując w celu uproszczenia liniowe zależności, otrzymujemy produkt krańcowy

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

pracy, który w poszczególnych sektorach równy jest produktowi przeciętnemu, tj. wydajności pracy:

$$P_i\left[\delta Y_i / \delta L_i\right] = P_i\left[Y_i / L_i\right] = W, \text{ dla } i = T, N$$

czyli:

$$P_T\lambda_T = P_N\lambda_N = W, \tag{6.36}$$

gdzie $\lambda_T = Y_T / L_T$ i $\lambda_N = Y_N / L_N$ oznaczają średnie wydajności pracy kolejno w sektorze wymiennym i niewymiennym. Cały dochód przypada zasobowi pracy jako jedynemu czynnikowi. W modelu bardziej rozwiniętym z dwoma czynnikami (kapitał i praca) całkowita produktywność (ang. *total factor productivity – TFP*) jest sumą krańcowych produktów kapitału i pracy. W takim modelu wynagrodzenie czynnika pracy będzie z reguły niższe niż w modelu jednoczynnikowym[35]. Aby nadmiernie nie komplikować analizy, pozostaniemy przy tym ostatnim modelu[36]. Zmiana poziomu cen *dP* jest równa:

$$dP = (1-\alpha)dP_T + \alpha dP_N. \tag{6.37}$$

[35] Większą liczbę czynników produkcji można bez kłopotu uwzględnić, przyjmując założenie, że stronę podażową gospodarki opisują dwie sektorowe funkcje produkcji typu Cobba-Douglasa o następującej postaci (Froot i Rogoff 1995, Cipriani 2000, Rotter 2000, Rosati 2002):

$$Y_i = A_i K_i^{(1-\beta)} L_i^{\beta}, \text{ dla } i = T, N \tag{6.38}$$

gdzie Y oznacza produkcję, K jest zasobem kapitału, L zasobem pracy, a wskaźniki T i N oznaczają kolejno sektor wymienny i niewymienny. Różniczkując wzór (6.38) względem L i K, otrzymujemy równania określające poziom płac nominalnych w warunkach konkurencji. Płace W będą równe krańcowym produktom pracy w każdym sektorze, gdzie P_t, P_N to odpowiednio ceny towarów wymiennych i niewymiennych:

$$W_T = P_T \frac{\partial Y_T}{\partial L_T} = P_T \beta_T A_T \left(\frac{K_T}{L_T}\right)^{1-\beta_T} \tag{6.39}$$

oraz

$$W_N = P_N \frac{\partial Y_N}{\partial L_N} = P_N \beta_N A_N \left(\frac{K_N}{L_N}\right)^{1-\beta_N}.$$

Gdy rynek pracy jest konkurencyjny, a siła robocza mobilna, płace w obu sektorach wyrównują się, czyli $W_T = W_N$.

[36] Wybór modelu nie ma znaczenia w przypadku określania wielkości efektu B-S, choć pozostaje istotny przy wyznaczaniu proporcji, w jakich efekt B-S rozkłada się między inflację cen krajowych i aprecjację nominalnego kursu waluty krajowej. W modelu jednoczynnikowym całkowity przyrost wydajności pracy przekłada się na wzrost płac, a efekt B-S występuje w formie inflacji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W systemie stałego kursu walutowego i przy braku ograniczeń w handlu ceny towarów wymiennych są w pełni określone przez ceny rynku międzynarodowego, na mocy prawa „jednej ceny" (*the law of one price*). Ponieważ celem naszej analizy jest jednak oszacowanie „dodatkowej" inflacji wywołanej przez efekt B-S, możemy chwilowo założyć (bez uszczerbku dla ogólności rozważań), że ceny towarów wymiennych nie ulegają zmianie, a więc $dP_T = 0$. Różniczkując równanie (6.36) i przekształcając je, otrzymujemy zmianę cen towarów niewymiennych dP_N[37]:

$$dP_N = P_N\left[r_{\lambda_T} - r_{\lambda_N}\right], \tag{6.40}$$

gdzie r_{λ_T} i r_{λ_N} oznaczają stopy wzrostu wydajności pracy odpowiednio w sektorze wymiennym i niewymiennym. Podstawiając otrzymany wzór (6.40) do równania (6.37), a następnie przyjmując, że $dP_T = 0$, oraz dzieląc obie strony przez P, otrzymujemy wzór na „dodatkową" inflację, czyli efekt B-S:

$$\pi = \alpha\left(r_{\lambda_T} - r_{\lambda_N}\right)\frac{P_N}{P}, \tag{6.41}$$

gdzie $\pi = \frac{dP}{P}$. Wielkość efektu B-S zależy więc od trzech czynników: udziału sektora niewymiennego w gospodarce (parametr α), różnicy w tempie wzrostu wydajności pracy między sektorami oraz relatywnych cen towarów niewymiennych w stosunku do ogólnego poziomu cen. Równanie (6.41) można alternatywnie zapisać jako:

$$\pi = \alpha\left(r_W - r_{\lambda_N}\right)\frac{P_N}{P}, \tag{6.41a}$$

o ile tempo wzrostu płac r_W jest równe tempu wzrostu wydajności pracy w sektorze wymiennym r_{λ_T}.

W warunkach sprawnie rozwijającej się gospodarki otwartej efekt B-S powoduje szybszy wzrost ogólnego poziomu cen. Kraje znajdujące się na niższym poziomie rozwoju osiągają na ogół wyższe tempo wzrostu wydajności pracy w sektorach wymiennych niż kraje rozwinięte, ponieważ sprowadzają i imitują istniejącą już

[37] Różniczkując wzór (6.36) i dzieląc obustronnie przez $W\lambda_N$, otrzymujemy $\left[dW/W\right]\left(1/\lambda_N\right) = \left(dP_N/W\right) + \left(P_N/W\right)\left(d\lambda_N/\lambda_N\right)$, a następnie podstawiamy $dP_N = W/\lambda_N$ i rozwiązujemy dla dP_N.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

technologię oraz importują kapitał, co umożliwia szybszy wzrost technicznego uzbrojenia pracy. Z kolei ceny towarów wymiennych zależą od wartości cen międzynarodowych – na mocy prawa „jednej ceny". Wzrost wydajności pracy pociąga za sobą wzrost płac realnych najpierw w sektorach wymiennych, a następnie w sektorach niewymiennych. Wywołuje to szybszy niż za granicą wzrost cen towarów niewymiennych i ogólnie wyższą inflację niż za granicą. Zauważmy, że ta „dodatkowa" inflacja, wywołana efektem B-S, nie ma nic wspólnego z typowymi źródłami inflacji, takimi jak nadmierny deficyt budżetowy lub zbyt szybka akcja kredytowa banków, a jest wyłącznie wynikiem przyspieszonego wzrostu wydajności pracy, który w całości lub w części przekłada się na wzrost płac realnych (zob. ramka 6.11).

Efekt B-S mierzy „dodatkową" inflację, notowaną powyżej inflacji będącej skutkiem oddziaływania standardowych czynników pieniężno-popytowych. Tę „dodatkową" inflację można zapisać w postaci równania (6.41) (zob. wyprowadzenie wzoru w ramce 6.10):

$$\pi = \alpha\left(r_W - r_{\lambda_N}\right)\frac{P_N}{P}.$$

Jak pokazano w ramce 6.10, na „dodatkową" inflację mają wpływ trzy czynniki. Parametr α uzależnia efekt B-S od udziału sektora niewymiennego w gospodarce. Im niższy będzie stopień międzynarodowego otwarcia gospodarki, tym wyższy stanie się efekt B-S. Duże gospodarki posiadają rozległe rynki wewnętrzne i udział sektora niewymiennego jest tam z reguły wyższy niż w małych otwartych gospodarkach. Efekt B-S będzie więc zasadniczo wyższy w przypadku gospodarek dużych i mniej otwartych. Drugim czynnikiem kształtującym „dodatkową" inflację jest różnica między tempem wzrostu płac a tempem wzrostu wydajności pracy w sektorze niewymiennym – im szybciej rosną średnie płace w relacji do wydajności pracy w sektorze niewymiennym, tym wyższa staje się inflacja. Innymi słowy, im wyższe okaże się tempo wzrostu wydajności pracy w sektorze wymiennym, tym silniejsze będzie oddziaływanie efektu B-S. Taka sytuacja występuje w krajach „doganiających" (*catching-up countries*) kraje rozwinięte. Dzięki temu, że mogą one importować oraz imitować nowe technologie z krajów rozwiniętych, są w stanie osiągać wysokie przyrosty wydajności pracy. Wreszcie, o „dodatkowej" inflacji decydują także ceny względne: im niższe stają się ceny towarów niewymiennych w relacji do cen towarów wymiennych, tym niższy będzie nominalny udział sektora niewymiennego w całkowitej produkcji oraz

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

tym niższa okaże się „dodatkowa" inflacja. Taka właśnie sytuacja ma miejsce w krajach rozwijających się.

Ceny towarów niewymiennych w danej gospodarce zależą od popytu wewnętrznego. Ponieważ w krajach ubogich siła nabywcza ludności jest niska, towary niewymienne (np. usługi) są generalnie tanie, co sprawia, że udział tego sektora w PKB pozostaje niski. W tych warunkach efekt B-S okazuje się także niewielki. Lecz w miarę rozwoju gospodarczego relatywne ceny towarów niewymiennych rosną, a ich stopień „niedowartościowania" maleje.

Szybszy wzrost cen krajowych na skutek oddziaływania efektu B-S w krajach „doganiających" pozostaje zjawiskiem naturalnym i nie oznacza spadku konkurencyjności. Jest przejawem szybszego wzrostu wydajności pracy w sektorze wymiennym, co oznacza szybsze tempo wzrostu gospodarczego. Nie należy więc przeciwdziałać mu za pomocą restrykcyjnej polityki pieniężnej. Wyższa inflacja będąca skutkiem efektu B-S może jednak utrudniać konwergencję nominalną w ramach unii walutowej. Jak wiemy, w procesie akcesji do strefy euro państwa członkowskie UE muszą między innymi spełnić kryterium niskiej inflacji. Jest ono sformułowane w sposób bardzo wymagający[38] i państwa członkowskie doświadczające efektu B-S mogą mieć duże trudności z jego urzeczywistnieniem (Rosati 2002).

6.8.4. Efekt Balassy-Samuelsona a zmiany kursu walutowego

Stosunek $\frac{P_N}{P}$ w równaniu (6.41) odzwierciedla relatywną różnicę między cenami towarów niewymiennych i ogólnym poziomem cen. Stosunek ten może być przedstawiony jako funkcja różnicy między kursem parytetu siły nabywczej (PPP) a kursem rynkowym. Niech e_m oznacza nominalny rynkowy kurs walutowy, a e_p kurs walutowy PPP (oba wyrażone w jednostkach waluty obcej na jednostkę waluty krajowej). Wówczas:

$$e_m = P_T^* / P_T \tag{6.42}$$

oraz

$$e_p = P^* / P, \tag{6.43}$$

gdzie P^* oznacza ceny obowiązujące na rynku międzynarodowym. W celu uproszczenia wywodu normalizujemy ceny zagraniczne na poziomie równym jedności,

[38] Kryteria konwergencji nominalnej zostały omówione w rozdziale 8.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

czyli $P_T^* = P_N^* = P^* = 1$[39]. Podstawiając równanie (6.37) do wzoru (6.43) i porządkując, otrzymujemy[40]:

$$\frac{P_N}{P} = \frac{1}{\alpha} - \left(\frac{(1-\alpha)}{\alpha}\right)\left(\frac{e_P}{e_m}\right). \tag{6.44}$$

Wstawiając zaś równanie (6.44) do wzoru (6.41), uzyskujemy ostatecznie wzór na „dodatkową" inflację wywołaną efektem B-S:

$$\pi = \left(r_W - r_{\lambda_N}\right)\left[1-(1-\alpha)\frac{e_P}{e_m}\right]. \tag{6.45}$$

Jak widać, siła oddziaływania efektu B-S zależy od trzech czynników. Po pierwsze, im większa jest różnica między średnim tempem wzrostu płac i tempem wzrostu wydajności pracy w sektorze niewymiennym, tym silniejszy staje się efekt B-S[41]. Po drugie, im wyższe wartości notuje udział sektora niewymiennego α w PKB, tym silniejszy wpływ wywiera efekt B-S. Po trzecie zaś, im mniejszy jest stosunek kursu PPP do kursu rynkowego, tym większy okazuje się udział efektu B-S. Ta ostatnia zależność potwierdza wcześniejsze obserwacje. Im większa jest różnica w poziomie rozwoju między krajem rozwiniętym a krajem „doganiającym", tym niższy staje się stosunek $\frac{e_P}{e_m}$ i tym wyższa „dodatkowa" inflacja.

Gdy płace w kraju „doganiającym" rosną wolniej niż wydajność pracy w sektorze wymiennym, ogólny poziom cen także będzie wzrastał wolniej, a wówczas „dodatkowa" inflacja okaże się niższa niż siła oddziaływania efektu B-S. W takiej sytuacji część przyrostu wydajności pracy, stanowiąca różnicę miedzy (wyższym) wzrostem wydajności pracy a (niższym) wzrostem płac, ujawni się w formie aprecjacji *nominalnego* kursu walutowego (zob. ramka 6.12).

[39] Możemy tak uczynić, ponieważ ceny międzynarodowe są punktem odniesienia, a poszukujemy odchyleń krajowych cen od cen międzynarodowych.

[40] Wykorzystujemy równania (6.42) i (6.43) oraz podstawiamy $P_T^* = P_N^* = P^* = 1$.

[41] Ponieważ $r_W = r_{\lambda_T}$, efekt B-S będzie tym większy, im szybciej wrośnie wydajność w sektorze wymiennym w porównaniu z wydajnością pracy w sektorze niewymiennym.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 6.12. Wymienność między inflacją a aprecjacją nominalnego kursu walutowego w warunkach oddziaływania efektu Balassy-Samuelsona

Załóżmy, że ceny zagraniczne utrzymują się na stałym poziomie (inflacja za granicą jest równa zeru), co oznacza, że w warunkach zmiennego kursu walutowego ceny krajowe P_T zmieniają się wraz z kursem nominalnym. Oznacza to, że $dP_T \neq 0$. Różniczkując całkowicie równanie (6.36) i przekształcając, otrzymujemy wzór na zmiany cen towarów niewymiennych[42]:

$$\pi_N = \pi_T + \left(r_{\lambda_T} - r_{\lambda_N}\right), \tag{6.46}$$

gdzie π_T, π_N oznaczają inflację odpowiednio w sektorze dóbr wymiennych i niewymiennych. Na podstawie równania (6.37) można przyjąć, że ogólny wskaźnik wzrostu cen π jest średnią ważoną sektorowych wskaźników cen π_T i π_N:

$$\pi = (1-\alpha)\pi_T + \alpha\pi_N. \tag{6.47}$$

Podstawiając wzór (6.46) równania do (6.47), otrzymujemy:

$$\pi = (1-\alpha)\pi_T + \alpha\left(\pi_T + r_{\lambda_T} - r_{\lambda_N}\right). \tag{6.48}$$

W przypadku $P_T^* = 1 = \text{const}$, zmiany cen towarów wymiennych są ściśle określone przez zmiany kursu nominalnego. Wówczas na podstawie równania (6.42) $\pi_T = -r_{e_m}$, gdzie r_{e_m} jest stopą zmian kursu nominalnego waluty obcej. Podstawiając do (6.48), otrzymujemy:

$$\alpha\left(r_{\lambda_T} - r_{\lambda_N}\right) = \pi - r_{e_m}. \tag{6.49}$$

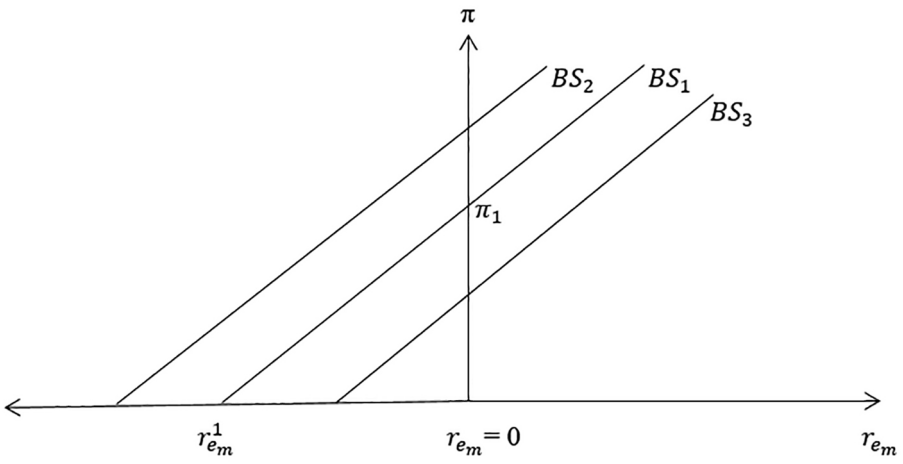
Powyższe równanie (6.49) ukazuje wymienność między inflacją a aprecjacją kursu nominalnego w warunkach występowania efektu B-S. Wobec danej wielkości efektu B-S (lewa strona równania) im niższa staje się inflacja π, tym większa musi być aprecjacja kursu r_{e_m} (pamiętajmy, że e_m wyraża stosunek cen zagranicznych do

[42] Różniczkując, otrzymujemy $dP_T\lambda_T + P_T d\lambda_T = dP_N\lambda_N + P_N d\lambda_N$. Dzieląc zaś obustronnie przez $P_T P_N$ oraz przez $\lambda_T\lambda_N$, a następnie porządkując i pamiętając, że $P_T\lambda_T = P_N\lambda_N$, otrzymujemy równanie (6.46).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

krajowych, a zatem $r_{e_m} < 0$ oznacza deprecjację waluty zagranicznej i aprecjację waluty krajowej). W skrajnym przypadku zerowej inflacji ($\pi = 0$), cały efekt B-S przejawia się w aprecjacji kursu.

Rysunek 6.17. Efekt Balassy-Samuelsona: wymienność między inflacją a nominalną aprecjacją kursu



Źródło: opracowanie własne.

Wymienność między inflacją i aprecjacją ukazana jest na rysunku 6.17, gdzie na osi poziomej odłożono zmiany kursu, a na osi pionowej inflację. Równanie (6.49) ilustrują linie *BS* o nachyleniu 45°. W przypadku danej wielkości efektu B-S, na przykład dla $\alpha\left(r_{\lambda_T} - r_{\lambda_N}\right) = BS_1$, możliwe kombinacje inflacji i aprecjacji znajdują się na linii BS_1. W warunkach stałego kursu efekt B-S będzie wywoływał inflację na poziomie π_1. Z kolei w warunkach kursu płynnego, w zależności od stopnia restrykcyjności polityki pieniężnej, wobec inflacji mniejszej od π_1 wystąpi aprecjacja kursu nominalnego, a w przypadku inflacji wyższej od π_1 będzie miała miejsce deprecjacja kursu. W skrajnym przypadku, gdzie $\pi = 0$, aprecjacja kursu wyniesie $r^1_{e_m}$. Wobec danej różnicy sektorowych stóp wzrostu wydajności pracy im wyższa staje się wartość α, tym wyżej położona jest linia *BS*. Podobnie dzieje się w przypadku danej wartości α – im większa okaże się różnica sektorowych stóp wzrostu wydajności pracy, tym wyżej położona będzie linia *BS*. Linia BS_2 odzwierciedla silniejszy efekt B-S, z kolei linia BS_3 ilustruje jego słabsze oddziaływanie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Trzeba jednocześnie podkreślić, że w obu przypadkach będziemy mieli do czynienia z aprecjacją *realnego* kursu walutowego. Pamiętamy, że przy danych cenach zagranicznych, kurs realny stanowi funkcję kursu nominalnego i cen krajowych. Oznaczmy tym razem kurs nominalny jako cenę jednostki waluty obcej wyrażoną w walucie krajowej, e_N[43]. Wówczas – na podstawie równania (6.1) – kurs realny równy jest:

$$e_R = e_m \left(\frac{P^*}{P} \right).$$

Efekt B-S oznacza więc albo „dodatkową" inflację ponad poziom inflacji zagranicznej, czyli wzrost P w stosunku do P^*, albo aprecjację kursu nominalnego, czyli spadek e_m. W obu przypadkach wartość e_R maleje, a zatem kurs realny waluty krajowej ulega aprecjacji.

6.8.5. Wpływ zmian podaży i popytu na kurs walutowy w długim okresie

Jak podkreślaliśmy kilkakrotnie, mimo że w tendencji rzeczywisty kurs walutowy zbliża się do kursu PPP, w procesie tym mogą występować długotrwałe odchylenia od parytetu siły nabywczej. Są one spowodowane czynnikami strukturalnymi, a także różnym tempem zmian zachodzących po stronie podaży i popytu na dobra i usługi w poszczególnych krajach.

Ogólnie rzecz biorąc, wzrost popytu w kraju A w relacji do wzrostu popytu w kraju B wywołuje proporcjonalnie większy wzrost absolutnych wydatków na produkty wytwarzane w kraju A niż wzrost absolutnych wydatków na produkty wytwarzane w kraju B. Dzieje się tak dlatego, że popyt w każdym kraju nakierowany jest przede wszystkim na produkty krajowe. Na przykład w wydatkach konsumpcyjnych polskich gospodarstw domowych udział produktów polskich jest większy niż udział produktów wytwarzanych w jakimkolwiek innym kraju (choć *suma* wydatków na import z innych krajów może być większa niż *suma* wydatków przeznaczanych na produkcję lokalną). Tak ukształtowana struktura wydatków sprawia, że nawet w przypadku proporcjonalnego wzrostu wydatków w kraju A na wszystkie produkty – zarówno krajowe, jak i importowane – w kategoriach bezwzględnych

[43] Pamiętajmy, że w dotychczasowej analizie kurs rynkowy e_m stanowił cenę jednostki waluty krajowej wyrażoną w walucie obcej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wzrost wydatków na produkcję krajową jest większy niż wzrost wydatków na towary z innych krajów. Oznacza to, jeśli wydatki rosną szybciej w kraju A niż w kraju B, ceny w kraju A będą rosły w relacji do cen obowiązujących w kraju B. W efekcie kurs realny waluty kraju A ulega (względnej) aprecjacji.

Odnosząc się zaś do podaży, przyjmijmy, że całkowita produktywność zasobów (łączna wydajność czynników produkcji) w kraju A wzrasta szybciej niż w kraju B. Ponieważ struktura wydatków w kraju A nie zmienia się (udział towarów krajowych i towarów importowanych jest stały), wzrost podaży towarów krajowych na skutek wzrostu produktywności powoduje spadek relatywnych cen towarów wytwarzanych w kraju A w stosunku do cen towarów wytwarzanych w kraju B. To zaś wywołuje wzrost popytu na dobra krajowe i przywraca równowagę na rynku towarów. Lecz skutkiem zmian cen relatywnych staje się również realna deprecjacja waluty kraju A w stosunku do waluty kraju B.

Gdy różnice w zakresie tempa zmian w kształtowaniu się popytu i podaży między poszczególnymi krajami utrzymują się dostatecznie długo, na rynkach walutowych obserwuje się długookresowe tendencje poszczególnych walut do aprecjacji lub deprecjacji w stosunku do innych walut. Przykładem długotrwałej realnej aprecjacji w stosunku do większości walut wymienialnych był kurs japońskiego jena w latach 70. i 80., którego umocnienie spowodowane było szybkim wzrostem produktywności zasobów w gospodarce japońskiej w tym okresie. Z kolei przykładem długotrwałej realnej deprecjacji stał się kurs funta brytyjskiego odnotowany w tym samym czasie.

6.8.6. Wpływ trwałych zmian w zakresie polityki makroekonomicznej na kurs walutowy w długim okresie

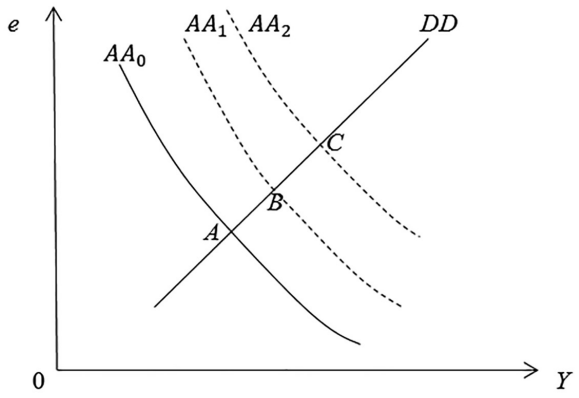
W przeciwieństwie do zmian przejściowych trwałe zmiany w zakresie polityki pieniężnej i fiskalnej wywierają wpływ nie tylko na bieżący kurs walutowy, ale także na kurs walutowy w długim okresie. Z kolei zmiany kursu w długim okresie kształtują bieżące oczekiwania co do poziomu kursu w przyszłości. Aby wyjaśnić skutki trwałego wzrostu podaży pieniądza na produkcję i kurs walutowy, sięgniemy ponownie po krzywe *DD* i *AA*, których konstrukcję opisaliśmy w punkcie 6.6.

Zbadajmy najpierw wpływ jednorazowego trwałego (permanentnego) wzrostu podaży pieniądza na gospodarkę. Jak wiadomo z wcześniejszych obserwacji, już przejściowy wzrost podaży pieniądza powoduje przesunięcie krzywej *AA* w prawo i wzrost bieżącego kursu waluty obcej (deprecjację waluty krajowej). Gdy wzrost

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

podaży pieniądza pozostaje jednak trwały, wpływa on także na wzrost cen i oczekiwanego kursu walutowego w przyszłości. W długim okresie procentowy wzrost kursu, podobnie jak wzrost wszystkich innych cen, staje się zaś równy procentowemu wzrostowi podaży pieniądza. Wynika z tego, że oczekiwana wartość kursu e^p rośnie, co – zgodnie z równaniem parytetu – wymaga silniejszej bieżącej deprecjacji. Oznacza to, że przesunięcie krzywej *AA* w prawo będzie większe niż w przypadku przejściowego wzrostu podaży pieniądza[44]. Na rysunku 6.18 krzywa *AA* przesuwa się więc najpierw do położenia AA_1, a następnie do AA_2. Punkt równowagi *C* charakteryzuje się jeszcze wyższym poziomem produkcji i kursu waluty niż punkt *B*, opisujący równowagę w przypadku przejściowego wzrostu podaży pieniądza.

Rysunek 6.18. Krótkookresowy efekt jednorazowego i trwałego wzrostu podaży pieniądza: wpływ oczekiwań kursowych



Źródło: opracowanie własne.

Przyjmijmy, że gospodarka początkowo znajduje się w równowadze, a poziom produkcji Y^F zapewnia pełne zatrudnienie. Uchylimy obecnie założenie o stałości cen, ponieważ rozpatrujemy proces dostosowawczy w długim okresie. Oznacza to, że trwały wzrost podaży pieniądza wywoła wzrost produkcji ponad poziom pełnego zatrudnienia Y^F, co w konsekwencji doprowadzi do wzrostu płac i cen. Wzrost cen w długim okresie pociąga za sobą dwa skutki. Po pierwsze, przy danym kursie obniża konkurencyjność produkcji krajowej, co powoduje spadek eksportu netto *X*

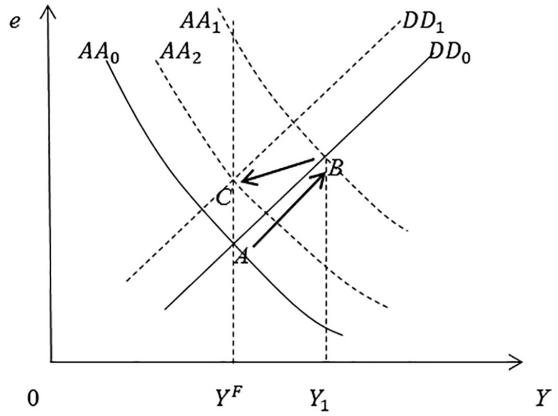
[44] Wzrost oczekiwanego kursu walutowego e^p w przyszłości powoduje wzrost oczekiwanej deprecjacji oraz przekłada się na wzrost stóp procentowych, co przesuwa krzywą *AA* w prawo (zob. punkt 6.6.4).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

oraz spadek produkcji. Efekt ten wywołuje więc przesunięcie krzywej *DD* w lewo. Po drugie, w miarę jak pod wpływem trwałego wzrostu podaży pieniądza następuje stopniowy wzrost płac i cen, realna podaż pieniądza systematycznie maleje. Efekt ten wpływa zatem na przesunięcie krzywej *AA* także w lewo.

Zmiany te zostały przedstawione na rysunku 6.19. W wyniku trwałego wzrostu podaży pieniądza gospodarka najpierw przesuwa się z punktu *A* do *B*, gdzie poziom produkcji Y_1 przekracza poziom potencjalny Y^F, a następnie powraca stopniowo do punktu *C*, gdzie produkcja zapewnia pełne zatrudnienie przy stabilnych cenach.

Rysunek 6.19. Długookresowy efekt trwałego wzrostu podaży pieniądza: wpływ wzrostu cen



Źródło: opracowanie własne.

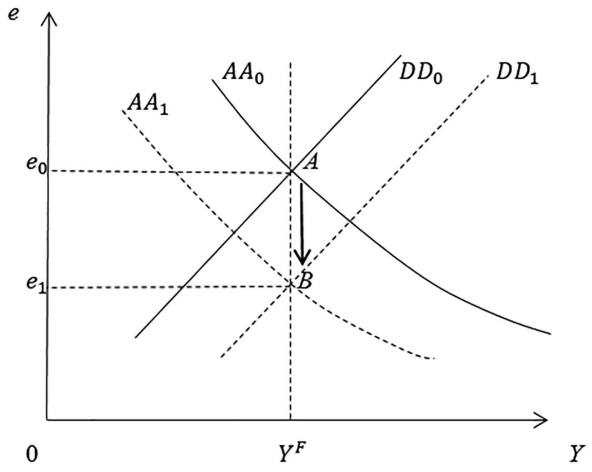
Zbadajmy obecnie wpływ jednorazowego trwałego wzrostu wydatków budżetowych na gospodarkę. Ekspansja fiskalna nie tylko przekłada się bezpośrednio na wzrost zagregowanego popytu, powodując przesunięcie krzywej *DD* w prawo, lecz kształtuje także oczekiwania dotyczące przyszłego poziomu kursu walutowego. Jak wykazaliśmy w punkcie 6.7.5, trwały wzrost wydatków (w relacji do innych krajów) wywołuje długookresową aprecjację kursu, co oznacza spadek e^p. To z kolei pociąga za sobą konieczność przeprowadzenia jeszcze silniejszej *bieżącej* aprecjacji kursu, aby przywrócić krótkookresową równowagę opisaną równaniem parytetu. W efekcie krzywa *AA* przesuwa się w lewo. Ostateczny rezultat tych zmian ukazany został na rysunku 6.20.

Początkowo gospodarka znajduje się w punkcie *A*, gdzie produkcja osiąga poziom potencjalny, zapewniający pełne zatrudnienie (Y^F). Permanentna ekspansja

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

fiskalna przesuwa krzywą DD do położenia DD_1, zwiększając produkcję ponad poziom potencjalny. Spadek oczekiwanego kursu e^p przekłada się jednak na przesunięcie krzywej AA do położenia AA_1, podczas gdy gospodarka stabilizuje się w punkcie B. Poziom produkcji powraca do poziomu potencjalnego, a kurs walutowy ulega trwałej aprecjacji z e_0 do e_1. Należy jednak zauważyć, że stałe zwiększanie wydatków rządowych nie przynosi trwałego skutku w postaci wzrostu produkcji. Efekt wzrostu zagregowanego popytu jest neutralizowany efektem aprecjacji kursu i w rezultacie produkcja utrzymuje się na niezmienionym poziomie. Gdy produkcja znajduje się zaś na poziomie potencjalnym, trwała ekspansja fiskalna staje się nieskuteczna w pobudzaniu produkcji.

Rysunek 6.20. Skutki trwałego wzrostu wydatków rządowych



Źródło: opracowanie własne.

6.9. Dewaluacja jako instrument polityki gospodarczej

6.9.1. Argumenty na rzecz dewaluacji

Dewaluacja jest świadomym obniżeniem przez władzę gospodarczą kursu własnej waluty wobec walut obcych w ramach systemu stałego kursu walutowego. Należy ją odróżniać od deprecjacji, która oznacza spadek kursu waluty krajowej względem walut obcych na rynku walutowym, dokonujący się w ramach systemu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

płynnego (rynkowego) kursu walutowego. Dlatego też dewaluacja, w odróżnieniu od deprecjacji, jest *instrumentem* polityki gospodarczej. Osłabienie kursu własnej waluty może być jednak dokonane także w systemie częściowo kontrolowanego kursu, na przykład poprzez świadomą interwencję banku centralnego, polegającą na zakupie walut obcych w zamian za walutę krajową. Efektem tych działań jest wymuszona deprecjacja kursu, która przynosi takie same skutki jak formalna dewaluacja.

W wyniku dewaluacji (deprecjacji) zagraniczne ceny towarów krajowych spadają, a ceny krajowe towarów zagranicznych rosną. Pojawia się wówczas tendencja do wzrostu eksportu i spadku importu o tyle, o ile elastyczności popytu i podaży są większe od zera. Gdy elastyczności popytu są dostatecznie wysokie – tak, że ich suma jest większa od jedności (a elastyczność podaży eksportowej pozostaje większa od zera), dewaluacja wpływa na poprawę bilansu płatniczego (zgodnie z warunkiem Marshalla-Lernera-Robinson – zob. punkt 6.6.1). Ten związek między osłabieniem kursu własnej waluty a strumieniami handlu zagranicznego stanowił tradycyjne uzasadnienie w przypadku wykorzystywania dewaluacji w polityce przywracania równowagi zewnętrznej danego kraju (zob. Hirschman 1949, Alexander 1952).

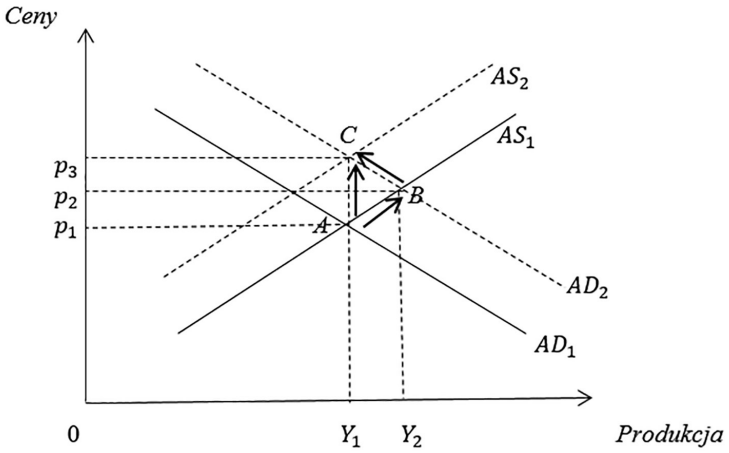
Inną jeszcze przyczyną zastosowania takiej strategii może być dążenie do przywrócenia równowagi wewnętrznej. Gdy kraj boryka się z wysokim bezrobociem lub chce przyspieszyć rozwój gospodarczy poprzez wzrost eksportu, obniżenie kursu własnej waluty może być narzędziem pobudzania koniunktury. Wywołany dewaluacją wzrost eksportu oznacza zwiększenie krajowej produkcji w sektorach eksportowych, podczas gdy spadek importu oznacza wzrost krajowej produkcji w sektorach konkurujących z importem. Przekłada się to na wzrost zatrudnienia i przyspieszenie wzrostu PKB. Zastosowanie dewaluacji okazuje się szczególnie atrakcyjne w przypadku krajów, które z różnych powodów mogą nie być w stanie pobudzić koniunktury innymi metodami, na przykład przy pomocy wzrostu wydatków rządowych lub rozluźnienia polityki pieniężnej.

Wpływ dewaluacji na równowagę wewnętrzną można przedstawić, wykorzystując w tym celu model AS-AD, co ilustruje rysunek 6.21 (będący częściową repliką rysunku 5.5). Początkowa równowaga wyznaczona jest przez punkt przecięcia krzywej łącznego popytu AD_1 i krzywej łącznej podaży AS_1 (punkt A). Dewaluacja pociąga za sobą wzrost popytu na produkcję krajową (zarówno eksportową, jak i konkurującą z importem), co oznacza przesunięcie się krzywej popytu do położenia AD_2. W efekcie rośnie produkcja i gospodarka przesuwa się z punktu A do punktu B. Nowemu punktowi równowagi odpowiada wyższy poziom produkcji,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gdzie $Y_2 > Y_1$, oraz wyższy poziom cen, kiedy $p_2 > p_1$. Skala wzrostu produkcji i skala wzrostu cen zależy od nachylenia krzywych *AD* i *AS*. Im bardziej płaska jest krzywa *AS* przy danej krzywej *AD*, tym większy staje się przyrost produkcji i tym mniejszy wzrost cen. Z kolei im bardziej płaska będzie krzywa *AD* przy danej krzywej *AS*, tym mniejszy okaże się przyrost produkcji i wzrost cen.

Rysunek 6.21. Wpływ dewaluacji na równowagę wewnętrzną

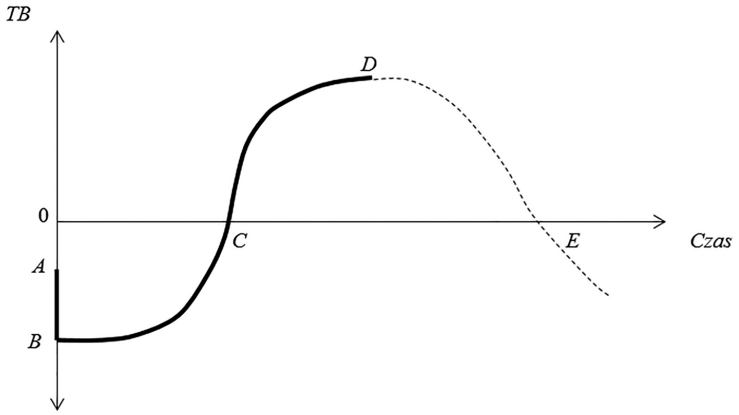


Źródło: opracowanie własne.

W obu przypadkach efekt dewaluacji nie jest natychmiastowy, lecz rozłożony w czasie. Obserwacje empiryczne pokazują, że bezpośrednio po dewaluacji deficyt w bilansie płatniczym rośnie, a produkcja spada. Dzieje się tak dlatego, że wpływ dewaluacji na import jest szybszy niż jej wpływ na eksport. Importerzy w krótszym czasie reagują na wzrost cen importowych w wyniku dewaluacji, redukując zakupy importowe. Z kolei po stronie eksporterów dostosowanie do nowych cen jest wolniejsze, ponieważ uzyskiwanie kolejnych zamówień i zwiększanie produkcji zajmuje więcej czasu. W rezultacie pozytywny wpływ dewaluacji na poprawę bilansu płatniczego ujawnia się zawsze z pewnym opóźnieniem. Zjawisko to nazywa się efektem „krzywej *J*" (*J-curve effect*), której przebieg został przedstawiony na rysunku 6.22 (zob. Salvatore, s. 560–561). Zakładając, że w punkcie wyjścia znajduje się umiarkowany deficyt w bilansie handlowym (punkt *A*), bezpośrednim skutkiem dewaluacji jest pogłębienie się deficytu handlowego (punkt *B*). W miarę upływu czasu bilans handlowy ulega stopniowej poprawie, najpierw powracając do stanu równowagi (punkt *C*), a następnie osiągając nadwyżkę (punkt *D*).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rysunek 6.22. Wpływ dewaluacji na bilans handlowy (*TB*) – „krzywa *J*"



Źródło: Salvatore (2004), s. 560.

6.9.2. Argumenty przeciwko dewaluacji

Dewaluacja była dość popularnym narzędziem polityki gospodarczej w drugiej połowie XX wieku, gdy w większości państw obowiązywały stałe kursy walutowe, a transgraniczne przepływy kapitału były niewielkie ze względu na ograniczenia w zakresie wymienialności walut. Dotyczyło to zwłaszcza krajów rozwijających się, które miały trudności z utrzymaniem równowagi płatniczej w warunkach niskiej konkurencyjności swoich gospodarek. Wraz z upowszechnianiem się płynnych kursów, wzrostem skali przepływów kapitałowych i pogłębianiem się integracji rynków finansowych w skali globalnej dewaluacja straciła jednak na znaczeniu. W rzeczywistości niesie ona bowiem za sobą szereg negatywnych zjawisk, których skumulowany efekt może całkowicie przekreślić sens jej funkcjonowania.

Po pierwsze, dewaluacja wcale nie musi pobudzać wzrostu produkcji i zatrudnienia. Jak wykazali m.in. Krugman i Taylor (1978) oraz Kasa (1998), w pewnych warunkach może ona nawet wywoływać recesję gospodarczą. Sytuacja ta wystąpi przede wszystkim na skutek oddziaływania efektu dochodowego – wzrost eksportu i produkcji konkurującej z importem w warunkach sztywnych (początkowo) płac powoduje wzrost zysków w tych sektorach, co oznacza transfer dochodów do grup o wyższej skłonności do oszczędzania. Wzrost oszczędności krajowych obniża zaś popyt zagregowany i pociąga za sobą spadek produkcji. Co więcej, gdy w punkcie wyjścia kraj notuje duży deficyt handlowy, wzrost cen towarów eksportowych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i importowanych przekłada się na spadek realnego dochodu w kraju, ponieważ płatności zagraniczne okazują się wyższe niż należności otrzymywane z zagranicy. W rezultacie popyt zagregowany spada.

Po drugie, pozytywne skutki w zakresie wzrostu gospodarczego – jeśli wystąpią – są krótkotrwałe. Zgodnie z zasadą neutralności pieniądza, mówiącą o tym, że przy pomocy zmiennych nominalnych nie można w sposób trwały wpłynąć na zmienne realne, dewaluacja staje się na dłuższą metę nieskuteczna. Ilustruje to rysunek 6.21. Choć dewaluacja początkowo powoduje wzrost produkcji z punktu *A* do *B*, to związany z tym wzrost krajowych cen importowych i eksportowych przekłada się na wzrost rzeczowych kosztów produkcji i spadek realnych dochodów ludności. Wywołuje to presję na wzrost wynagrodzeń, a w efekcie rosną także osobowe koszty produkcji. Wzrost kosztów sprawia zaś, że krzywa *AS* przesuwa się do położenia AS_2 – produkcja wówczas spada, a ceny rosną[45]. Gospodarka przesuwa się z punktu *B* do punktu *C*, w którym produkcja powraca do początkowego poziomu, lecz pułap cen jest znacznie wyższy ($p_3 > p_2$). Korzyść w postaci wzrostu produkcji okazuje się więc przejściowa. Ta krótkotrwałość efektu dewaluacji wynika z tego, że czyni ona produkcję krajową przejściowo *tańszą*, z kolei nie zwiększa wydajności pracy ani nie poprawia jakości produkcji. Jedynym trwałym skutkiem dewaluacji pozostaje więc wyższa inflacja.

Wzrost kosztów produkcji wywołany dewaluacją sprawia, że równie krótkotrwały jest dodatni wpływ tego zjawiska na równowagę zewnętrzną. Jak pokazano na rysunku 6.22, po pewnym czasie, gdy wzrost cen importu i eksportu przeniesie się w pełni na koszty produkcji i ceny krajowe, konkurencyjność międzynarodowa danego kraju obniża się i dodatnie saldo handlowe stopniowo maleje, przesuwając się z punktu *D* do punktu *E* (przejście zaznaczone linią przerywaną). Bilans handlowy kraju powraca wówczas do pozycji wyjściowej.

Po trzecie, każda dewaluacja jest kosztowna, ponieważ zwiększa oczekiwania kolejnej dewaluacji, co skłania uczestników rynku do podnoszenia cen i zwiększa premię za ryzyko kursowe w stopach procentowych. Częste stosowanie dewaluacji podważa wiarygodność polityki, a pojawienie się dodatkowych oczekiwań rodzi ponadto problem niespójności czasowej w jej zakresie (zob. rozdział 5). Uczestnicy rynku uwzględniają skutki polityki rządu w swoich zachowaniach. Gdy utrwalą się oczekiwania dewaluacyjne, przenoszą się one natychmiast na ceny krajowe i zanika

[45] Efekt kosztowo-cenowy dewaluacji jest silniejszy w małych otwartych gospodarkach, o dużym udziale sektora wymiennego w PKB.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

także krótkookresowy pozytywny efekt dewaluacji. Powracając do rysunku 6.21, w warunkach utrwalonych oczekiwań dewaluacyjnych mamy do czynienia z *jednoczesnym* przesunięciem krzywych *AD* i *AS* oraz przejściem gospodarki z punktu *A* bezpośrednio do punktu *C*. Oznacza to, że nie nastąpi żaden wzrost produkcji, nawet w krótkim okresie.

Po czwarte, dewaluacja może wywołać reakcje obronne ze strony partnerów handlowych, których pozycja konkurencyjna ulega pogorszeniu. Typową reakcją retorsyjną jest dokonanie „konkurencyjnej" dewaluacji przez partnera handlowego, tak aby przywrócić dotychczasowy parytet kursu walut. Wówczas kraj, który pierwszy zdewaluował swoja walutę, może ulec pokusie dokonania kolejnej dewaluacji, aby utrzymać przewagę konkurencyjną. Taka polityka wzajemnego „podcinania" konkurencyjności cenowej przy pomocy konkurencyjnych dewaluacji nosiła w pierwszej połowie XX wieku nazwę polityki „zubożania partnera" (*beggar-thy-neighbour policy* – zob. np. Krugman i Obstfeld 1991, s. 272). Ponieważ polega ona na podejmowaniu prób odniesienia korzyści kosztem partnerów handlowych, często prowadziła do „wojen walutowych" (*currency wars*) miedzy państwami (ramka 6.13). W takich warunkach dewaluacja nie przynosi żadnych korzyści, a wręcz przeciwnie – prowadzi do obniżania poziomu dobrobytu w krajach stosujących „konkurencyjne" dewaluacje.

Ramka 6.13. Wojny walutowe

Wojna walutowa polega na serii konkurencyjnych dewaluacji dokonywanych przez partnerów handlowych. Gdy poszczególne kraje chcą zwiększać udziały w rynku poprzez wzrost eksportu kosztem swoich partnerów, mogą podjąć działania zmierzające do obniżenia kursu własnej waluty wobec walut partnerów handlowych. Jeśli inne kraje odpowiedzą w ten sam sposób, mamy do czynienia z konkurencyjnymi dewaluacjami, czyli z wojną walutową. Skutkiem tego typu działań staje się ograniczenie skali handlu międzynarodowego i spadek korzyści handlowych w przypadku wszystkich uczestników.

Historycznie rzecz biorąc, najgłośniejszą wojną walutową była seria dewaluacji dokonana w okresie, który nastąpił po Wielkim Kryzysie w latach 30., gdy kraje rozwinięte odeszły od systemu waluty złotej i dewaluowały swoje waluty w dążeniu do pobudzenia produkcji i ograniczenia bezrobocia. Wywołało to działania retaliacyjne ze strony pozostałych państw, które broniły swojej pozycji konkurencyjnej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ekonomiczny bezsens stosowania takiej polityki w przekonujący sposób wyjaśniła brytyjska ekonomistka Joan Robinson (1937).

Kolejna odsłona wojny walutowej, tym razem już bardziej współczesna, miała miejsce w latach 2010–2012, bezpośrednio po kryzysie finansowym z lat 2008–2009. Objęła ona tym razem swoim zasięgiem kraje stosujące płynne lub płynne kontrolowane kursy walutowe. Pod wpływem oddziaływania bardzo luźnej polityki pieniężnej zastosowanej głównie przez Stany Zjednoczone (a także Wielką Brytanię i strefę euro) kurs dolara amerykańskiego w stosunku do walut innych krajów uległ znacznemu osłabieniu, co uderzyło w eksport tych państw. Na przykład w latach 2005–2010 efektywny nominalny kurs walutowy dolara kanadyjskiego umocnił się w stosunku do dolara amerykańskiego o 17%, chińskiego juana i dolara australijskiego o 21%, jena japońskiego o 25%, a brazylijskiego reala o ponad 38%. Sprowokowało to ówczesnego brazylijskiego ministra finansów Guido Mantegę do oskarżenia Stanów Zjednoczonych o rozpoczęcie „wojny walutowej"[46]. Mantegę poparło wówczas wielu przedstawicieli innych krajów, w tym m.in. Chin, Japonii, Malezji, Indonezji, a także Tajwanu.

Ten ostatni epizod zakończył się stosunkowo szybko, bo już w 2012 roku odpływ kapitału z wielu rynków „wschodzących" doprowadził do osłabienia ich walut w stosunku do dolara oraz euro. Jednocześnie kraje członkowskie grupy G7 i grupy G20 zadeklarowały w lutym 2013 roku, że będą unikać konkurencyjnych dewaluacji. W 2015 roku efektywne nominalne kursy walut większości krajów wobec dolara amerykańskiego (z wyjątkiem Chin) powróciły do poziomów zbliżonych do tych, które obserwowano w 2005 roku.

Wreszcie, jak wskazywał amerykański ekonomista Richard Cooper, dewaluacja okazuje się również politycznie bardzo kosztowna. Ponieważ wywołuje ona spadek realnych dochodów ludności i przyczynia się do wzrostu inflacji, staje się niemal zawsze kłopotliwym świadectwem słabości i symbolem porażki w zakresie polityki gospodarczej rządu (Cooper 1971). Dewaluacja podnosi także koszty obsługi długu denominowanego w walutach obcych, co w przypadku mocno zadłużonych krajów rozwijających się wymusza dodatkową inflację. Co więcej, dewaluacja może zniechęcać zagranicznych przedsiębiorców do inwestowania w danym kraju. Dlatego też, choć dewaluacja była stosowana relatywnie często, traktowano

[46] "Financial Times", 27 September, 2010.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ją głównie jako przykrą konieczność, na którą decydowano się w sytuacji, gdy dany kraj nie był w stanie wdrożyć innych instrumentów pobudzania popytu (na przykład w postaci luzowania polityki pieniężnej lub zwiększania deficytu fiskalnego) bądź nie znajdował innej możliwości poprawy konkurencyjności międzynarodowej.

6.10. Interwencje walutowe

6.10.1. Cele i zasady interwencji walutowych

Interwencją walutową nazywamy transakcję sprzedaży lub kupna walut obcych przez bank centralny na rynku walutowym. Jej celem jest najczęściej dążenie do utrzymania kursu walutowego na określonym poziomie lub w określonym przedziale wahań bądź (rzadziej) dążenie do utrzymania rezerw walutowych na określonym poziomie. Transakcje te określa się mianem interwencji, ponieważ są przeprowadzane przez organ władzy publicznej i oznaczają ingerencję w funkcjonowanie rynku walutowego w celu dokonania zmian w zakresie rynkowych relacji podaży i popytu, a w konsekwencji także kursu walutowego.

W systemie funkcjonowania sztywnego kursu walutowego interwencje walutowe mają charakter ciągły, ponieważ bank centralny, broniąc określonego poziomu kursu, nieustannie kupuje i sprzedaje waluty obce. W systemie kursu płynnego interwencje walutowe mają z reguły charakter sporadyczny, choć w sytuacjach nadzwyczajnych mogą występować częściej. Celem takiej interwencji jest najczęściej utrzymanie inflacji pod kontrolą. Gdy waluta krajowa ulega znacznemu osłabieniu, deprecjacja kursu wywołuje wzrost cen towarów wymiennych i przyczynia się do przyspieszenia inflacji. Interwencja walutowa ma wówczas służyć podtrzymywaniu bieżących notowań kursu i łagodzeniu presji inflacyjnej. Innym jeszcze celem tych działań jest utrzymanie zewnętrznej równowagi płatniczej, a zwłaszcza konkurencyjności eksportu. Gdy waluta krajowa ulega silnej aprecjacji, może to ograniczać eksport i pobudzać import, a w efekcie prowadzić do znacznego pogorszenia salda obrotów bieżących. Z kolei wysoki deficyt na rachunku bieżącym zwiększa ryzyko wystąpienia kryzysu walutowego. Interwencje tego rodzaju są dokonywane najczęściej wówczas, gdy ma miejsce znaczne i dość trwałe odchylenie bieżącego kursu walutowego od poziomu równowagi średniookresowej. Wreszcie, celem interwencji walutowych może być także ograniczenie nadmiernych wahań kursu. Silne fluktuacje zwiększają bowiem niepewność w gospodarce i mogą destabilizować system

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

finansowy poprzez niekorzystny wpływ na bilanse banków i przedsiębiorstw zadłużonych w walutach obcych.

Interwencji walutowych można dokonywać w sposób systematyczny lub doraźny, w pewnych określonych sytuacjach. W pierwszym przypadku bank centralny określa na własny użytek poziom kursu, który uważa za pożądany z perspektywy realizacji swoich celów, takich jak na przykład utrzymanie stabilności cen. Bardzo często poziom kursu ustalany jest w formie przedziału dopuszczalnych wahań. Gdy kurs rynkowy zbliża się do dolnej lub górnej granicy przyjętego przedziału, bank centralny podejmuje interwencję. W systemie kursu płynnego ustalone pasmo wahań jest na ogół dość szerokie, tak aby uniknąć zbyt częstych interwencji. Bank centralny podaje niekiedy ustalone pasmo wahań do publicznej wiadomości, chcąc wpłynąć w ten sposób na oczekiwania rynku, lecz znacznie częściej nie ujawnia tej informacji, aby nie zachęcać do spekulacji. W tym samym celu, nawet jeśli występują oficjalnie ogłoszone granice pasma wahań, bank centralny stosuje wewnętrzne, węższe pasmo zmienności, aby nie dopuścić do sytuacji, w której uczestnicy rynku są w stanie przewidzieć interwencje walutową i podejmować spekulację. W przypadku, gdy interwencje są dokonywane w sposób systematyczny, uczestnicy rynku mogą jednak łatwo odczytać z danych rynku walutowego, jakich granic pasma broni bank centralny.

Dlatego też banki centralne wolą stosować interwencje doraźne, dając rynkom finansowym do zrozumienia, że nie bronią żadnego wcześniej ustalonego poziomu kursu, co zwiększa tym samym ryzyko związane ze spekulacją walutową. Banki centralne interweniują wówczas w sposób nieregularny, na różną skalę i z różną częstotliwością. Głównym celem takich interwencji jest hamowanie niekorzystnych tendencji rynkowych w kształtowaniu się kursu walutowego.

6.10.2. Rodzaje interwencji walutowych

Najprostszą formą interwencji na rynku walutowym są interwencje werbalne. Gdy kurs odchyla się nadmiernie od pożądanego poziomu (średniookresowej równowagi), bank centralny wysyła sygnały mające na celu zahamowanie i odwrócenie tej tendencji. Na przykład, kiedy kurs waluty krajowej ulega osłabieniu z powodu pojawienia się informacji o złych wynikach gospodarczych, prezes banku centralnego (lub inny członek władz decyzyjnych) zabiera głos, starając się przekonać uczestników rynku, że waluta krajowa jest fundamentalnie stabilna, a osłabienie kursu ma charakter przejściowy. Tego rodzaju komunikat przekazywany jest,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przykładowo, na specjalnie zwołanej konferencji prasowej, w formie wywiadu lub w postaci okazjonalnej wypowiedzi dla mediów. Czasami tego typu działania odnoszą zamierzony skutek, w szczególności gdy bank centralny jest instytucją wiarygodną. Wówczas można ustabilizować kurs waluty bez sięgania po bardziej kosztowne metody, takie jak sprzedaż walut obcych przez bank centralny.

Często interwencje werbalne nie odnoszą jednak oczekiwanego skutku. W takiej sytuacji obrona kursu wymaga sprzedaży walut obcych na rynku walutowym. Lecz bank centralny posiada w swoich aktywach jedynie niewielki zasób walut obcych w gotówce. Główną część rezerw międzynarodowych stanowią bowiem papiery wartościowe denominowane w walutach obcych. Interwencja walutowa banku centralnego polega więc, ściśle mówiąc, na sprzedaży obcych papierów wartościowych w zamian za pieniądz krajowy. Przebieg typowy takiej transakcji został opisany w ramce 6.14 oraz na rysunku 6.23.

Ramka 6.14. Schemat interwencji walutowej ze strony banku centralnego

Aby zrozumieć istotę interwencji walutowej, przedstawimy jej przebieg na prostym przykładzie. Zakładamy, że polski bank centralny ma do dyspozycji aktywa krajowe oraz aktywa zagraniczne, w tym głównie obligacje emitowane przez rządy najbardziej wiarygodnych krajów (takich jak na przykład USA, Niemcy, Wielka Brytania), z kolei pasywa banku centralnego składają się z gotówki pozostającej w obiegu i rezerw banków komercyjnych przechowywanych na rachunkach w banku centralnym. Przyjmijmy, że na skutek deficytu w bilansie obrotów bieżących występuje presja na deprecjację złotego w stosunku do euro i że polski bank centralny chce tej presji przeciwdziałać w trosce o utrzymanie stabilności cen. W tym celu bank postanawia dokonać interwencji walutowej w postaci sprzedaży aktywów zagranicznych ze swoich rezerw. Załóżmy, że w ramach tej operacji bank centralny sprzedaje za złote obligacje rządu Niemiec denominowane w euro na kwotę 100 mln PLN. Nabywcą jest niemiecki bank komercyjny, który potrzebuje obligacji swojego rządu jako zabezpieczenia w operacjach otwartego rynku i skupuje te obligacje za pośrednictwem polskiego banku komercyjnego. Polski bank komercyjny obciąża kwotą transakcji rachunek niemieckiego banku komercyjnego. W wyniku tych działań ilość pieniądza krajowego spada w relacji do pieniądza zagranicznego (euro), co wywołuje tendencję do umacniania kursu złotego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Operacja ta została przedstawiona schematycznie na rysunku 6.21 (zob. Weber 1986). W interwencji uczestniczą trzy podmioty – polski bank centralny (PBC), polski bank komercyjny (PBK) i niemiecki bank komercyjny (NBK). Najpierw bank centralny przekazuje niemieckie obligacje rządowe na kwotę 100 mln PLN do banku niemieckiego – transakcja (1). Następnie bank niemiecki kupuje za euro sumę 100 mln PLN od polskiego banku komercyjnego na rynku walutowym[47] i przelewa tę kwotę na rachunek tego banku – transakcja (2). Polski bank komercyjny uiszcza następnie wpłatę na rzecz banku centralnego ze swojego rachunku rezerwowego – transakcja (3).

Rysunek 6.23. Schemat interwencji walutowej (dane w PLN)

PBC		PBK		NBK	
Aktywa	Pasywa	Aktywa	Pasywa	Aktywa	Pasywa
(1) obligacje niem. – 100 mln				obligacje niem. + 100 mln	
(2)		przelew z NBK + 100 mln (w euro)		przelew do PBK – 100 mln	
(3)	spadek rezerw PBK – 100 mln	spadek rezerw w BC – 100 mln			

Źródło: opracowanie własne.

Skutkiem interwencji walutowej staje się spadek pieniądza rezerwowego o 100 mln PLN. Umocnienie kursu złotego jest spowodowane dwoma tendencjami. Po pierwsze, podaż pieniądza krajowego maleje w relacji do pieniądza zagranicznego. Po drugie zaś, w stopniu, w jakim niemiecki bank skupuje złote na rynku walutowym, kurs złotego ulegnie dalszej aprecjacji.

To właśnie sprzedaż (i zakup) aktywów zagranicznych stanowi najbardziej typowy rodzaj interwencji walutowych podejmowanych przez bank centralny. Sprzedaje on aktywa zagraniczne z rezerw oficjalnych za walutę krajową, broniąc kursu przed niepożądaną deprecjacją, lub skupuje aktywa zagraniczne za walutę krajową, chroniąc kurs waluty przed niepożądaną aprecjacją. Interwencje te mogą być i często są dokonywane bezpośrednio przez bank centralny, lecz czasami

[47] Bank niemiecki może uiścić wymaganą płatność na rzecz banku polskiego także z własnego rachunku prowadzonego w banku polskim.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

korzysta on także z pośrednictwa wyspecjalizowanego agenta finansowego, jakim staje się, przykładowo, jeden z państwowych banków komercyjnych (inwestycyjnych). Tego rodzaju rozwiązanie jest stosowane wówczas, gdy bank centralny nie chce ujawniać uczestnikom rynku, że dokonuje interwencji – po to, na przykład, aby nie pogłębiać niepokojów na rynku walutowym.

Czy interwencje walutowe dokonywane przez bank centralny powinny być jawne czy niejawne? To zależy od konkretnych okoliczności. Zaletą jawnej interwencji jest to, że stanowi ona jasny sygnał dla uczestników rynku, że bank centralny uważa aktualny poziom kursu walutowego za zbyt wysoki lub zbyt niski. Na przykład w obliczu nasilającej się spekulacji na obniżkę kursu waluty krajowej bank centralny może zainterweniować po to, aby podnieść kurs i w ten sposób spowodować straty u inwestorów, którzy spekulowali na zniżkę kursu. Podejmując interwencję, bank centralny w sposób świadomy zwiększa ryzyko związane ze spekulacją, co ma zniechęcać inwestorów do podejmowania takich działań. Jednocześnie jawne interwencje, jeśli dokonywane są w sposób spójny z oficjalnie ogłoszonym pożądanym poziomem kursu (lub pasmem kursowym), zwiększają przewidywalność polityki banku centralnego i ograniczają niepewność na rynku.

Na ogół banki centralne znajdują jednak istotne powody po temu, aby nie powiadamiać o interwencji przeprowadzanej na rynku walutowym. Przede wszystkim mogą obawiać się o jej skuteczność. Interwencje nie zawsze pozwalają bowiem zmienić niekorzystną rynkową tendencję zmian kursu (warunkami skuteczności zajmiemy się później). Gdyby okazało się, że jawna interwencja nie przyniosła oczekiwanej poprawy sytuacji, bank centralny poniósłby straty, nie osiągając zakładanego celu. Naraziłby w ten sposób na szwank swoją reputację i wiarygodność. Innym powodem podejmowania niejawnych interwencji jest chęć zasugerowania, że to sam rynek wpływa na aktualną ocenę kondycji waluty krajowej. Często też banki centralne ukrywają fakt przeprowadzenia interwencji, ponieważ nie chcą ujawniać, że sytuacja na rynku walutowym jest na tyle poważna, iż wymaga podjęcia nadzwyczajnych działań.

Ogólnie rzecz biorąc, wybór strategii komunikowania – a w jej ramach decyzja o tym, czy interwencje mają mieć charakter jawny, czy też ukryty – zależy od skali nierównowagi na rynku walutowym, wiarygodności banku, wielkości posiadanych rezerw walutowych i stopnia odchylenia kursu od poziomu równowagi. Decydujące okazują się więc konkretne warunki – teoria ekonomii nie przesądza w tym przypadku, czy bardziej skuteczna będzie jawna, czy też niejawna interwencja.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Interwencje walutowe może także podejmować rząd danego kraju (lub w jego imieniu i na jego rachunek wyspecjalizowana instytucja finansowa – na przykład bank państwowy). Cele interwencji rządowej definiowane są jednak z reguły inaczej niż założenia banku centralnego. Gdy rząd kupuje lub sprzedaje waluty obce, chcąc doprowadzić do określonej zmiany kursu walutowego, dąży najczęściej do poprawy sytuacji w przypadku sektora eksportowego, redukcji obciążenia budżetu kosztem obsługi zadłużenia zagranicznego lub zmniejszenia wielkości długu publicznego denominowanego w walutach obcych. Interwencje banku centralnego są bardziej symetryczne, zasadniczo nakierowane na ograniczenie odchyleń kursu walutowego od średniookresowego kursu równowagi.

Interwencje walutowe mogą być podejmowane jednostronnie przez zainteresowany kraj lub przyjmować postać skoordynowanej akcji realizowanej jednocześnie przez kilka państw. W pierwszym przypadku ich skuteczność jest z reguły mniejsza, a ryzyko niepowodzenia większe. Wynika to z głównie ograniczonych zasobów walutowych znajdujących się w posiadaniu banków centralnych, zwłaszcza w relacji do środków pozostających do dyspozycji inwestorów międzynarodowych. Skoordynowane, wielostronne interwencje, zwłaszcza z udziałem krajów, których waluty są pod presją aprecjacji, okazują się z reguły bardzo skuteczne, przynajmniej w krótkim okresie. Do udanych interwencji skoordynowanych należą wspólne akcje krajów grupy G5 i G7 (skupiających największe demokratyczne kraje świata), takie jak np. porozumienie z Plaza (*Plaza Agreement*, z września 1985 roku) w sprawie obniżenia wartości dolara w stosunku do innych najważniejszych walut[48] lub porozumienie z Luwru (*Louvre Accord*, z lutego 1987 roku) w sprawie stabilizowania kursu głównych walut w czasie kryzysu giełdowego.

6.10.3. Skuteczność interwencji walutowych

Interwencja walutowa oddziałuje na kurs waluty za pośrednictwem kilku kanałów (Dominguez i Frankel 1993). Najważniejszym z nich jest kanał monetarny. Interwencje wpływają na zmiany bazy monetarnej, a za ich pośrednictwem na podaż pieniądza i rynkowe stopy procentowe. Z kolei zmiany stóp procentowych – w systemie kursu płynnego – oddziałują na podaż i popyt na waluty obce,

[48] 22 września 1985 roku w Hotelu Plaza w Nowym Jorku przedstawiciele rządów państw grupy G5 (Stany Zjednoczone, Japonia, Republika Federalna Niemiec, Wielka Brytania i Francja) podpisali porozumienie w sprawie wspólnej interwencji mającej na celu obniżenie wartości dolara względem innych walut. W okresie następnych sześciu miesięcy efektywny kurs dolara spadł o 20%.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

a w konsekwencji także na ich cenę, czyli kurs walutowy. Na przykład sprzedaż walut obcych z rezerw banku centralnego w zamian za pieniądz krajowy zmniejsza ilość pieniądza w obiegu, podnosi krajowe stopy procentowe i wywołuje presję na osłabienie kursu walut obcych. Skuteczność funkcjonowania tego kanału zależy przede wszystkim od skali interwencji i elastyczności popytu na walutę względem stopy procentowej.

Drugim kanałem oddziaływania jest kanał struktury portfela aktywów. W warunkach rynkowych inwestorzy optymalizują strukturę posiadanych aktywów z perspektywy oczekiwanych stóp przychodu i ryzyka. Interwencja walutowa, polegająca na zwiększeniu (lub zmniejszeniu) podaży aktywów denominowanych w walutach obcych, zaburza optymalną strukturę portfela. Jej przywrócenie wymaga wzrostu (spadku) dochodowości innych aktywów. Na przykład sprzedaż walut obcych przez bank centralny oznacza wzrost podaży tych środków w relacji do podaży pieniądza krajowego. Gdy aktywa zagraniczne i aktywa krajowe nie są doskonale substytucyjne, następuje spadek kursu waluty obcej. Sprzedaż walut zagranicznych za walutę krajową przekłada się na wzrost stóp procentowych, co zwiększa atrakcyjność lokat w walucie krajowej. W tych warunkach inwestorzy są skłonni zwiększać udział aktywów krajowych kosztem aktywów zagranicznych w swoim portfelu.

Trzecim kanałem oddziaływania jest kanał komunikacji. Sama zapowiedź interwencji ze strony banku centralnego wpływa na zmianę oczekiwań i może wywołać pożądane dostosowanie na rynku (np. spadek popytu na waluty obce). Podobne w skutkach są również komunikaty podkreślające stabilność fundamentów ekonomicznych. Skuteczność tego kanału pozostaje wysoka w warunkach asymetrii informacji na temat rzeczywistej kondycji gospodarki, występującej między uczestnikami rynku a bankiem centralnym.

Wreszcie, wyróżnia się także czwarty kanał – kanał analizy technicznej (*noise trading*). Ponieważ interwencje dokonywane są najczęściej na krańcach typowego przedziału wahań kursu, powoduje to przyłączanie się innych uczestników rynku wykorzystujących analizę techniczną do trendu wyznaczanego przez transakcje dokonywane przez bank centralny. Funkcjonowanie tego kanału nasila skutki interwencji, lecz tylko w krótkim okresie.

Skuteczność interwencji, czyli siła i trwałość jej oddziaływania na kurs walutowy, zależy jednak od wielu czynników i nie daje gwarancji skuteczności. Co więcej, nigdy do końca nie wiadomo, czy zaobserwowana zmiana kursu powstaje na skutek interwencji, czy też dokonuje się pod wpływem innych czynników. Badania empiryczne sugerują, że interwencje walutowe mają w zasadzie ograniczoną

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

skuteczność w zakresie kształtowania pożądanego poziomu kursu, zwłaszcza gdy występuje wyraźna rozbieżność między oczekiwaniami rynku a celami banku centralnego. W szczególności zaś, jeśli fundamenty ekonomiczne gospodarki wskazują wyraźnie, że kurs jest zawyżony (lub zaniżony) w stosunku do poziomu równowagi, interwencja na ogół nie przyniesie oczekiwanych rezultatów. W takich warunkach lepsza jest strategia „opierania się wiatrowi" (*leaning against the wind*), polegająca na łagodzeniu tendencji rynkowej, lecz bez podejmowania wysiłków zmierzających do jej odwrócenia. Lepsze rezultaty osiąga się wówczas, gdy bank centralny działa zgodnie z „fundamentalnymi" czynnikami, a odchylenia kursu, którym chce zapobiegać, mają charakter przejściowy lub czysto spekulacyjny.

Skuteczność interwencji jest wyraźnie asymetryczna. O wiele łatwiej przychodzi interweniować, broniąc kursu przed aprecjacją niż przed deprecjacją. W pierwszym przypadku zasoby banku centralnego są teoretycznie nieograniczone, ponieważ bank interweniuje, sprzedając własną walutę, której jest jedynym dostawcą. Kosztem przeprowadzenia takiej operacji jest jednak utrata kontroli nad podażą pieniądza rezerwowego, co może prowadzić do powstawania baniek spekulacyjnych oraz inflacji. Dlatego też w praktyce żaden bank centralny nie będzie skłonny akumulować aktywów zagranicznych w nieskończoność. Przykładem jest działanie banku centralnego Szwajcarii, który starał się w latach 2011–2014 ograniczyć skalę aprecjacji franka – jak się okazało, bezskutecznie (zob. ramka 6.15). W drugim przypadku interwencja polega na sprzedaży walut obcych z rezerw. Wówczas granice interwencji są wyznaczone przez zasób oficjalnych aktywów rezerwowych znajdujących się w posiadaniu banku centralnego. Ponieważ uczestnicy rynku wiedzą, jakimi rezerwami dysponuje bank centralny oraz że rezerwy te są zawsze ograniczone, to jeśli zabraknie przekonania co do tego, że fundamenty gospodarki pozostają solidne albo ulegną poprawie, presja na deprecjację waluty krajowej stanie się tak silna, że bank centralny nie będzie w stanie skutecznie jej się przeciwstawić. Ogólnie rzecz biorąc, gdy fundamenty okażą się nietrwałe, żaden bank centralny nie wygra z rynkiem.

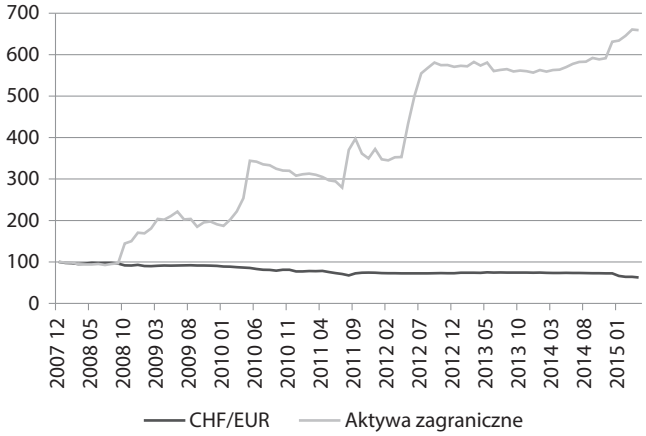
Ramka 6.15. Nieskuteczność interwencji walutowej: polityka Szwajcarskiego Banku Narodowego w latach 2011–2014

Po wybuchu kryzysu finansowego w 2008 roku Szwajcaria odgrywała rolę finansowej „bezpiecznej przystani" dla inwestorów wycofujących się z zagrożonych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kryzysem zadłużenia państw strefy euro. Napływ kapitału powodował systematyczne umacnianie się kursu franka, który wzrósł ze średniego poziomu 1,57 CHF/EUR w latach 2005–2009 do 1,12–1,15 CHF/EUR w lecie 2011 roku, czyli prawie o 30%. W obliczu tak silnej aprecjacji waluty, obniżającej międzynarodową konkurencyjność gospodarki szwajcarskiej i grożącej załamaniem sektora eksportowego, bank centralny Szwajcarii (SNB) zdecydował się we wrześniu 2011 roku na przeprowadzenie interwencji walutowych na dużą skalę, a celem tych działań było utrzymanie kursu franka wobec euro na poziomie nie niższym niż 1,2 CHF/EUR. Interwencje te w latach 2011–2014 doprowadziły do wzrostu sumy zagranicznych aktywów rezerwowych znajdujących się w posiadaniu SNB z 85 mld do 550 mld CHF i niemal dziesięciokrotnego wzrostu podaży pieniądza rezerwowego (z 46 mld do 450 mld CHF). Mimo podjęcia tych działań napływ kapitału nie ustawał i nadal utrzymywała się presja na aprecjację.

Wykres 6.10. Zmiany poziomu zagranicznych aktywów SNB (w CHF) i kurs CHF/EUR w latach 2008–2015 (grudzień 2007 = 100)



Źródło: Swiss National Bank, obliczenia własne.

Rosnące obawy związane z niekontrolowanym wzrostem podaży pieniądza krajowego oraz perspektywa dalszego napływu kapitału w związku z zapowiedzią rozpoczęcia „luzowania ilościowego" przez Europejski Bank Centralny na początku 2015 roku skłoniły ostatecznie SNB do wycofania się z interwencji. W styczniu 2015 roku bank centralny Szwajcarii ogłosił, że zaprzestaje bronić kursu na poziomie 1,20 CHF/EUR. W ciągu kilku dni kurs franka umocnił się o ponad 10%.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 6.10 pokazuje, że interwencje walutowe SNB okazały się mało skuteczne: nie zahamowały napływu kapitału zagranicznego (suma tych aktywów w posiadaniu SNB uległa podwojeniu w okresie od początku 2012 roku do początku 2015 roku) i nie zmniejszyły presji na aprecjację franka.

6.10.4. Interwencja sterylizowana

Interwencja walutowa polega na tym, że bank centralny skupuje i sprzedaje walutę obcą, broniąc określonego poziomu kursu lub starając się przeciwdziałać tendencji do niepożądanej aprecjacji bądź deprecjacji waluty krajowej. Ponieważ zakup walut obcych oznacza wzrost krajowej bazy monetarnej, a sprzedaż walut zagranicznych przekłada się na spadek krajowej bazy monetarnej, interwencje mają bezpośredni wpływ na podaż pieniądza krajowego. Jeśli bank centralny chce jednak utrzymać podaż pieniądza krajowego na stałym poziomie (np. w obawie przed wzrostem inflacji), musi sterylizować pieniężne efekty interwencji walutowej. W związku z tym, gdy bank centralny skupuje waluty obce w zamian za pieniądz krajowy, musi jednocześnie ściągać z rynku nadwyżkę pieniądza krajowego poprzez absorbujące operacje otwartego rynku (operacje *reverse repo*). Podobnie, gdy bank centralny sprzedaje waluty obce w zamian za pieniądz krajowy, musi jednocześnie dostarczyć na rynek dodatkową podaż pieniądza krajowego poprzez zasilające operacje otwartego rynku (operacje *repo*). Operacje niwelujące pieniężne skutki interwencji walutowej nazywamy operacjami sterylizującymi. Sprawiają one, że w warunkach dokonywania interwencji walutowych podaż pieniądza rezerwowego nie ulega zmianie.

Czy to oznacza, że bank centralny jest w stanie całkowicie i w każdych warunkach neutralizować wpływ interwencji walutowych na podaż pieniądza krajowego? Niestety, tak się nie dzieje, i to z dwóch powodów. Po pierwsze, sterylizacja jest kosztowna. Rozważmy przypadek sterylizacji związanej z aprecjacją waluty krajowej. W takiej sytuacji dokonywana przez bank centralny sprzedaż obligacji bankom komercyjnym w ramach operacji *reverse repo* wywołuje presję na wzrost stóp procentowych. Aby skłonić banki do zwiększenia w ich portfelach zasobu obligacji, bank centralny musi zaoferować bardziej atrakcyjne oprocentowanie. Im większa staje się skala sterylizacji, tym wyższe musi być oprocentowanie obligacji. Oznacza to, że koszty sterylizacji rosną bardziej niż proporcjonalnie w stosunku do skali sterylizacji. Co dzieje się natomiast w przypadku sterylizacji związanej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

z deprecjacją waluty krajowej? Bank centralny, sprzedając waluty obce, prowadzi równolegle operacje *repo*, czyli skupuje obligacje i zapewnia płynność w sektorze bankowym. Wywołuje to presję na spadek stóp procentowych, co może dodatkowo nasilać tendencje inflacyjne spowodowane deprecjacją.

Po drugie, sterylizacja bywa na dłuższą metę nieskuteczna. W przypadku sterylizacji związanej z aprecjacją waluty krajowej wzrost stóp procentowych, konieczny w celu skłonienia banków komercyjnych do dokonywania dalszych zakupów obligacji, pociąga za sobą zarazem zwiększony napływ kapitału zagranicznego. Im większy zasób obligacji bank centralny stara się ulokować w portfelach banków, tym wyższe musi być ich oprocentowanie i tym silniejszy dodatkowy napływ kapitału z zagranicy (Claessens i in. 1993). Ten wzmożony udział kapitału, wywołany wzrostem stóp procentowych na skutek operacji sterylizującej, przekłada się na wzrost podaży pieniądza krajowego. Relację spadku podaży pieniądza na skutek sterylizacji do wzrostu podaży pieniądza pod wpływem wzrostu stóp procentowych wywołanego sterylizacją mierzy tzw. współczynnik kompensacji (*offset coefficient*). Im wyższa staje się wartość tego współczynnika, tym mniej skuteczna jest sterylizacja (Kouri i Porter 1974, Bond 1998). Przy poziomie współczynnika bliskim jedności napływ kapitału całkowicie eliminuje wpływ sterylizacji, w związku z czym nie ma sposobu, aby ograniczyć pieniężne skutki interwencji walutowej. Z kolei gdy wartość współczynnika kompensacji jest większa od jedności, sterylizacja dodatkowo potęguje pieniężne efekty interwencji walutowej. Wynika z tego, że sterylizacja ma sens tylko wtedy, gdy współczynnik kompensacji jest niewielki, najlepiej bliski zeru[49].

Ograniczenia interwencji sterylizowanej sprawiają, że pod względem wpływu na kurs walutowy operacja ta jest mniej skuteczna niż interwencja niesterylizowana (zob. Frenkel i in. 2001, Sarno i Taylor 2001). Dzieje się tak dlatego, że pieniężne skutki interwencji wpływają na kurs walutowy, nadając mu pożądany kierunek. Na przykład skup walut obcych podnosi ich cenę i w tym samym kierunku oddziałuje również związany z tym wzrost podaży pieniądza krajowego. W rezultacie próba zniwelowania pieniężnych skutków interwencji poprzez sterylizację osłabia wpływ interwencji na kurs walutowy.

[49] Rosati (1999) oszacował wielkość współczynnika kompensacji w przypadku Polski w okresie od stycznia 1995 roku do marca 1998 roku na poziomie 0,561.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

6.11. Polityka kursowa w Polsce

Polityka kursu walutowego w Polsce po 1989 roku dzieli się wyraźnie na dwa okresy. W pierwszym z nich, obejmującym początkowe lata transformacji, obowiązywał najpierw system kursu stałego, a następnie system kursu pełzającego o stopniowo malejącym tempie dewaluacji i systematycznie poszerzanym paśmie dopuszczalnych wahań. Z kolei od kwietnia 2000 roku w Polsce funkcjonuje w pełni płynny kurs walutowy, a jednocześnie brak jest jakichkolwiek ograniczeń w zakresie wymienialności i transgranicznych przepływach kapitału. Narodowy Bank Polski w zasadzie nie interweniuje na rynku walutowym (z wyjątkiem kilku epizodów w latach 2010–2012). Oznacza to, że w praktyce od 2000 roku nie prowadzi się w Polsce polityki kursowej w sensie ustalania przez władzę publiczną pożądanego poziomu kursu złotego, co pozostaje spójne z realizowaną od 1999 roku strategią bezpośredniego celu inflacyjnego[50].

6.11.1. Polityka kursowa w Polsce w latach 1990–1999

W styczniu 1990 roku, w ramach realizowanego wówczas programu stabilizacyjnego, władze polskie zdewaluowały kurs złotego wobec dolara amerykańskiego o 42% i wprowadziły wymienialność polskiej waluty w przypadku większości transakcji bieżących bilansu płatniczego (tzw. „wymienialność wewnętrzna", wprowadzona na podstawie art. XIV MFW – zob. ramka 6.4). Głęboka dewaluacja miała umożliwić odzyskanie równowagi w bilansie płatniczym, a usztywnienie kursu powinno było służyć jako kotwica nominalna w walce z inflacją, która w tamtym okresie sięgała kilkudziesięciu procent miesięcznie. Skala dewaluacji okazała się jednak zbyt duża, ponieważ nie wzięto pod uwagę reakcji uczestników rynku wywołanej efektem wiarygodnej zmiany systemowej (tzw. *credible regime change* – piszemy o tym w rozdziale 1, zob. ramka 1.7), co pogłębiło spadek produkcji, a jednocześnie utrudniło hamowanie inflacji. W warunkach stałego nominalnego kursu szybki wzrost inflacji powodował aprecjację kursu realnego i erozję konkurencyjności sektora eksportowego. Zmusiło to władze – już po kilkunastu

[50] Przypomnijmy, że niespójność polityki kontrolowanego kursu z polityką celu inflacyjnego w warunkach swobody przepływów kapitału wynika z tzw. trójkąta niemożności polityki makroekonomicznej, opisanego w poprzednim rozdziale (punkt 5.10.2).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

miesiącach – do kolejnej dewaluacji kursu (o 17%). Jednocześnie, dążąc do stabilizacji kursu efektywnego, powiązano kurs złotego z koszykiem pięciu walut wymienialnych o największym udziale w polskim handlu zagranicznym. Ponieważ utrzymująca się nadal wysoka inflacja powodowała niepożądaną aprecjację kursu realnego, w październiku 1991 roku system stałego kursu zastąpiono systemem pełzającej dewaluacji (*crawling peg*), aby lepiej dostosować tempo zmian kursu do szybko rosnących cen. Na początku następnego roku dokonano kolejnej dewaluacji (o 12%), a po kilku miesiącach system pełzającego kursu zastąpiono systemem pełzającego pasma kursowego (*crawling band*) o szerokości +/–1% wokół centralnego parytetu, co miało na celu zwiększenie elastyczności i odporności systemu kursowego (Rosati 1996).

W miarę stabilizowania się gospodarki po pierwszym okresie transformacji skala nierównowagi zewnętrznej malała. W sierpniu 1993 roku dokonano jeszcze jednej – już ostatniej – dewaluacji (o 8%), lecz systematyczny spadek inflacji i poprawa sytuacji w ramach bilansu płatniczego pozwoliły na stopniowe zmniejszanie tempa pełzającej dewaluacji z 1,8% miesięcznie w 1991 roku do 1,4% miesięcznie w 1994 roku (Rosati 1998, rozdział 6).

Zasadnicza zmiana nastąpiła w końcu 1994 roku po podpisaniu przez Polskę porozumień o restrukturyzacji zadłużenia zagranicznego z wierzycielami rządowymi (Klub Paryski) i wierzycielami prywatnymi (Klub Londyński). Umożliwiły one Polsce wejście na międzynarodowe rynki finansowe, co wywołało masowy napływ kapitału zagranicznego i rosnącą presję na aprecjację złotego. System kontrolowanego kursu zmuszał Narodowy Bank Polski do skupu walut obcych w celu obrony dopuszczalnego pasma wahań. W rezultacie, o ile w 1994 roku stan oficjalnych rezerw walutowych powiększył się o niecałe 2 mld USD, o tyle w następnym roku przyrost ten wyniósł prawie 9 mld USD. Presja aprecjacyjna zmusiła NBP do dokonania w 1995 roku dwukrotnej rewaluacji parytetu centralnego (łącznie o ponad 11%), obniżenia tempa pełzającej dewaluacji do 1,0% miesięcznie i do poszerzenia pasma dopuszczalnych wahań kursu do +/–7%. W związku z poprawą sytuacji w zakresie bilansu płatniczego Polska rozszerzyła zakres wymienialności waluty krajowej, przyjmując w 1995 roku standard opisany w art. VIII MFW (zob. ramka 6.4).

Stały napływ kapitału zagranicznego podtrzymywał presję aprecjacyjną w kolejnych latach. Broniąc pełzającej dewaluacji kursu w dążeniu do utrzymania konkurencyjności międzynarodowej polskich eksporterów, NBP nie był jednak w stanie w pełni kontrolować inflacji, która systematycznie przekraczała wyznaczone cele.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ta wewnętrznie sprzeczna „eklektyczna" strategia została ostatecznie zarzucona w 1999 roku, kiedy została zastąpiona polityką bezpośredniego celu inflacyjnego (zob. rozdział 5.7). W kwietniu 2000 roku zniesiono system pełzającego pasma i nastąpiło całkowite uwolnienie kursu. Syntetyczne zestawienie głównych decyzji z zakresu polityki kursowej w tym czasie przedstawia tabela 6.7.

Tabela 6.8. Polityka kursowa w Polsce w latach 1990–2000

Data	Zmiana kursu / zmiana systemowa	Waluta odniesienia	Szerokość pasma wahań kursu
1.01.1990	dewaluacja o 42%	dolar amerykański	
1.01.1990	ustalenie stałego kursu (9500 PLN/USD)	dolar amerykański	brak
20.05.1991	dewaluacja o 17% wprowadzenie koszyka walut	koszyk 5 walut[a]	
14.10.1991	kurs pełzający, tempo dewaluacji 1,8% na miesiąc		
26.02.1992	dewaluacja o 12%		
1.10.1992	wprowadzenie pasma kursowego		+/–1%
27.08.1993	dewaluacja parytetu centralnego o 8%		
13.09.1994	tempo dewaluacji 1,5% na miesiąc		
30.11.1994	tempo dewaluacji 1,4% na miesiąc		
16.02.1995	tempo dewaluacji 1,2% na miesiąc		
6.03.1995	poszerzenie pasma wahań		+/–2%
16.05.1995	rewaluacja parytetu centralnego o 5% poszerzenie pasma wahań		+/–7%
22.12.1995	rewaluacja parytetu centralnego o 6%		
8.01.1996	tempo dewaluacji 1,0% na miesiąc		
26.02.1998	tempo dewaluacji 0,8% na miesiąc poszerzenie pasma wahań		+/–10%
17.07.1998	tempo dewaluacji 0,65% na miesiąc		
10.09.1998	tempo dewaluacji 0,5% na miesiąc		
29.10.1998	poszerzenie pasma wahań		+/–12,5%
1.01.1999	poszerzenie pasma wahań zmiana koszyka tempo dewaluacji 0,3% na miesiąc	koszyk 2 walut[b]	+/–15%
12.04.2000	kurs płynny		brak

a) Koszyk obejmował pięć walut wymienialnych w następujących proporcjach: dolar amerykański (45%), marka niemiecka (35%), funt brytyjski (10%), frank francuski (5%) i frank szwajcarski (5%).
b) Koszyk obejmował dwie waluty wymienialne w następujących proporcjach: euro (55%) i dolar amerykański (45%).

Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Polityka kursowa prowadzona w Polsce w pierwszym okresie transformacji okazała się zasadniczo dość skuteczna w realizacji celów stabilizacyjnych. Stały kurs w początkowym okresie przyczynił się do obniżenia inflacji, choć nadmierna początkowa dewaluacja zapewne wydłużyła ten proces. Jednocześnie system pełzającej dewaluacji okazał się skuteczny w utrzymywaniu międzynarodowej konkurencyjności polskiej gospodarki i przyczynił się do szybkiego wzrostu eksportu i przywrócenia równowagi zewnętrznej. W kolejnych latach, w miarę otwierania się polskiej gospodarki na przepływy kapitału, coraz trudniej było jednak osiągać jednocześnie cele równowagi wewnętrznej i zewnętrznej w warunkach kontrolowanego kursu.

6.11.2. Polityka kursowa w Polsce w latach 2000–2015

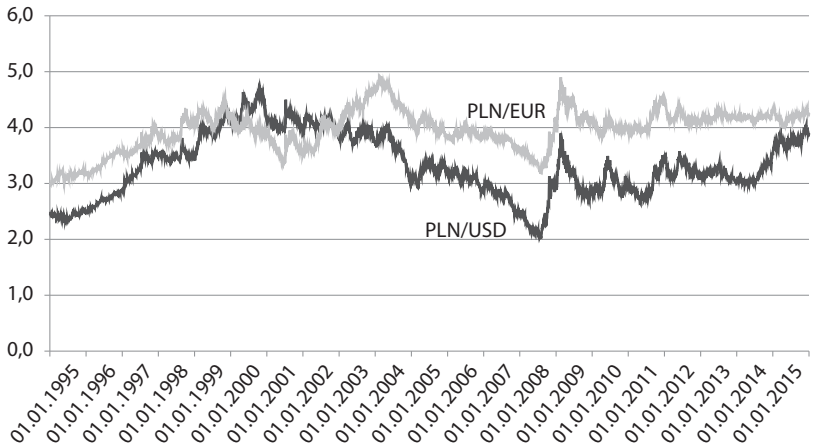
Jak już wspomnieliśmy, system płynnego kursu i generalny brak interwencji walutowych oznaczał, że po 2000 roku w Polsce nie prowadzono polityki walutowej w sensie świadomego kształtowania poziomu kursu. Argumentem przemawiającym na rzecz takiego rozwiązania był nie tylko potencjalny konflikt celów wynikający z „trójkąta niemożności" (w warunkach swobody przepływu kapitałów nie można jednocześnie i niezależnie wyznaczać celu kursowego oraz celu inflacyjnego), ale także chęć uniknięcia niebezpieczeństw związanych z reżimem stałego kursu i ryzyka ewentualnego przewartościowania waluty krajowej w warunkach inflacji znacznie wyższej niż u partnerów handlowych.

Uwolnieniu kursu złotego towarzyszyły jednak też obawy. Z jednej strony istniało spore prawdopodobieństwo nadmiernego umocnienia kursu waluty w związku z napływem kapitału, co mogło nie tylko osłabić sektor eksportowy, lecz także grozić załamaniem kursu w sytuacji nagłego odpływu kapitału. Czy obawy te okazały się uzasadnione? Wykres 6.11 przedstawia kształtowanie się kursu złotego wobec euro i dolara amerykańskiego w latach 1995–2015. Jak widać, między 1995 a 1999 rokiem system pełzającej dewaluacji (choć w malejącym tempie) powodował systematyczne osłabienie kursu złotego wobec obu walut. Ta tendencja wyhamowała po uwolnieniu kursu i w kolejnych latach kurs złotego nie wykazywał żadnej trwałej tendencji do aprecjacji lub deprecjacji. Wyjątkiem jest okres trwający od początku 2007 do połowy 2009 roku, kiedy kurs złotego najpierw został poddany spekulacyjnej presji na aprecjację, a następnie przeszedł silną korektę w dół w związku z wybuchem kryzysu finansowego. W następnych latach kurs powrócił do długookresowego poziomu równowagi, który kształtował się w okolicach 4 PLN/EUR

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i 3,35 PLN/USD[51]. W istocie można więc stwierdzić, że w długim okresie nominalny kurs złotego wykazywał dość dużą stabilność – w końcu 2015 roku kurs złotego względem dolara był silniejszy o 4,8% niż w dniu upłynnienia kursu piętnaście lat wcześniej (12 kwietnia 2000 roku), a kurs wobec euro stał się słabszy o 8,5%.

Wykres 6.11. Kurs PLN/EUR i PLN/USD w latach 1995–2015 (notowania dzienne)



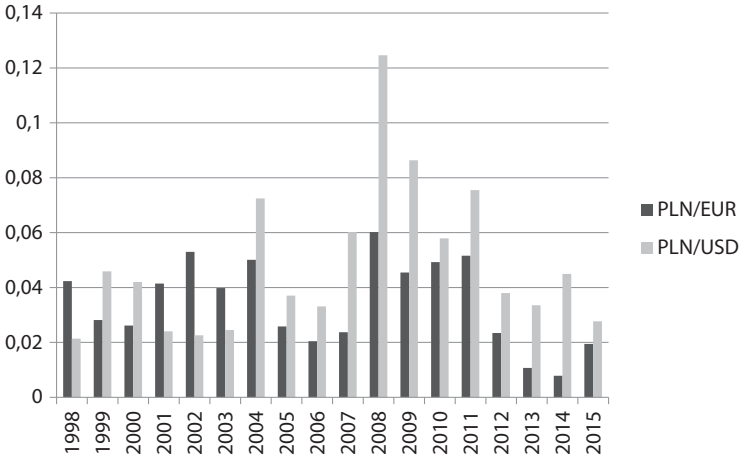
Źródło: dane NBP.

Nie sprawdziły się również obawy dotyczące tego, że uwolnienie kursu może wywołać znaczny wzrost krótkookresowej zmienności kursu waluty i wiązać się z negatywnymi skutkami w przypadku wymiany międzynarodowej. Jak pokazuje wykres 6.12, po wprowadzeniu kursu płynnego zmienność kursu waluty mierzona relacją odchylenia standardowego do średniej wzrosła nieznacznie. Wyraźny przyrost zmienności kursu miał miejsce dopiero w latach 2008–2011 w związku z kryzysem finansowym i jego konsekwencjami. W późniejszym czasie zmienność kursu powróciła do bardzo niskiego poziomu. Wykres 6.12 ukazuje także rosnące powiązanie polskiej waluty z euro – o ile jeszcze w latach 1998–2003 zmienność kursu złotego w stosunku do euro była wyższa niż wobec dolara, co wskazywało na to, że złoty w większym stopniu „podążał" za dolarem amerykańskim, o tyle po wejściu do UE w 2004 roku zmienność kursu złotego względem euro stała się już znacznie niższa niż wobec dolara.

[51] Średnie arytmetyczne kursów złotego wobec euro i dolara w okresie od 12.04.2000 roku (uwolnienie złotego) do 31.12.2015 roku wyniosły odpowiednio 4,0552 PLN/EUR i 3,3625 PLN/USD.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 6.12. Zmienność kursu PLN/EUR i PLN/USD w latach 1998–2015 (stosunek odchylenia standardowego do średniej rocznej)



Źródło: opracowanie własne na podstawie danych NBP.

W latach 2002–2004 dokonano też dalszej liberalizacji w zakresie wymienialności waluty krajowej na waluty obce. Na podstawie ustawy z 27 lipca 2002 roku Prawo dewizowe, rozszerzono zakres swobody wymienialności w Polsce, zezwalając rezydentom na posiadanie wartości dewizowych w kraju i za granicą, a nierezydentom na przechowywanie wartości dewizowych w kraju. Nadal utrzymywano jednak generalny obowiązek natychmiastowego transferu tych wartości z zagranicy do kraju. Ograniczenie to, a także większość pozostałych restrykcji w zakresie wymienialności, w tym w transakcjach kapitałowych i finansowych bilansu płatniczego, zniesiono ostatecznie w związku z wejściem Polski do Unii Europejskiej w 2004 roku.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Literatura cytowana

Acocella N., *Economic Policy in the Era of Globalisation*, Cambridge University Press, Cambridge 2005.

Alexander S.S., *The Effects of Devaluation on a Trade Balance*, "IMF Staff Papers" 1952, no. 2.

Balassa B., *The Purchasing Power Parity Doctrine: A Reappraisal*, "Journal of Political Economy" 1964, vol. 72.

Barro R.J., *Economic Growth in a Cross-Section of Countries*, "Quarterly Journal of Economics" 1991, vol. 106, no. 2.

Barro R.J., Sala-i-Martin X., *Convergence*, "Journal of Political Economy" 1992, vol. 100, no. 2.

Benassy-Quéré A., Qoeuré P., Jacquet P., Pisani-Ferry J., *Economic Policy. Theory and Practice*, Oxford University Press, Oxford–New York 2010.

Blanchard O., Giavazzi F., *Current Account Deficits in the Euro Area: The End of the Feldstein-Horioka Puzzle?*, "Brookings Papers on Economic Activity" 2002, no. 2.

Bond J.T., *Capital Flows to Asia: The Role of Monetary Policy*, "Empirica" 1998, vol. 25, no. 2.

Burda M., Wyplosz C., *Makroekonomia. Podręcznik europejski*, wydanie 2 zmienione, PWE, Warszawa 2000.

Calvo G.A., Reinhart C.M., *Fear of Floating*, "NBER Working Papers" 2000, no. 7993.

Cipriani M., *The Balassa-Samuelson Effect in Transition Economies*, IMF, Washington 2000, mimeo.

Claessens S., Dooley M., Warner A., *Portfolio Capital Flows: Hot or Cool?*, "World Bank Discussion Paper" 1993, no. 228.

Cline W.B., Williamson J., *New Estimates of Fundamental Equilibrium Exchange Rates*, "Policy Brief" 2008, no. PB08–7, Paterson Institute for International Economics, Washington.

Cooper R., *Currency devaluations in developing countries*, "Essays in International Finance" 1971, no. 86.

Darvas Z. i in., *Analysis of developments in EU capital flows in the global context*, Final Report Bruegel, no. MARKT/2013/50/F, Brussels, November 2014.

Dominguez K., Frankel J., *Does Foreign Exchange Intervention Work?*, Institute for International Economics, Washington 1993.

Eichengreen B., Razo-Garcia R., *The International Monetary System in the Last and Next 20 Years*, "Economic Policy" 2006, no. 21(47).

Fischer S., *Exchange Rate Regimes: Is the Bipolar View Correct?*, Distinguished Lecture on Economics in Government, "Journal of Economic Perspectives" 2001, vol. 15, no. 2.

Fisher I., *The Theory of Interest*, MacMillan, New York 1930.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Frankel J., *Monetary and Portfolio-Balance Models of Exchange Rate Determination*, w: *Economic Interdependence and Flexible Exchange Rates*, red. J. Bhandari, B. Putnam, MIT Press, Cambridge 1983.

Frankel J., *No Single Currency Regime is Right for All Countries or at All Times*, "NBER Working Paper" 1999, no. 7338.

Frenkel M., Pierdzioch Ch., Stadtmann G., *The Interventions of the European Central Bank: Effects, Effectiveness and Policy Implications*, Koblenz University, Mainz 2001.

Froot K.A., Rogoff K., *Perspectives on PPP and long-run real exchange rates*, "NBER Working Paper" 1994, no. 4952.

Ghosh A.R., Ostry J.D., Qureshi M.S., *Exchange Rate Management and Crisis Susceptibility. A Reassessment,* "IMF Working Paper" 2014, no. WP/14/11.

Hanke S.H., *Currency Boards*, "The Annals of the American Academy of Political and Social Science" 2002, no. 579.

Hanke S.H., *Currency Wars, the Devaluation Delusion*, "Globe Asia", April 2016.

Hanke S.H., Schuler K., *Currency boards for developing countries*, Institute for Contemporary Studies, San Francisco 1994.

Hirschman A.O., *Devaluation and the Trade Balance: A Note*, "Review of Economics and Statistics" 1949, vol. 31.

IMF, *Review of Exchange Arrangements, Restrictions, and Controls*, 27 November, 2007, Washington, mimeo.

IMF, *Annual Report on Exchange Arrangements and Exchange Restrictions 2014*, October 2014, Washington.

Kasa K., *Contractionary effects of devaluation*, "FRBSF Economic Letter" 1998, no. 34.

Kouri P., Porter M., *International Capital Flows and Portfolio Equilibrium*, "Journal of Political Economy" 1974, no. 4.

Krugman P.R., Taylor L., *Contractionary Effects of Devaluation*, "Journal of International Economics" 1978, no. 8.

Krugman P.R., Obstfeld M., *International Economics*, wydanie II, Harper Collins Publishers, New York 1991.

Lane Ph.R., Milesi-Ferretti G.M., *The external wealth of nations, mark II: Revised and extended estimates of foreign assets and liabilities 1970–2004*, "Journal of International Economics" 2007, no. 73.

Lerner A., *Economics of Control*, MacMillan, London 1949.

Lucas R.E. Jr., *Why Doesn't Capital Flow from Rich to Poor Countries?*, "The American Economic Review – American Economic Association" 1990, vol. 80, no. 2.

Marshall A., *Money, Credit and Commerce*, MacMillan, London 1924.

Mendoza E.G., *Sudden Stops, Financial Crises, and Leverage*, "American Economic Review – American Economic Association" 2010, vol. 100, no. 5.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Obstfeld M., *Expanding gross asset positions and the international monetary system*, Proceedings – Economic Policy Symposium, Federal Reserve Bank of Kansas City, Jackson Hole 2010.

Obstfeld M., Rogoff K., *The Mirage of Fixed Exchange Rates*, "Journal of Economic Perspectives" 1995, vol. 9, no. 4.

Obstfeld M. Rogoff K., *Global Current Account Imbalances and Exchange Rate Adjustments*, "Brookings Papers on Economic Activity" 2005, vol. 36, no. 1.

Robinson J., *Essays in the Theory of Employment*, Basil Blackwell, Oxford 1937.

Rogoff K., *The Purchasing Power Parity Puzzle*, "Journal of Economic Literature" 1996, no. 34(2).

Rosati D.K., *Polityka proeksportowa*, PWE, Warszawa 1990.

Rosati D.K., *Output decline during transition from plan to market: A reconsideration*, "The Economics of Transition" 1994, vol. 2, no. 4.

Rosati D.K., *Exchange rate policies during transition from plan to market*, "The Economics of Transition" 1996, vol. 4, no. 1.

Rosati D.K., *Polska droga do rynku*, PWE, Warszawa 1998.

Rosati D.K., *Dylematy polityki pieniężnej w warunkach swobody przepływów kapitału*, „Bank i Kredyt" 1999, nr. 7–8.

Rosati D.K., *Członkostwo Polski w Unii Gospodarczo-Walutowej. Kiedy i dlaczego?*, „Wykłady Otwarte" 2001, nr 14, Wyższa Szkoła Zarządzania i Administracji w Zamościu.

Rosati D.K., *The Balassa-Samuelson Effect in the EU Candidate Countries*, w: *Trade, Integration and Transition*, red. R. Grawe, A. Inotai, Institute for World Economics of the Hungarian Academy of Sciences, Budapest 2002.

Rosati D.K., *The Impact of EU Enlargement on Economic Disparities in Central and Eastern Europe*, w: *Shaping the New Europe. Economic Policy Challenges of European Union Enlargement*, red. M.A. Landesmann, D.K. Rosati, Palgrave MacMillan, London 2004.

Rother Ph.C., *The Impact of Productivity Differentials on Inflation and the Real Exchange Rate: an Estimation of the Balassa-Samuelson Effect in Slovenia*, "IMF Staff Report" 2000, no. 00/41.

Salvatore D., *International Economics*, wydanie 8, Wiley, Prentice Hall 2004.

Samuelson P.A., *Theoretical Notes on Trade Problems*, "Review of Economics and Statistics" 1964, vol. 23.

Sarno L., Taylor M.P., *Official Intervention in the Foreign Exchange Market: Is It Effective and, If So, How Does It Work?*, "Journal of Economic Literature" 2001, vol. XXXIX.

Weber W.E., *Do Sterilized Interventions Affect Exchange Rates?*, "Federal Reserve Rank of Minneapolis Quarterly Review" 1986, vol. 10, no. 3.

Williamson J., *The Exchange Rate System*, "Policy Analyses in International Economics" 1983, no. 5.

Williamson J., *What Washington Means by Policy Reform*, w: *Latin American Readjustment: How Much Has Happened*, red. J. Williamson, Institute for International Economics, Washington 1989.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozdział 7

POLITYKA MAKROOSTROŻNOŚCIOWA

7.1. Wprowadzenie

Polityka makroostrożnościowa stanowi nową dziedzinę polityki gospodarczej. Jej zadaniem jest regulowanie pracy sektora finansowego jako całości, w tym zwłaszcza sektora bankowego, w celu zapobiegania narastaniu ryzyka systemowego, wynikającego z istnienia szkodliwej cykliczności w zakresie funkcjonowania tego sektora. Zjawisko to polega na okresowym podejmowaniu przez instytucje finansowe decyzji inwestycyjnych obarczonych nadmiernym ryzykiem, co może prowadzić do występowania niebezpiecznych załamań rynku i kryzysów finansowych. O ile polityka makoostrożnościowa dotyczy w zasadzie wyłącznie sektora finansowego, o tyle jego znaczenie w przypadku prawidłowego funkcjonowania sektora realnego gospodarki oraz ogólnogospodarcza skala konsekwencji załamań rynkowych występujących w tym sektorze sprawiają, że polityka ta ma ważny wymiar makroekonomiczny.

Polityka makroostrożnościowa znalazła się w centrum uwagi wśród ekonomistów i polityków w okresie kryzysu finansowego i gospodarczego z lat 2008–2010. Okazało się wówczas, że tradycyjne instrumenty prowadzenia polityki regulacji i nadzoru nad indywidualnymi instytucjami finansowymi nie zapobiegły narastaniu zjawisk kryzysowych. W reakcji na tę sytuację Międzynarodowy Fundusz Walutowy wezwał do włączenia zagadnień związanych z polityką makroostrożnościową w zakres polityki gospodarczej na szczeblu makroekonomicznym (Blanchard, Dell'Ariccia i Mauro 2010). W niniejszym rozdziale przedstawimy najważniejsze argumenty na rzecz stosowania polityki makroostrożnościowej oraz omówimy jej główne cele, instrumenty i metody.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Kryzys finansowy z lat 2008–2010 ujawnił szereg niedoskonałości w funkcjonowaniu sektora bankowego w krajach rozwiniętych. Okazało się, że w głównej mierze to właśnie banki przyczyniały się do powstania kryzysu, udzielając nadmiernej ilości kredytów oraz lekkomyślnie inwestując w ryzykowne aktywa. Cykliczne pogorszenie się koniunktury gospodarczej, jakie miało miejsce w 2007 roku w krajach rozwiniętych, wywołało spadek popytu na nieruchomości i aktywa finansowe, co z kolei doprowadziło do obniżenia ich wartości rynkowej. To spowodowało zaś wzrost należności zagrożonych i konieczność tworzenia rezerw na pokrycie ewentualnych strat, wskutek czego wiele banków znalazło się w finansowych tarapatach. Rozbudowany system powiązań finansowych i właścicielskich między bankami oraz ich rosnąca współzależność w ramach globalnie zintegrowanych rynków finansowych sprawiały, że kłopoty poszczególnych banków rozprzestrzeniały się szybko na inne instytucje finansowe, zwiększając ryzyko załamania całego systemu bankowego. W ten sposób cykliczne osłabienie koniunktury w sektorze realnym ujawniło istotne strukturalne słabości sektora bankowego i w konsekwencji przerodziło się w 2008 roku w głęboki kryzys finansowy.

Trudności i zagrożenia, jakie pojawiły się wówczas w wielu bankach, przyniosły skutki daleko wykraczające poza sam sektor finansowy. Utrata wartości licznych aktywów posiadanych przez banki spowodowała z jednej strony zablokowanie rynku międzybankowego i ograniczenie skali kredytów dla gospodarki, z drugiej zaś strony zmusiła rządy poszczególnych państw do udzielania bankom zagrożonym niewypłacalnością pomocy finansowej na dużą skalę. W ten sposób pogarszająca się sytuacja finansowa wśród banków pogłębiała kryzys gospodarczy w sektorze realnym oraz stwarzała zagrożenie w zakresie stabilności finansów publicznych. Ta ostatnia zależność okazała się szczególnie groźna w krajach, w których miała miejsce hipertrofia sektora bankowego, a banki zgromadziły aktywa o wartości kilkakrotnie przekraczającej wielkość PKB (Islandia, Irlandia, Cypr). Skala pomocy finansowej, jakiej poszczególne rządy były zmuszone udzielić bankom, okazała się tak duża, że daleko wykraczała poza realne możliwości budżetowe poszczególnych państw. Zobowiązania podjęte z tego tytułu wywołały realną groźbę bankructwa rządów (bankructwa suwerennego – *sovereign default*). Kryzys lat z 2008–2010 ujawnił siłę wzajemnych powiązań między sektorem bankowym a sektorem finansów publicznych, które powodowały, że występujący w danym kraju kryzys bankowy prowadził do kryzysu zadłużenia publicznego, i odwrotnie – kryzys zadłużenia publicznego (suwerennego) zagrażał stabilności banków.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Działania na rzecz zlikwidowania tej szkodliwej współzależności stały się jednym z głównych kierunków polityki makroostrożnościowej.

Celem niniejszego rozdziału jest wyjaśnienie funkcjonowania mechanizmu narastania ryzyka systemowego – zarówno w poszczególnych instytucjach finansowych, jak i w całym sektorze bankowym – oraz pokazanie, w jaki sposób określone regulacje makroostrożnościowe mogą zapobiegać nasilaniu się tego zjawiska. Rozpoczniemy od przypomnienia specyfiki banków, uzasadniającej stosowanie zewnętrznego nadzoru nad ich działalnością. Następnie opisane zostaną źródła oraz mechanizm powstawania ryzyka systemowego w sektorze bankowym i pojawi się także krótkie przedstawienie wybranych sposobów pomiaru tego ryzyka. W dalszej kolejności omówione zostaną cele, zakres oraz instrumenty polityki makroostrożnościowej. Same zaś regulacje i politykę makroostrożnościową (*macroprudential regulations and policy*) będziemy definiować jako całościowe działanie władzy gospodarczej, którego celem jest identyfikacja i ograniczanie ryzyka systemowego oraz zapewnianie dzięki temu stabilności sektora finansowego jako całości (Crockett 2000, Brunnermeier i in. 2009, Billio i in. 2010, Clement 2010, IMF 2011a). W końcowej części rozdziału zajmiemy się zagadnieniami instytucjonalnymi i kwestiami aplikacyjnymi związanymi z polityką makroostrożnościową[1].

7.2. Potrzeba stosowania nadzoru nad sektorem bankowym

7.2.1. Specyfika banków

Banki stanowią specyficzną formę przedsiębiorstwa rynkowego. Ich działalność polega głównie na pośrednictwie finansowym, czyli na pozyskiwaniu oszczędności od gospodarstw domowych i przedsiębiorstw oraz na pożyczaniu ich innym gospodarstwom domowym i przedsiębiorstwom, a także sektorowi publicznemu. Zgodnie z obowiązującą w Polsce ustawą, działalność banków polega na „przyjmowaniu od osób fizycznych i prawnych środków finansowych w celu obciążania ich ryzykiem" w związku z udzielaniem kredytów i inwestowaniem w inne aktywa[2].

[1] Niniejszy rozdział stanowi rozwiniętą i zaktualizowaną wersję artykułu opublikowanego przez autora w czasopiśmie „Bank i Kredyt" (Rosati 2014).

[2] Art. 2 ustawy Prawo bankowe z dnia 29 sierpnia 1997 r. z późn. zm. stanowi, że „bank jest osobą prawną utworzoną zgodnie z przepisami ustaw, działającą na podstawie zezwoleń uprawniających do

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Specyfika banków polega więc po pierwsze na tym, że zarządzają one (a także inne instytucje finansowe, takie jak kasy oszczędnościowo-kredytowe) *środkami powierzonymi*, stanowiącymi własność innych osób prawnych i fizycznych. Dlatego też banki muszą podlegać szczególnie restrykcyjnym regulacjom prawnym i publicznemu nadzorowi w celu zapewnienia ich stabilności i wypłacalności oraz bezpieczeństwa środków, którymi obracają. Regulacje te wyznaczają ramy funkcjonowania poszczególnych banków i określają zasady prowadzonej przez nie działalności w zakresie udzielania kredytów, nabywania aktywów płynnych, zaciągania pożyczek, gromadzenia depozytów, wyceny ryzyka, tworzenia rezerw i zabezpieczeń oraz sprawozdawczości i komunikowania się z otoczeniem. Z kolei nadzór bankowy weryfikuje na bieżąco stopień przestrzegania przez poszczególne banki obowiązujących regulacji. Ponieważ regulacje te dotyczą z reguły funkcjonowania poszczególnych, indywidualnych instytucji finansowych, przyjęło się nazywać je regulacjami mikroostrożnościowymi.

Pomijając jednak kwestię zaufania, warto wskazać jeszcze na dwa inne ważne powody, w związku z którymi instytucje bankowe wymagają zewnętrznego, publicznego nadzoru. Jednym z nich jest *procykliczność* działalności banków (*procyclicality*), wyraźnie widoczna w silnym przyspieszaniu akcji kredytowej w okresach dobrej koniunktury i ostrym spowalnianiu akcji kredytowej w okresach kryzysu i recesji.

---

wykonywania czynności bankowych obciążających ryzykiem środki powierzone pod jakimkolwiek tytułem zwrotnym" (podkreślenia – D.R.). Z kolei art. 5.1 stanowi, że do czynności bankowych należy:

1) przyjmowanie wkładów pieniężnych płatnych na żądanie lub z nadejściem oznaczonego terminu oraz prowadzenie rachunków tych wkładów;
2) prowadzenie innych rachunków bankowych;
3) udzielanie kredytów;
4) udzielanie i potwierdzanie gwarancji bankowych oraz otwieranie i potwierdzanie akredytyw;
5) emitowanie bankowych papierów wartościowych;
6) przeprowadzanie bankowych rozliczeń pieniężnych;
7) wykonywanie innych czynności przewidzianych wyłącznie dla banku w odrębnych ustawach.

Do czynności bankowych zalicza się również następujące działania, o ile są one wykonywane przez banki:

8) udzielanie pożyczek pieniężnych;
9) operacje czekowe i wekslowe oraz operacje, których przedmiotem są warranty;
10) świadczenie usług płatniczych oraz wydawanie pieniądza elektronicznego;
11) terminowe operacje finansowe;
12) nabywanie i zbywanie wierzytelności pieniężnych;
13) przechowywanie przedmiotów i papierów wartościowych oraz udostępnianie skrytek sejfowych;
14) prowadzenie skupu i sprzedaży wartości dewizowych;
15) udzielanie i potwierdzanie poręczeń;
16) wykonywanie czynności zleconych, związanych z emisją papierów wartościowych;
17) pośrednictwo w dokonywaniu przekazów pieniężnych oraz rozliczeń w obrocie dewizowym.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Mamy w tym przypadku do czynienia z cyklem kredytowym, który w znacznej mierze pokrywa się z cyklem koniunkturalnym. Dlaczego banki udzielają jednak kredytów w sposób procykliczny?

Ma to związek ze sposobem mierzenia poziomu ryzyka w bilansie banku. Pomiar ten jest silnie uzależniony od stanu koniunktury. Ponieważ kredyty w bilansie banku figurują jako aktywa, to poziom udzielonych kredytów ma wpływ na ogólną wartość aktywów ważonych ryzykiem oraz liczony na ich podstawie współczynnik wypłacalności, zwany inaczej współczynnikiem adekwatności kapitałowej (*capital adequacy ratio – CAR*). Współczynnik ten wyraża stosunek kapitałów własnych netto banku do wartości aktywów ważonych ryzykiem:

$$CAR = \frac{K}{\sum_i w_i A_i}, \tag{7.1}$$

gdzie K oznacza kapitał własny banku, A_i jest wielkością bilansową (nominalną) aktywu i, podczas gdy $0 \leq w_i \leq 1$ odpowiada wadze ryzyka przypisanej danemu aktywowi.

Współczynnik wypłacalności banku pokazuje, jaką część utraty wartości aktywów bank jest w stanie pokryć kapitałem własnym. Władze nadzorcze ustalają wymaganą minimalną wielkość współczynnika CAR dla banków[3]. Przy obliczaniu wartości tego współczynnika poszczególnym pozycjom aktywów przydziela się określone stopnie ryzyka kredytowego, wyrażone w postaci wag procentowych. Wartość bilansową każdej pozycji aktywów mnoży się następnie przez stosowną wagę ryzyka w celu wyliczenia wartości ważonej ryzykiem.

Standardowo mierzony poziom ryzyka jest niski w okresie dobrej koniunktury (w tzw. „dobrych czasach"), co pozwala pomnażać aktywa – czyli podaż kredytów – bez konieczności zwiększania kapitału własnego. Również odwrotnie, w okresie złej koniunktury (w tzw. „złych czasach") poziom ryzyka rośnie, co zmusza banki do wyprzedaży aktywów i ograniczania kredytów przy danym poziomie kapitału własnego lub do zwiększania tego kapitału. W efekcie w „dobrych czasach" wzrost aktywów nominalnych nie zwiększa udziału aktywów ważonych ryzykiem, co może powodować nadmierną ekspansję kredytu. Z kolei w „złych czasach", kiedy rośnie ryzyko, wzrasta też wartość aktywów ważonych ryzykiem, co wymaga obniżania poziomu aktywów nominalnych (lub wzrostu kapitału własnego), a to nasila spadek cen aktywów i pogłębia kryzys. Zachodzi więc konieczność stosowania

[3] W przypadku banków europejskich wymagane minimum wynosi 8%.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zewnętrznych ograniczeń w ramach wspomnianej procykliczności, aby nie dopuścić do nadmiernych wahań kredytu i powstawania baniek spekulacyjnych.

Drugim powodem przemawiającym za koniecznością nadzorowania banków jest dominujący udział długu w finansowaniu aktywów, czyli wysoki poziom tzw. *dźwigni* (zwanej też lewarem – *leverage*). Utrzymywanie wysokiego udziału aktywów w relacji do kapitału własnego banku jest możliwe dzięki systemowi ułamkowej rezerwy obowiązkowej, który pozwala bankom utrzymywać w postaci płynnych rezerw tylko niewielką część pozyskanych depozytów (*fractional reserve system* – zob. rozdział 5). W ramach tego systemu banki korzystają z obcych źródeł finansowania na znacznie większą skalę niż wszystkie inne podmioty gospodarcze. Typowy udział wartości pożyczonych przez banki środków w stosunku do wartości wszystkich posiadanych aktywów w bankach detalicznych waha się w granicach 85–95%, a wśród banków inwestycyjnych jest często jeszcze wyższy.

Zależność ta jest szczególnie silna w okresach hossy, gdy rośnie udział aktywów. Gdy banki nabywają aktywa (papiery wartościowe, kredyty) obciążone niskim lub zerowym współczynnikiem ryzyka, wówczas nie wymaga to zwiększenia kapitału własnego, co oznacza, że przyrost aktywów może być finansowany długiem. W rezultacie wartość nominalna aktywów rośnie wraz z wartością zadłużenia w proporcji niemal 1:1, a to oznacza, że banki decydują się podejmować coraz większe potencjalne ryzyko. Przy wysokim poziomie zadłużenia (w relacji do kapitału własnego) nawet niewielki wzrost ujawnionego ryzyka lub spadek wartości aktywów wymaga jednak albo szybkiego zredukowania akcji kredytowej, albo gwałtownego zwiększenia kapitału własnego. W przeciwnym razie bankom grozi utrata płynności, a nawet niewypłacalność. Zewnętrzna regulacja jest więc niezbędna, aby nie dopuścić do nadmiernego „lewarowania" banków.

7.2.2. Nadzór mikroostrożnościowy i jego ograniczenia

Specyfika działalności banków wymaga zatem szczególnego nadzoru, który zapewni bezpieczeństwo wkładów, skłoni banki do utrzymywania dostatecznie dużego kapitału własnego i zapobiegnie dokonywaniu nadmiernie ryzykownych inwestycji. Funkcję tę tradycyjnie pełniły regulacje mikroostrożnościowe oraz instytucje publiczne odpowiedzialne za ich wdrażanie. Obostrzenia mikroostrożnościowe okazały się dość skuteczne w zapewnianiu stabilności i wypłacalności poszczególnych banków w tak zwanych „dobrych czasach", czyli w okresie dobrej koniunktury. Kryzys finansowy z lat 2008–2010 pokazał jednak, że nie były one

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w stanie zapewnić bezpieczeństwa i stabilności całego systemu finansowego. Stało się tak dlatego, że regulacje mikroostrożnościowe traktują sektor finansowy nie jako wewnętrznie powiązaną całość, lecz raczej jako zbiór odrębnych, niezależnych od siebie instytucji. W szczególności zaś w ramach tych regulacji nie bierze się pod uwagę trzech istotnych okoliczności.

Po pierwsze, wspomniana procykliczność banków sprawia, że w czasie kryzysu, gdy mamy do czynienia z rosnącym ryzykiem utraty wartości posiadanych aktywów, zapotrzebowanie na płynność wzrasta we wszystkich bankach jednocześnie. Tymczasem rynek międzybankowy i konwencjonalne źródła kapitału są tak skalibrowane, aby dostarczać płynność pojedynczym instytucjom, lecz nie wszystkim naraz. Stosowne regulacje mikroostrożnościowe, w trosce o pojedyncze banki, nie uwzględniają sytuacji *systemowego* niedoboru płynności lub kapitału, czyli niedoboru w skali całego sektora.

Po drugie, na rynkach finansowych może dochodzić do sytuacji, w których występuje konflikt między racjonalnością indywidualną a racjonalnością globalną. Działania, które wydają się racjonalne z perspektywy indywidualnego uczestnika rynku (a także w opinii regulatora oceniającego sytuację pojedynczej instytucji), mogą prowadzić do kryzysu systemowego, gdy podejmą je równocześnie (zbiorowo) wszyscy uczestnicy rynku. Tradycyjna polityka mikroostrożnościowa napotyka tu na tzw. błąd uogólnienia (*the fallacy of composition*), który polega na tym, że gdy działania rekomendowane na poziomie mikroekonomicznym w przypadku indywidualnej instytucji zostaną zastosowane przez wszystkie instytucje naraz, mogą doprowadzić do systemowego załamania całego sektora (IMF 2011a). Na przykład pozbywanie się przez indywidualnego inwestora aktywów uznanych przez niego za nadmiernie ryzykowne jest działaniem z jego punktu widzenia racjonalnym, lecz kiedy zaczynają postępować tak wszyscy inwestorzy, mamy wówczas do czynienia z masową, często paniczną wyprzedażą, która prowadzi do gwałtownego spadku cen aktywów, załamania rynku i ogromnych strat w bilansach banków. Podobne skutki pociąga za sobą zasada ograniczania strat (*stop-loss*), która narzuca bankom obowiązek sprzedaży aktywu, gdy jego cena rynkowa spadnie o określoną wartość procentową (lub poniżej określonego poziomu). W takich sytuacjach bardziej racjonalne – z punktu widzenia całego sektora – może być wstrzymanie się z wyprzedażą aktywów, aby nie doprowadzić do uruchomienia mechanizmu samospełniających się obaw.

Po trzecie, ze względu na wzajemne powiązania instytucji finansowych na rynku finansowym – zwłaszcza kredytowym – występują negatywne efekty zewnętrzne

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(*negative externality*). Są one spowodowane występowaniem obustronnych zależności i zobowiązań między bankami. Ta wzajemna ekspozycja banków powstaje bezpośrednio w ramach transakcji na rynku międzybankowym albo pośrednio w wyniku wspólnej ekspozycji wobec określonych rynków (np. rynku kredytów hipotecznych) bądź sektorów (np. sektora budowlanego). Istnienie wzajemnych ekspozycji oznacza, że udzielenie kredytu (lub nabycie innych aktywów) przez dany bank przyczynia się do większego obciążenia ryzykiem nie tylko tego banku, lecz także innych instytucji finansowych oraz pozostałych uczestników rynku kredytowego. Ponadto banki z reguły inwestują w podobne rodzaje aktywów, co czyni je podatnymi na spadek cen tych zasobów wywołany ich awaryjną wyprzedażą (*fire sale*), spowodowaną nagłą koniecznością pozyskania płynnych środków.

Mamy tu do czynienia ze szczególną odmianą efektu zewnętrznego (pozarynkowego), który sprawia, że obarczone określonym ryzykiem decyzje podejmowane przez jeden bank zawsze zwiększają ryzyko strat wśród pozostałych banków. Oznacza to, że na skutek decyzji pojedynczej instytucji rośnie nie tylko indywidualne ryzyko obciążające właśnie tę instytucję, ale wzrasta także ryzyko niestabilności całego sektora, które określa się w tym przypadku mianem *ryzyka systemowego*.

Powiązania występujące między bankami nie muszą oczywiście oznaczać automatycznie zwiększonej podatności systemu bankowego na szoki. W pewnych granicach i przy normalnych warunkach stają się one pożyteczne, ponieważ pozwalają absorbować szoki poprzez możliwość korzystania z pożyczek międzybankowych w celu uzupełniania płynności. Korzyści płynące z szerszego dostępu do źródeł finansowania przeważają wówczas nad zagrożeniami wynikającymi z rosnącej zależności między bankami. Na pewnym poziomie intensywności powiązania te zaczynają jednak kształtować silną współzależność, która w warunkach zwiększonego stresu coraz bardziej sprzyja przenoszeniu szoków między poszczególnymi instytucjami i w ten sposób przyczynia się do zwiększenia ryzyka niestabilności w ramach całego systemu (Haldane 2009; Acemoglu, Ozdaglar i Tahbaz-Salehi 2013).

To właśnie istnienie wzajemnych ekspozycji i wynikających z nich negatywnych efektów zewnętrznych w bankowości pozwala wyjaśnić, dlaczego upadek instytucji finansowej pociąga za sobą ryzyko systemowe, podczas gdy bankructwo przedsiębiorstwa działającego w innym sektorze gospodarki – na przykład w przemyśle lub górnictwie – nie stwarza takiego ryzyka. Upadek firmy produkującej jogurty bądź samochody nie wywołuje spadku ogólnej podaży tych wyrobów, ponieważ miejsce bankrutujących firm szybko zajmują inni dostawcy. Z kolei, gdy na rynku kredytowym upada jeden z banków, ograniczone zostaje wówczas dostarczanie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kredytu także przez inne banki. Przyczyną takiego stanu rzeczy jest występowanie wzajemnych ekspozycji, co w warunkach asymetrycznej informacji zmusza banki do gromadzenia rezerw i ograniczania akcji kredytowej.

Te trzy okoliczności – niedobór kapitału w okresie kryzysu, rozbieżności między racjonalnością na poziomie mikroekonomicznym a racjonalnością w skali makroekonomicznej oraz negatywne efekty zewnętrzne – sprawiają, że same regulacje mikroostrożnościowe, choć pozostają konieczne w celu zapewnienia stabilności pojedynczych instytucji, okazują się niewystarczające w przypadku stabilizowania całego systemu. W takiej sytuacji trzeba uzupełniać je regulacjami makroostrożnościowymi, których celem jest ograniczenie pojawiającego się ryzyka systemowego. Na tym właśnie polega główne zadanie polityki makroostrożnościowej.

7.3. Ryzyko systemowe w sektorze bankowym

7.3.1. Słabości wspólne oraz instytucje „ważne systemowo"

Ryzyko systemowe oznacza możliwość pojawienia się zaburzeń w funkcjonowaniu całego systemu finansowego i rynków finansowych, co związane jest z utratą wartości aktywów i prowadzi do gwałtownego ograniczenia lub wręcz wstrzymania dostarczania uczestnikom rynku ważnych produktów finansowych, takich jak kredyty, pożyczki, ubezpieczenia czy instrumenty zabezpieczające przed ryzykiem, które są niezbędne do prawidłowego funkcjonowania realnej sfery gospodarki (De Bandt i Hartmann 2000, s. 11; Adrian i Brunnermeier 2011; IMF 2011a, s. 7; IMF 2013, s. 6). Kluczowe znaczenie ma tu słowo „systemowe", wskazujące na możliwość wystąpienia nieprawidłowości w funkcjonowaniu całego lub niemal całego systemu finansowego.

Typowym przykładem takiej anomalii jest zjawisko „zacieśnienia kredytowego" (*credit crunch*), które powstaje wówczas, gdy straty – nie tylko rzeczywiste, ale także potencjalne – pojedynczych banków blokują sprawne działanie całego rynku międzybankowego, co z kolei prowadzi do gwałtownego wstrzymania udzielania kredytu dla całej gospodarki. Częstą przyczyną *credit crunch* jest pęknięcie bańki spekulacyjnej na rynkach aktywów. Napędzane pożyczkami zakupy aktywów wywołują początkowo silny, spekulacyjny wzrost ich cen. W efekcie ceny odchylają się od tzw. fundamentalnej wartości aktywów, określonej sumą zdyskontowanych przyszłych dochodów z danego aktywu. Nadmierny wzrost cen powoduje

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wreszcie zahamowanie popytu na dane aktywa, a to z kolei wpływa na załamanie cen. Wahania te nazywane są cyklami typu *boom-bust*, a ich skutkiem staje się często bankructwo uczestników rynku, spowodowane nadmiernymi inwestycjami finansowanymi długiem.

Wspomniane nieprawidłowości pojawiają się w dwóch sytuacjach. Po pierwsze, ich źródłem mogą być tzw. słabości wspólne (*common weakness, aggregate weakness*), czyli niekorzystne ukształtowanie się ważnych pozycji w bilansach banków, przy czym zjawisko to jest wspólne dla całego sektora bankowego lub jego zdecydowanej większości. Przykładem może być gromadzenie w aktywach banków kredytów hipotecznych zabezpieczanych nieruchomościami, których wartość pozostaje jednak zależna od stanu koniunktury. Takie słabości wspólne na ogół narastają w bilansach banków w fazie silnego ożywienia w cyklu koniunkturalnym i zanikają w fazie recesji. Ich powstawanie ma ścisły związek z omówioną wyżej procyklicznością działalności banków. Po drugie, zaburzenia te mogą wiązać się z upadkiem indywidualnych instytucji, które z określonych powodów są „ważne systemowo" (*systemically-important financial institutions* – SIFI), czyli podmiotów dostatecznie dużych oraz takich, które są silnie powiązane z wieloma innymi dużymi instytucjami[4]. Zgodnie z jedną z definicji, instytucja finansowa jest systemowo ważna, jeśli jej „niezdolność do regulowania zobowiązań wobec klientów i wierzycieli miałaby znaczne negatywne konsekwencje dla całego sektora finansowego i pozostałych części gospodarki" (Acharya, Engle i Richardson 2012)[5].

Pierwsza z opisanych sytuacji, czyli wspólna słabość bilansowa sektora, pojawia się wówczas, gdy sektor bankowy jako *całość* staje się narażony na nadmierne ryzyko z określonego tytułu na skutek pewnych zbiorowych zachowań banków. Słabości takie narastają z reguły systematycznie, zgodnie z przebiegiem cyklu koniunkturalnego, i przyjmują postać nadmiernego (w stosunku do możliwej absorpcji potencjalnych strat) wzrostu kredytu, przesadnego zadłużania się (lewarowania), w tym także w walutach obcych, oraz wygórowanego spekulacyjnego wzrostu cen

[4] Rozróżnienie między tymi dwoma sytuacjami pojawia się m.in w tzw. raporcie de Larosiere'a, gdzie czytamy: „O ile ryzyko systemowe może w zasadzie pojawić się w wyniku upadku pojedynczej instytucji finansowej, o ile jest ona dostatecznie duża w relacji całej gospodarki, i/lub posiada liczne oddziały w innych krajach, znacznie ważniejsze, globalne ryzyko systemowe powstaje na skutek istnienia wspólnej ekspozycji wielu instytucji finansowych wobec tego samego czynnika ryzyka. Analiza makroostrożnościowa musi zatem zwracać szczególną uwagę na wspólne, lub skorelowane, szoki, i szoki dla tych części system finansowego, które powodują efekty zarażenia" (de Larosière 2009, s. 38).

[5] Jobst i Grey (2013, s. 6–7) traktują obie przywołane sytuacje (słabość wspólna i upadek instytucji ważnej systemowo) jako podstawę do rozróżnienia modeli ryzyka systemowego odpowiednio na „partycypacyjne" (*participation risk*) i „kontrybucyjne" (*contribution risk*).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

aktywów i zabezpieczeń (*collaterals*). Oznacza to, że występuje wówczas wysoka dodatnia korelacja między stopami przychodu z aktywów w poszczególnych bankach (Acharya 2009). Wzrasta wówczas ogólne ryzyko systemowe, obejmujące poszczególne odcinkowe rodzaje ryzyka, takie jak ryzyko kredytowe (dłużnicy są nadmiernie zadłużeni, tracą pracę i nie spłacają kredytów), ryzyko rynkowe (wartość posiadanych przez banki zabezpieczeń spada), ryzyko płynności (mnożą się trudności w sprzedaży aktywów i w pozyskaniu płynnych środków na rynku finansowym), a także ryzyko stopy procentowej i ryzyko kursowe. Jeśli wysoka ekspozycja na wskazane rodzaje ryzyka jest wspólna w przypadku wielu instytucji (czyli zmiany dochodowości aktywów są dodatnio skorelowane w poszczególnych instytucjach sektora), to w sytuacji wystąpienia kryzysu wszystkie te organizacje wpadną w kłopoty. Wynika to stąd, że posiadane przez nie aktywa jednocześnie tracą na wartości, co powoduje równoczesny, powszechny wzrost zapotrzebowania na płynność i skokowy wzrost kosztu pozyskiwania kapitału. Ponieważ rynek pieniężny nie jest w stanie zaspokoić rosnących potrzeb płynnościowych wszystkich banków równocześnie, dochodzi do ograniczenia możliwości w zakresie dostarczania ważnych produktów finansowych dla gospodarki, takich jak kredyty czy ubezpieczenia[6].

Druga sytuacja, która rodzi ryzyko powstawania zaburzeń na rynku finansowym, to upadek dużej instytucji finansowej. Jak już wspomniano, bankructwo pojedynczej instytucji finansowej wpływa ujemnie na zdolność innych podmiotów do dostarczania ważnych produktów finansowych dla gospodarki, co stwarza zagrożenie dla prawidłowego funkcjonowania całego sektora. Dotyczy to zwłaszcza instytucji dużych, mających silne i rozległe powiązania z innymi instytucjami – podmioty takie określa się często mianem „zbyt dużych, aby upaść" (*too-big-to-fail* – TBTF) i „zbyt powiązanych, aby upaść" (*too-interconnected-to-fail* – TICTF). Instytucje TBTF można identyfikować na podstawie różnych kryteriów, takich jak wielkość poszczególnych podmiotów mierzona sumą aktywów w porównaniu z wielkością całego krajowego lub międzynarodowego systemu bankowego lub ich

[6] Przykładem słabości wspólnej była nadmierna ekspozycja banków amerykańskich na papiery wartościowe emitowane pod zastaw kredytów hipotecznych (*mortgage-backed securities* – MBS), zwłaszcza typu *subprime*. Gdy załamał się rynek nieruchomości, wzrosło ryzyko kredytowe w przypadku banków, ponieważ kredytobiorcy o niskiej zdolności kredytowej nie byli w stanie spłacać swoich zobowiązań, a ryzyko rynkowe wzrosło, gdyż wartość zabezpieczeń kredytów (ceny domów) spadała. Ze względu na to, że wysoka ekspozycja na MBS była słabością wspólną sektora bankowego, problemy te dotknęły wiele instytucji jednocześnie. Ponadto, ponieważ wiele banków finansowało się na rynku hurtowym, który uległ zamrożeniu, nastąpił wzrost ryzyka płynności. Wszystko to doprowadziło do głębokiego kryzysu.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

udział w rynku określonych usług finansowych z uwzględnieniem ewentualnych barier wejścia na dany rynek. Z kolei instytucje TICTF identyfikowane są nie tyle na podstawie ich wielkości, ile przede wszystkim pod względem intensywności powiązań, jakie łączą je z innymi instytucjami (miarą może być korelacja zmian wartości aktywów w poszczególnych instytucjach), oraz potencjalnego wpływu ich upadłości na pozostałe sektory gospodarki. Ponadto trzeba pamiętać, że ryzyko systemowe może powstać także w wyniku równoczesnego upadku większej liczby małych instytucji finansowych, z których żadna nie należy do kategorii TBTF lub TICTF (np. SKOK-i w polskim systemie bankowym). W takiej sytuacji również zagrożona staje się stabilność całego sektora (Brunnermeier i in. 2009).

Wspólnym elementem obu opisanych sytuacji jest zjawisko transmisji ryzyka indywidualnego między poszczególnymi bankami. Mechanizm ten przenosi ryzyko na pozostałe banki – jest to tak zwany efekt zarażenia – *contagion effect* – i w ten sposób stwarza zagrożenie dla stabilności całego systemu bankowego (De Bandt i Hartmann 2000). W mechanizmie tym można wyodrębnić cztery główne kanały przenoszenia negatywnych impulsów z banku zagrożonego na pozostałe instytucje:

- bezpośrednią ekspozycję innego banku wobec banku zagrożonego: jeśli inny bank posiada wierzytelności wobec banku zagrożonego, upadek tego drugiego wywołuje straty kapitałowe w banku posiadającym wierzytelności;
- awaryjną wyprzedaż aktywów (*fire sale*) przez zagrożony bank, co powoduje spadek cen rynkowych aktywów znajdujących się w posiadaniu innych banków oraz konieczność zaabsorbowania powstałych strat w bilansach tych instytucji (zgodnie z zasadą wyceny rynkowej aktywów *mark-to-market*);
- sytuację, w której inny bank jest stałym odbiorcą produktów finansowych z banku zagrożonego (np. korzysta z kredytu), a upadek tego drugiego powoduje nagłe wstrzymanie podaży określonych produktów i konieczność ograniczenia działalności przez bank z nich korzystający;
- wzrost kosztów pozyskania kapitału na rynku: w wyniku wycofywania kapitału z banku zagrożonego oraz innych instytucji i przenoszenia środków do bezpiecznych instrumentów finansowych występuje ostry brak kapitału na rynku międzybankowym, co przekłada się na wzrost kosztów pozyskania kapitału[7].

[7] Dobrym przykładem funkcjonowania mechanizmu transmisji ryzyka, ilustrującym skalę oraz siłę powiązań między bankami, jest upadek banku inwestycyjnego Lehman Brothers. Pociągnęło to za sobą nie tylko bezpośrednie straty samego banku (Lehman miał ponad 600 mld USD aktywów) i jego wierzycieli oraz straty instytucji, które Lehman finansował. Pośrednim, lecz niesłychanie ważnym skutkiem tego bankructwa, stał się gwałtowny wzrost kosztów finansowania na rynku. Ponieważ zobowiązania Lehman

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W celach analitycznych wyróżnia się często dwa wymiary ryzyka systemowego – wymiar czasowy oraz wymiar sektorowy. Z wymiarem czasowym mamy do czynienia wówczas, gdy ryzyko systemowe narasta w czasie dobrej koniunktury i maleje w warunkach złej koniunktury. Przyczyną takiego stanu rzeczy jest wspomniana procykliczność banków, czyli ich skłonność do nadmiernego zwiększania kredytu (finansowanego długiem) w okresie hossy (w fazie *boom* cyklu kredytowego) i gwałtownego ograniczania ryzykownych ekspozycji w okresie bessy (w fazie *bust* cyklu kredytowego). Procykliczność sprawia więc, że w okresie dobrej koniunktury w systemie finansowym następuje kumulacja ryzyka, co czyni go znacznie bardziej podatnym na szoki (Bank of England 2009; IMF 2011a, s. 8).

Z kolei z wymiarem sektorowym (lub przekrojowym – *cross-sectional*) mamy do czynienia wtedy, gdy wzrost ryzyka systemowego dotyczy określonego rynku (np. rynku kredytów hipotecznych) albo określonego sektora (np. sektora budownictwa), w ramach którego występują pewne wspólne słabości. Wymiar sektorowy uwzględnia zróżnicowany rozkład ryzyka w obrębie sektora finansowego i związany jest z nadmierną koncentracją działalności obarczonej ryzykiem w określonych instytucjach oraz ze wzajemnymi powiązaniami między nimi, co sprzyja rozpowszechnianiu się oraz intensyfikacji negatywnych szoków. Prowadzi to ostatecznie do kryzysu obejmującego cały sektor[8].

7.3.2. Rodzaje ryzyka systemowego

Ryzyko systemowe może powstawać w różnych obszarach działalności bankowej. Pozwala to wyróżnić jego cztery główne kategorie:

– ryzyko spowodowane (nadmiernie) szybkim wzrostem kredytów bankowych i wywołanym tym nadmiernym wzrostem cen aktywów. Zbyt duża ekspansja kredytu jest z reguły spowodowana nadmiernie optymistyczną oceną poziomu i zmian ryzyka w okresie dobrej koniunktury przez uczestników

Brothers były rozproszone praktycznie w całym sektorze bankowym, wiele instytucji musiało tworzyć odpisy na zagrożone aktywa ulokowane w bankrutującym banku, co spowodowało gwałtowne pogorszenie ich sytuacji finansowej. W warunkach braku pełnej informacji o stopniu zagrożenia tych ekspozycji silnie wzrastało ryzyko kontrahenta (*counterparty risk*). W takiej sytuacji banki dysponujące nadwyżką płynności wstrzymywały się z udzielaniem pożyczek, ponieważ nie wiedziały, w jakiej kondycji finansowej znajdują się ich potencjalni partnerzy. Skutkiem tego stało się zamrożenie całego rynku pożyczek międzybankowych w USA. To z kolei zmusiło banki mające kłopoty z płynnością do wyprzedaży aktywów w trybie awaryjnym, panicznym (*fire sale*), co doprowadziło do silnego spadku ich cen oraz poważnych strat bilansowych.

[8] W praktyce oba wymiary ryzyka są ściśle ze sobą powiązane (IMF 2011a).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

rynku, co skłania ich do przeprowadzania nadmiernych inwestycji finansowanych z długu. Zjawisko to dostrzegał już Keynes, który traktował je jako typowy przejaw tzw. „instynktów zwierzęcych" (*animal spirits*), cechujących uczestników rynku. Szczególną rolę w tym procesie odgrywa mechanizm akceleratora finansowego, który wyjaśnia, dlaczego zachowania uczestników rynku w warunkach powiązania sektora realnego i sektora finansowego przyczyniają się do wzrostu ryzyka (zob. ramka 7.1);

– ryzyko związane z nadmiernym lewarowaniem – czyli nabywaniem aktywów finansowanych długiem i skutkami następującego później delewarowania. Lewar (*leverage*) stanowi wielkość długu zaciągniętego na zakup aktywów i mierzony jest stosunkiem wartości (księgowej) wszystkich aktywów do kapitału własnego (*assets/equity*). Relację tę uznaje się umownie za nadmierną, gdy jej wielkość przekracza poziom historycznego trendu o jedno odchylenie standardowe (Lim i in. 2011, s. 8). Kiedy lewar jest nadmierny, nawet niewielki spadek ceny aktywów wywołuje techniczną niewypłacalność, ponieważ wartość aktywów spada poniżej wartości zadłużenia;
– ryzyko związane z niedostateczną płynnością. Systemowe ryzyko płynności powstaje wówczas, gdy większość banków odczuwa w danym momencie niedostatek płynności. W sektorze stanowiącym pewną całość występuje więc wówczas generalnie za mało płynności, a instytucje finansowe nie są w stanie jej szybko uzupełnić w normalnym trybie zasilania (np. zaciągając krótkookresowe pożyczki na rynku międzybankowym);
– ryzyko związane z transgranicznymi przepływami kapitału na wielką skalę, w tym w zakresie pożyczek i kredytów. W warunkach pełnej swobody transgranicznych przepływów kapitału może dochodzić do epizodów gwałtownego napływu i odpływu kapitału z danego kraju. W tym drugim przypadku migracja krótkoterminowego kapitału pożyczkowego może doprowadzić do drastycznego kryzysu płynności w sektorze bankowym i osłabienia kursu waluty, co skutkuje ostatecznie pełnowymiarowym kryzysem walutowym i bankowym. Ten rodzaj ryzyka zbliżony jest do ryzyka niedostatecznej płynności, lecz różni się od niego jednocześnie tym, że obejmuje także ryzyko zmian kursu walutowego i ryzyko makroekonomiczne kraju[9].

[9] Billio i in. (2010) proponują nieco inny podział ryzyka systemowego, wyróżniając ryzyko płynności (*liquidity*), ryzyko lewarowania (*leverage*), ryzyko powiązań (*linkages*) i ryzyko strat (*losses*).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 7.1. Akcelerator finansowy

Koncepcję akceleratora finansowego jako pierwsi opisali niezależnie od siebie amerykańscy ekonomiści Hyman Minski (1992) i Charles Kindleberger (1996). Akcelerator finansowy (*financial accelerator*) opisuje szczególny rodzaj zależności między sektorem realnym a sektorem finansowym w gospodarce. Aby inwestować, firmy muszą zaciągać kredyty, a skala uzyskanych kredytów zależy od posiadanych zabezpieczeń (*collateral*), wycenianych według bieżącej wartości rynkowej. Zgodnie z tym mechanizmem, przesadny optymizm (*irrational exuberance*) co do przyszłych zysków skłania przedsiębiorstwa i osoby prywatne do dokonywania nadmiernych inwestycji. Powoduje to napływ kapitału na rynek, wzrost cen aktywów oraz wzrost rynkowej wartości zabezpieczeń (wzrost względnych cen aktywów), co sprzyja dalszemu zadłużaniu się. W ten sposób nadmierny optymizm skłania firmy i gospodarstwa domowe do zwiększania poziomu zadłużenia oraz wydatków. Gdy– prędzej czy później – rzeczywistość okazuje się jednak gorsza od oczekiwań, ceny aktywów i wartości zabezpieczeń przestają rosnąć i zaczynają spadać. Akcelerator finansowy ulega odwróceniu – kapitał odpływa, ceny aktywów spadają, wartość zabezpieczeń maleje. Wywołuje to pogorszenie bilansów firm, spadek ich wartości rynkowej netto (*net worth*) i relatywny wzrost zadłużenia (relacja długu do aktywów rośnie), co ogranicza zdolność uczestników rynku do obsługi długu i pozyskiwania nowych kredytów. Miejsce przesadnego optymizmu zajmuje nadmierny pesymizm, a wartość inwestycji spada jeszcze bardziej. Akcelerator jest mechanizmem silnie procyklicznym – powoduje wzmocnienie i przedłużenie boomu oraz pogłębienie i przedłużenie recesji (zob. także Bernanke i Gertler 1999; Bernanke, Gertler i Gilchrist 1999).

Ryzyko systemowe powstaje pod wpływem kumulowania się różnych rodzajów ryzyka indywidualnego w poszczególnych instytucjach bankowych. Na ryzyko indywidualne składają się: ryzyko kredytowe, ryzyko płynności, ryzyko rynkowe, ryzyko stopy procentowej oraz ryzyko kursowe. Banki wyceniają skalę ryzyka w każdej z tych kategorii, otrzymując w ten sposób ogólną ocenę ryzyka w przypadku danego banku. Łączne ryzyko w zakresie całego systemu okazuje się jednak zawsze większe niż prosta suma ryzyka dotyczącego poszczególnych instytucji, ponieważ wycena ryzyka na poziomie mikroekonomicznym nie uwzględnia negatywnych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

efektów zewnętrznych, jakie działania określonych instytucji niosą dla pozostałych uczestników rynku. Nie bierze się też pod uwagę tego, że kapitałowe i płynnościowe potrzeby poszczególnych banków mogą okazać się niemożliwe do zaspokojenia w warunkach kryzysu. Pojedynczy bank nie jest bowiem w stanie samodzielnie ocenić, jaki wpływ na jego sytuację finansową oraz jego indywidualne ryzyko mogą mieć decyzje podejmowane przez inne banki. W związku z tym zachodzi potrzeba dokonywania analizy i ograniczania ryzyka także z perspektywy całego systemu finansowego (bankowego). Takiego ujęcia zabrakło w polityce regulacji i nadzoru w okresie poprzedzającym kryzys finansowy z lat 2008–2010. Ryzyko systemowe wymknęło się spod kontroli także dlatego, że w poszczególnych krajach zabrakło instytucji jednoznacznie odpowiedzialnej za wycenę i ograniczanie ryzyka systemowego, dysponującej niezbędnymi informacjami oraz instrumentami działania (Goodhard 2011). Żadna instytucja nadzorcza nie została wyposażona w stosowny mandat, nakładający na nią obowiązek wyceny i ograniczania ryzyka systemowego.

Wymienione rodzaje ryzyka indywidualnego mogą obciążać zarówno aktywa, jak i pasywa bilansu banków. Na przykład ryzyko kredytowe związane z udzielonym kredytem dotyczy aktywów banku, ryzyko płynności odnosi się do pasywów, a ryzyko kursowe może wiązać się ze stanem zarówno aktywów, jak i pasywów. Jest to obserwacja istotna, ponieważ w dotychczasowej praktyce władze nadzorcze wykorzystywały głównie regulacje mikroostrożnościowe dotyczące ryzyka powstającego po stronie aktywów bankowych, nie uwzględniając natomiast w dostatecznym stopniu ryzyka po stronie pasywów. Tymczasem skuteczność polityki makroostrożnościowej pozostaje zależna od tego, czy instrumenty stosowane w celu przeprowadzenia redukcji ryzyka systemowego uwzględniają umiejscowienie ryzyka indywidualnego i właściwie je adresują.

7.3.3. Pomiar ryzyka systemowego

Istotnym problemem praktycznym staje się identyfikacja i właściwy pomiar wielkości ryzyka systemowego. Wykorzystuje się w tym celu szereg modeli ekonometrycznych, w ramach których definiuje się i szacuje ryzyko systemowe jako funkcję ryzyka indywidualnego poszczególnych banków w zależności od ich wielkości, stopnia zadłużenia i skali wzajemnych powiązań. Modele te można podzielić na dwie kategorie – modele strukturalne, uwzględniające współzależności między wieloma zmiennymi, oraz modele o postaci zredukowanej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W przypadku pierwszej kategorii najczęściej wykorzystuje się tzw. analizę należności (roszczeń) uwarunkowanych (*contingent claims analysis* – CCA) służącą wycenie wartości aktywów obciążonych ryzykiem, a w konsekwencji także w celu oszacowania ryzyka niewypłacalności indywidualnej instytucji finansowej. W ramach tej metody popularne jest wykorzystanie klasycznego modelu wyceny opcji Blacka-Scholesa-Mertona (BSM) w celu ustalenia wartości aktywów i pasywów banku z uwzględnieniem ryzyka. Metodę tę można następnie zastosować do oszacowania ryzyka systemowego, definiowanego jako łączne ryzyko niewypłacalności wielu instytucji finansowych jednocześnie. W tym kontekście Grey i Jobst (2011) oraz Jobst i Grey (2013) rozwijają systemową analizę należności uwarunkowanych (*Systemic CCA* – SCCA). Składa się ona z dwóch etapów. Najpierw szacuje się rozkłady prawdopodobieństwa oczekiwanych strat indywidualnych instytucji finansowych w warunkach wystąpienia załamania rynkowego, przy czym straty te są definiowane jako opcje *put* właściciela danej instytucji[10]. W drugim kroku wycenia się całościowe ryzyko systemowe na podstawie łącznego rozkładu prawdopodobieństwa w przypadku oczekiwanych strat wszystkich objętych analizą instytucji. Metodę tę stosuje Międzynarodowy Fundusz Walutowy w badaniach dotyczących stabilności systemów bankowych w poszczególnych krajach (zob. m.in. IMF 2010b, 2011c i 2011d).

Również Segoviano i Goodhart (2009) wykorzystują koncepcję łącznego rozkładu prawdopodobieństwa, analizując różne miary stabilności sektora bankowego jako całości. Zaproponowana przez autorów funkcja gęstości prawdopodobieństwa wielu bankowych zmiennych bilansowych (*Banking System Multivariate Density* – BSMD) odnosi się zarówno do indywidualnych, jak i łącznych rozkładów zmian wartości aktywów sektora bankowego. Funkcja BSMD pozwala wychwycić wzajemne zależności między różnymi stopniami ryzyka finansowego załamania (*financial distress*) w poszczególnych bankach, dzięki czemu zapewnia potencjalną możliwość uwzględnienia negatywnych efektów zewnętrznych wynikających z istnienia wzajemnych ekspozycji między bankami[11]. Jest to istotna przewaga tego modelu nad tradycyjnymi miarami ryzyka, które z reguły odnoszą się jedynie do dwóch zmiennych (wartości rynkowej aktywów oraz określonego czynnika ryzyka np. kryzysu finansowego) i uwzględniają jedynie proste, liniowe zależności (korelacje) między nimi. Segoviano i Goodhart (2009) wprowadzają także ciekawą miarę

[10] Wystawca opcji *put* gwarantuje nabywcy opcji (właścicielowi instrumentu bazowego) prawo do sprzedaży instrumentu bazowego wystawcy opcji po z góry ustalonej cenie.

[11] Segoviano (2009) proponuje określoną ekonometryczną metodę szacowania funkcji BSM.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

systemowego znaczenia danego banku, zdefiniowaną jako wielkość warunkowa prawdopodobieństwa, że co najmniej jeden bank znajdzie się w stanie niewypłacalności, pod warunkiem, że pewien określony bank A doświadczy takiej właśnie sytuacji (*Probability that at Least One Bank becomes Distressed* – PAO). Im wyższa stanie się wartość PAO, tym większe będzie „systemowe" znaczenie banku A. Wskaźnik PAO można więc zastosować w celu określanie, które banki są „zbyt duże, aby upaść" (TBTF). Podejście to rozwija następnie Zhou (2010).

W ostatnich latach coraz popularniejsze stają się modele w postaci zredukowanej, gdzie wskaźnik ryzyka systemowego jest określoną funkcją niedoboru kapitału w warunkach kryzysowych. Do najczęściej omawianych należą te, w których miarą ryzyka systemowego jest potencjalny niedobór kapitału w systemie bankowym. W modelu MES krańcowy oczekiwany niedobór kapitału w systemie bankowym (*Marginal Expected Shortage* – MES) jest spowodowany niedoborem kapitału w indywidualnym banku (lub innej instytucji finansowej) w warunkach wystąpienia kryzysu (zob. np. Acharya i in. 2011; Brownlees i Engle 2011; Acharya, Engle i Richardson 2012). Punktem wyjścia przy określaniu miary MES jest wskaźnik VaR (*value-at-risk*), czyli najbardziej popularna miara ryzyka finansowego w przypadku pojedynczego banku. VaR definiuje się najprościej jako wartość progową (maksymalną) możliwej straty na portfelu aktywów banku w danym czasie i z określonym prawdopodobieństwem wystąpienia.

7.3.4. Modele ryzyka systemowego*

Poniżej przedstawimy krótkie omówienie najważniejszych metod szacowania ryzyka systemowego. Oznaczmy wartość aktywów w posiadaniu banku *i* jako X_i, a potencjalną stratę na aktywach tego banku jako L_i. W przypadku danego okresu, VaR można więc określić przy pomocy następującej formuły:

$$\Pr\left(L_i > VaR_i\right) = q, \tag{7.2}$$

która pokazuje, że prawdopodobieństwo, iż spadek wartości portfela aktywów X_i danego banku *i* będzie większy niż wartość progowa VaR_i, wynosi *q*. Oczekiwany niedobór kapitału (*Expected Shortage – ES*) w banku *i*, pod warunkiem wystąpienia straty na aktywach przekraczającej VaR z prawdopodobieństwem *q*, jest zatem równy[12]:

[12] Wielkość oczekiwanej straty zależy oczywiście od rozkładu prawdopodobieństwa: im grubszy staje się „ogon" rozkładu, tym większa będzie możliwa strata w przypadku danego *q*.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$ES_i^q = E\left(L_i \middle| L_i \rangle VaR_i^q\right). \tag{7.3}$$

Wówczas MES_i^q oznacza krańcowy, oczekiwany niedobór kapitału, który pojawia się w sektorze bankowym na skutek wystąpienia strat w banku *i*, pod warunkiem, że straty całego sektora przekroczą VaR (czyli w sytuacji kryzysu):

$$MES_i^q = E\left(L_i | L > VaR^q\right). \tag{7.4}$$

Równanie (7.4) można więc interpretować jako krańcowy, oczekiwany przyrost niedoboru kapitału w całym sektorze bankowym, spowodowany niedoborem kapitału w banku *i* w warunkach wystąpienia kryzysu finansowego. Wielkość MES jest więc miarą wpływu sytuacji pojedynczego banku na ryzyko systemowe.

Oszacowania wielkości MES można dokonać na podstawie analizy bilansu bankowego. Równanie bilansowe aktywów i pasywów banku *i* (dla $i = 1, 2,\ldots, n$) można zapisać następująco:

$$D_i + C_i = \sum_{j=1}^{k} X_{i,j}, \tag{7.5}$$

gdzie D_i oznacza wszystkie zobowiązania, C_i to kapitał własny banku według aktualnej wyceny rynkowej (kapitalizacja), a elementy $X_{i,j}$ (dla $j = 1,2,\ldots,m$) stanowią aktywa banku *i* według aktualnej wyceny rynkowej. $MES_{i,t}$, czyli niedobór kapitału w banku *i* w okresie *t*, wystąpi wówczas, gdy:

$$k\left(D_i + C_i\right) - C_i > 0, \tag{7.6}$$

czyli:

$$k > \frac{C_i}{\sum_{j=1}^{k} X_{i,j}}, \tag{7.7}$$

gdzie *k* oznacza nominalny współczynnik adekwatności kapitałowej (stosunek kapitału własnego do sumy aktywów).

Wielkość MES w odniesieniu do danego banku oznacza udział tego podmiotu w ryzyku systemowym całego sektora w warunkach powstania kryzysu. Miarą ryzyka systemowego jest więc w tym przypadku kwota kapitału, jakim należy dofinansować dany bank, aby przywrócić bezpieczny poziom kapitału w stosunku do

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

aktywów, zapewniający dalsze niezakłócone funkcjonowanie tego banku. Kwota niezbędnego dokapitalizowania powinna zostać wpłacona przez właścicieli albo pozyskana na rynku. Lecz w warunkach kryzysu dodatkowy kapitał może być niedostępny w sektorze bankowym i w takiej sytuacji ciężar dokapitalizowania spada często na rząd, czyli podatników, ewentualnie z udziałem instytucji gwarancyjnych, finansowanych przez same banki.

Techniczną odmianą MES jest model wykorzystujący miarę ryzyka systemowego SRISK (Acharya, Engle i Richardson 2011). Zgodnie z tą koncepcją, $SRISK_i$ jest równe oczekiwanej wielkości niedoboru wymaganego kapitału własnego banku *i* w warunkach wystąpienia kryzysu:

$$SRISK_i = E(\textit{niedobór kapitału} | \textit{kryzys}).$$

Tak więc:

$$\begin{aligned} SRISK_i &= E[k(D_i + C_i) - C_i \mid \textit{kryzys}] \\ &= kD_i + k(1 - LRMES_i)C_i - (1 - LRMES_i)C_i \\ &= kD_i - (1-k)(1 - LRMES_i)C_i, \end{aligned} \tag{7.8}$$

gdzie $LRMES_i$ (*long-run marginal expected shortfall)* jest oczekiwanym procentowym spadkiem wartości kapitału własnego banku w warunkach kryzysu w określonym czasie.

Równanie (7.8) pokazuje, że przy danym współczynniku *k* wielkość ryzyka systemowego generowanego przez bank *i* jest tym mniejsza, im większy staje się kapitał własny banku C_i, oraz tym większa, im większe będzie jego zadłużenie D_i i potencjalna utrata wartości kapitału w warunkach kryzysu $LRMES_i$. Gdy $SRISK_{i,t} > 0$, dany bank jest narażony na ryzyko płynności lub niewypłacalności w okresie kryzysu. Całkowite ryzyko systemowe w sektorze bankowym w danym okresie pozostaje równe sumie (dodatnich) wartości $SRISK_i$ dla wszystkich *i*, czyli sumie ryzyka indywidualnego poszczególnych banków.

Jeszcze inną techniczną odmianę modelu MES stanowi zaproponowana przez Huang, Zhou i Zhu (2011) koncepcja wykorzystania hipotetycznych stawek ubezpieczeń od niewypłacalności banków. Ryzyko systemowe jest w tym przypadku mierzone kosztem ubezpieczenia w warunkach kryzysu (*distress insurance premium* – DIP):

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$DIP = E(L \mid L \geq L_m). \tag{7.9}$$

Choć formuła (7.9) wykazuje wyraźne podobieństwo do wskaźników ES i MES, należy zauważyć też pewne różnice występujące między tymi miarami. Po pierwsze, wskaźnik DIP odnosi się do określonego progowego poziomu strat L_m, podczas gdy wskaźniki ES, MES i VaR dotyczą określonego przedziału funkcji rozkładu prawdopodobieństwa. Po drugie, wskaźnik DIP jest szacowany na podstawie empirycznych obserwacji rynkowych stawek CDS (*credit default swap*). Jak pokazują jednak Huang, Zhou i Zhu (2011), wyniki otrzymane w ramach szacowania ryzyka systemowego przy wykorzystaniu różnych miar są bardzo zbliżone[13].

Na całkowite ryzyko systemowe składają się indywidualne udziały poszczególnych banków. Każdy krańcowy „wkład" do ryzyka systemowego równy jest oczekiwanej stracie danego banku *i*, przy założeniu, że kryzys obejmie cały system:

$$\partial DIP \big/ \partial L_i = E(L_i \mid L \geq L_m), \tag{7.10}$$

podczas gdy całe ryzyko systemowe równe jest sumie indywidualnych „wkładów" poszczególnych banków: $\sum_i^n \left(\partial DIP \big/ \partial L_i \right)$.

Wskaźniki MES, SRISK i DIP mają kilka istotnych zalet, dzięki którym mogą znaleźć wiele praktycznych zastosowań w nadzorze bankowym. Po pierwsze, są wyrażone w wartościach pieniężnych, co umożliwia agregację w skali sektora i pozwala bezpośrednio szacować wielkość potencjalnych potrzeb kapitałowych sektora bankowego w sytuacji kryzysowej. Ponadto, ich wyliczenie nie nastręcza większych kłopotów dzięki dostępności danych bilansowych dotyczących poszczególnych banków, w tym kapitału akcyjnego[14]. Problemem jest jednak to, że wskaźniki te odzwierciedlają udział poszczególnych banków w niedoborze kapitału w całym sektorze bankowym, podczas gdy sam kryzys jest traktowany zasadniczo

[13] W grupie 19 banków współczynniki korelacji rang Spearmana z zastosowaniem różnych miar ryzyka (MES, DIP, CoVaR i SCAP) kształtują się między 0,67 a 0,90 (Huang, Zhou i Zhu 2011, tabela 6, s. 36).

[14] Wskaźnik SRISK w przypadku banków amerykańskich jest obliczany regularnie, a dostęp do niego można uzyskać na stronie Volatility Lab of NYU Stern School, z kolei analogiczny wskaźnik odnośnie banków europejskich można znaleźć na stronie Centrum Zarządzania Ryzykiem (Center of Risk Management – CRML) Szkoły Wyższych Studiów Handlowych Uniwersytetu w Lozannie (l'Ecole des Hautes Etudes Commerciales de l'Université de Lausanne).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jako zmienna egzogeniczna. Przyjmuje się bowiem założenie (Acharya, Engle i Richardson 2012, s. 2), że gdy w jednym banku zabraknie kapitału, inne banki uzupełnią powstałą lukę w dostarczaniu ważnych produktów finansowych na rynek. W rzeczywistości sytuacja ta wygląda jednak odmiennie. Upadek jednego banku, zwłaszcza dużego, powoduje najczęściej zamrożenie rynku międzybankowego ze względu na asymetrię i niedostatek informacji, wzrost ryzyka kontrahenta i w rezultacie ograniczenie dostarczania produktów finansowych na rynek. Oznacza to, że kryzys może wystąpić nie tylko wówczas, gdy wartość kapitału całego sektora spadnie poniżej określonego progu, lecz także w sytuacji, gdy wartość zagregowanego kapitału znajdzie się powyżej tego progu, ale mechanizm przepływu środków do banków potrzebujących zastrzyku kapitału zostanie zablokowany. W warunkach istniejących powiązań miedzy bankami kryzys finansowy stanowi więc najczęściej zmienną endogeniczną. Opisane modele nie uwzględniają zatem w dostatecznym stopniu wszystkich efektów zewnętrznych.

Istnieje wreszcie jeszcze inne podejście, które proponują Adrian i Brunnermeier (2011), rozwijając wcześniejsze wyniki badań Engle'a i Manganellego (2004). Podobnie jak w koncepcji MES punktem wyjścia jest indywidualny miernik ryzyka VaR. Ponieważ wskaźnik ten mierzy jednak ryzyko danej instytucji finansowej w izolacji od innych podmiotów, Adrian i Brunnermeier postulują wykorzystanie formuły VaR w odniesieniu do sektora bankowego jako całości, przy uwzględnieniu powiązań występujących między poszczególnymi instytucjami. Tę nową miarę ryzyka systemowego autorzy definiują jako CoVaR. Wówczas $CoVaR^{q}_{S|C(X_i)}$ oznacza warunkową wielkość VaR w przypadku całego systemu bankowego, zależną od pewnego wydarzenia kryzysowego $C(X_i)$, które obniża dochodowość aktywów X_i banku i. Zakładając, że prawdopodobieństwo wystąpienia spadku wartości wszystkich aktywów w ramach całego systemu bankowego X_S większego lub równego $CoVaR_{S|C(X_i)}$ wynosi q, i podstawiając otrzymane zmienne do równania (7.2), można zapisać następujący wzór:

$$Pr(X_S \leq CoVaR^{q}_{S|C(X_i)} \mid C(X_i)) = q. \tag{7.11}$$

Różnica pomiędzy CoVaR w warunkach kryzysu ($CoVaR^{q}_{S|C(X_i)}$) a CoVaR w „normalnych" warunkach rynkowych ($CoVaR^{q}_{S|Median(X_i)}$) mierzy udział banku i w całościowym ryzyku systemowym:

$$\Delta CoVaR^{q}_{S} = CoVaR^{q}_{S|C(X_i)} - CoVaR^{q}_{S|Mediana(X_i)}. \tag{7.12}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Podobnie jak MES także ΔCoVaR jest miarą krańcową. Główną zaletą koncepcji ΔCoVaR pozostaje to, że uwzględnia ona w pewnym zakresie negatywne efekty zewnętrzne wynikające z powiazań występujących między bankami[15]. O ile VaR odnosi się do ryzyka na poziomie pojedynczej instytucji, o tyle ΔCoVaR mierzy udział poszczególnych instytucji w całkowitym ryzyku systemowym sektora. Znajomość konkretnych wartości ΔCoVaR pozwala zidentyfikować banki o największym potencjalnie negatywnym wpływie na stabilność systemu. Adrian i Brunnermeier (2011) szacują ΔCoVaR na podstawie tygodniowych odczytów rynkowej wartości aktywów instytucji finansowych notowanych w obrocie publicznym, wykorzystując technikę regresji kwantylowej. Autorzy proponują także „odwróconą" odmianę wskaźnika CoVaR, która mierzy wpływ kryzysu finansowego na ryzyko obciążające pojedynczą instytucję finansową.

Omówione metody mierzą ryzyko systemowe jako „warunkową" wielkość oczekiwanych strat na aktywach systemu w danym okresie i przy określonym prawdopodobieństwie, a także pozwalają wyznaczyć związany z tym „warunkowy" niedobór kapitału własnego banku w sytuacji wystąpienia kryzysu. Można więc traktować je jako różne odmiany testu warunków skrajnych (*stress test*), jaki przeprowadza władza nadzorcza. Wskazane wcześniej mierniki ES, MES, SRISK i DIP odzwierciedlają skalę potrzeb w zakresie niezbędnego dokapitalizowania poszczególnych banków i całego sektora w warunkach gwałtownego pogorszenia się sytuacji rynkowej. Modele te nie uwzględniają jednak pozostałych negatywnych efektów zewnętrznych, związanych z podejmowaniem przez poszczególne banki decyzji inwestycyjnych obarczonych ryzykiem. Tymczasem występowanie wzajemnych powiązań i spirali zależności (*spillovers*) między bankami, powstających bezpośrednio na skutek ekspozycji w ramach transakcji na rynku międzybankowym albo pośrednio w wyniku wspólnej ekspozycji wobec określonych rynków lub sektorów, generuje koszty zewnętrzne w postaci „efektu zarażenia" (Rochet i Tirole 1996), awaryjnej wyprzedaży aktywów (*fire sale* – Allen i Gale 2007), spirali braku płynności (Brunnermeier i Pedersen 2007) oraz epizodów paniki bankowej (Diamond i Dybvig 1983). Ze ściślejszym pomiarem ryzyka systemowego mamy do czynienia w przypadku miernika CoVaR, który pośrednio uwzględnia także ryzyko związane z występowaniem wzajemnych powiązań między bankami (Adrian i Brunnermeier 2011). Wadą wskaźnika CoVaR jest natomiast to, że nie pozwala on prawidłowo oszacować całościowego udziału ryzyka indywidualnego

[15] Analiza Adriana i Brunnermeiera dotyczy wszystkich instytucji finansowych, nie tylko banków.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

poszczególnych banków w ryzyku systemowym. Z kolei metody wykorzystujące łączne funkcje prawdopodobieństwa wielu zmiennych, w tym zwłaszcza systemowa analiza CCA (SCCA), uwzględniają negatywne efekty zewnętrzne w sposób bezpośredni.

7.4. Pojęcie, cele i zakres polityki makroostrożnościowej

7.4.1. Regulacja „mikro" i regulacja „makro"

Sam termin „makroostrożnościowy" (*macroprudential*) pojawił się w literaturze finansowej stosunkowo niedawno, bo dopiero w końcu lat 70. (IMF 2011a; Clement 2010). Tymczasem polityka makroostrożnościowa (choć wówczas nie używano jeszcze tego określenia) i jej poszczególne instrumenty znajdowały jednak zastosowanie już znacznie wcześniej – na przykład w USA co najmniej już od lat 20. (Elliott, Feldberg i Lehnert 2013). Kryzys finansowy ostatnich lat spowodował wzrost zainteresowania polityką makroostrożnościową (zob. Tressel i Verdier 2014), do czego przyczyniły się, między innymi, prace prowadzone w Banku Rozliczeń Międzynarodowych (*Bank for International Settlements* – zob. Borio 2003, 2011) oraz raport przygotowany na potrzeby Unii Europejskiej przez Grupę Wysokiego Szczebla (*High-level Group*) pod przewodnictwem byłego prezesa Europejskiego Banku Odbudowy i Rozwoju Jacquesa de Larosière'a (de Larosière 2009)[16]. W polskiej literaturze o potrzebie szerszego wykorzystania polityki makroostrożnościowej pisali m.in. Szpunar (2012) i Tymoczko (2012).

Regulacje i politykę makroostrożnościową definiuje się najczęściej jako całościowe (holistyczne) działanie władzy gospodarczej, służące identyfikacji i ograniczaniu ryzyka systemowego oraz zapewniające dzięki temu stabilność sektora finansowego jako całości (Crockett 2000; Brunnermeier i in. 2009; Billio i in. 2010;, Clement 2010; IMF 2011a). Kontrolując ryzyko systemowe, regulacje makroostrożnościowe ograniczają realne koszty gospodarcze związane z zaburzeniami, jakie mogą powstawać w procesie dostarczania ważnych produktów finansowych dla gospodarki. Do kosztów tych należy m.in. spadek produkcji wywołany ograniczeniem akcji kredytowej, spadek zatrudnienia i wzrost bezrobocia, bankructwo

[16] Raport de Larosière'a powstał z inicjatywy Rady Unii Europejskiej bezpośrednio po wybuchu kryzysu finansowego w Europie w 2008 roku.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

firm i banków oraz wzrost wartości długu publicznego (Elliott, Feldberg i Lehnert 2013; Lim i in. 2011)[17].

Celem polityki makroostrożnościowej jest ograniczanie ryzyka systemowego w sektorze bankowym i zapobieganie w ten sposób powstawaniu szkodliwych zaburzeń na rynkach finansowych. O ile regulacje mikroostrożnościowe mają służyć przede wszystkim ochronie interesów konsumentów usług finansowych (depozytariuszy), o tyle celem regulacji makroostrożnościowych jest ograniczanie ryzyka systemowego i ochrona stabilności całego sektora finansowego (Cecchetti 1999). Główną funkcją polityki makroostrożnościowej powinna więc pozostawać funkcja prewencyjna, a dopiero w dalszej kolejności funkcja korekcyjna, czyli postępowanie naprawcze. Chodzi bowiem o to, aby przede wszystkim nie dopuszczać do powstawania zagrożeń. Funkcja prewencyjna ma tu więc kluczowe znaczenie, ponieważ w sytuacji wystąpienia kryzysu nawet najbardziej energiczne działania nie są w stanie zniwelować wszystkich jego skutków.

Ponieważ ryzyko systemowe może być spowodowane zarówno pojawieniem się słabości bilansowych wspólnych dla całego sektora, jak i wynikać z zagrożenia upadkiem instytucji ważnych systemowo, zakres polityki makroostrożnościowej musi obejmować cały sektor bankowy (wszystkie instytucje dostarczające kredyty ze środków pożyczonych bez względu na ich rozmiary). W przypadku ryzyka wynikającego ze wspólnych słabości bilansowych, które narastają cyklicznie wraz z przyspieszaniem wzrostu gospodarczego, polityka makroostrożnościowa powinna być zmienna w czasie (tj. mieć charakter antycykliczny). Z kolei wobec ryzyka upadku instytucji ważnych systemowo polityka makroostrożnościowa powinna pozostawać niezależna od przebiegu cyklu i mieć charakter trwały, strukturalny.

[17] Międzynarodowy Fundusz Walutowy, we współpracy z Bankiem Rozliczeń Międzynarodowych i Radą Stabilności Finansowej (Lim i in. 2011) charakteryzuje politykę makroostrożnościową za pomocą trzech głównych pojęć, takich jak:

- cel: ograniczenie ryzyka wystąpienia zaburzeń na dużą skalę w dostarczaniu ważnych usług finansowych, co może powodować znaczne straty w sektorze realnym gospodarki ze względu na ścisłe powiązania sektora finansowego i sektora realnego;
- zakres analityczny: przedmiotem uwagi jest system finansowy jako całość, także w interakcji z sektorem realnym;
- instrumenty i system zarządzania: wykorzystanie instrumentów ostrożnościowych przeznaczonych przede wszystkim do ograniczania ryzyka systemowego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

7.4.2. Polityka makroostrożnościowa a polityka makroekonomiczna

Ustalając pożądany zakres polityki makroostrożnościową, należy pamiętać, że wypełnia ona lukę między polityką mikroostrożnościową z jednej strony a polityką makroekonomiczną – w tym głównie polityką pieniężną – z drugiej strony (Elliott i in. 2011, Viñals 2011). O ile pierwsza z nich dąży do zapewnienia stabilności poszczególnych instytucji finansowych, o tyle zadaniem drugiej jest zapewnienie równowagi makroekonomicznej, a zwłaszcza stabilnego poziomu cen w całej gospodarce. Oczywiście, nie da się w pełni precyzyjnie wyznaczyć zakresu oddziaływania trzech wymienionych rodzajów polityki, ponieważ obszary ich funkcjonowania częściowo się pokrywają. W pewnej mierze można założyć, że im bardziej wymagające stają się regulacje mikroostrożnościowe (mowa tu na przykład o buforach płynnościowych czy wymogach kapitałowych), tym mniejszy może być wymagany zakres polityki makroostrożnościowej, i odwrotnie. Podobnie też im bardziej restrykcyjna będzie polityka pieniężna (na przykład pod względem poziomu stóp procentowych), tym mniej wymagający może okazać się niezbędny nadzór makroostrożnościowy. Niemniej skala wzajemnej substytucji tych rodzajów polityki jest ograniczona (zob. rysunek 7.1). Doświadczenia wyniesione z ostatniego kryzysu finansowego pokazały, że ani polityka mikroostrożnościowa, ani polityka pieniężna nie są w stanie samodzielnie i w każdych warunkach zabezpieczać stabilności całego systemu bankowego.

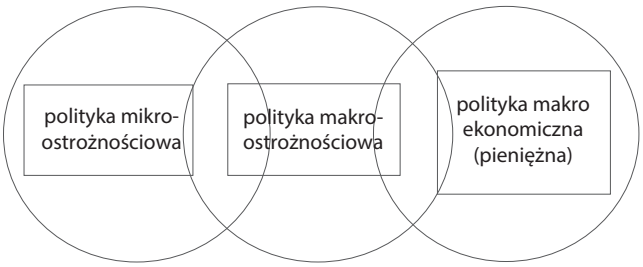
Szczególnie istotne jest rozdzielenie zakresu odpowiedzialności polityki mikroostrożnościowej i polityki makroostrożnościowej. Nie brakuje bowiem opinii, że prawidłowo funkcjonująca polityka nadzoru „mikro" zapobiega w większości przypadków narastaniu ryzyka systemowego, podczas gdy polityka „makro" sama w sobie niesie dodatkowe zagrożenia (zob. np. Calomiris 2013; Aiyar, Calomiris i Wielądek 2012). Pogląd ten pozostaje jednak kwestią dyskusyjną z następujących powodów.

Dotychczasowy system regulacji mikroostrożnościowych opierał się przede wszystkim na roli kapitału własnego banku jako bufora absorbującego możliwe straty. Pożądaną wielkość tego bufora określano na podstawie historycznych danych odnoszących się do poniesionych strat w stosunku do całości posiadanych przez bank aktywów. System ten sprawdzał się względnie dobrze w czasach „normalnej" koniunktury i w ramach tradycyjnej bankowości. Jednocześnie wykazywał jednak dwie główne słabości, które ze szczególną wyrazistością wyszły na jaw w czasie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ostatniego kryzysu. Po pierwsze, sama zdolność do absorbowania strat w historycznie określonej wysokości nie ogranicza bezpośrednio nadmiernego wzrostu aktywów, spowodowanego procyklicznością działalności banków. Innymi słowy, okazało się, że system buforów kapitałowych nie zapobiega narastaniu ryzyka. Odnosi się to zwłaszcza do nadmiernego wzrostu kredytów w czasach dobrej koniunktury, kiedy przypisane aktywom wagi ryzyka pozostają niskie, a wzrost wartości aktywów nie wymaga zwiększenia kapitału własnego. Po drugie, system ten ignoruje niekorzystne zmiany po pasywnej stronie bilansu banków, zwłaszcza gdy chodzi o nadmierne finansowanie się kapitałem hurtowym, krótkoterminowym (nadmierne lewarowanie). Taki sposób finansowania stwarza ryzyko gwałtownego zatrzymania napływu kapitału lub jego odpływu w sytuacji, gdy koniunktura ulega pogorszeniu. Niedoskonałości te decydują więc o potrzebie uzupełnienia systemu regulacji „mikro" regulacjami „makro".

Rysunek 7.1. Częściowo nakładające się obszary funkcjonowania polityki mikroostrożnościowej, makroostrożnościowej i makroekonomicznej



Źródło: opracowanie własne.

Częściowa substytucyjność polityki mikroostrożnościowej, polityki makroostrożnościowej i polityki makroekonomicznej sprawia, że optymalny zakres regulacji i polityki makroostrożnościowej pozostaje nadal przedmiotem dyskusji. Większość autorów skłania się do przekonania, że powinny być one kształtowane w relacji do ryzyka powstającego wyłącznie w sektorze finansowym (Viñals 2011), wskazując jednocześnie na niebezpieczeństwo nakładania się kompetencji i ewentualną niespójność stosowanych instrumentów przy szerszym określeniu mandatu. Lecz pojawiają się też zwolennicy szerszego zakresu odpowiedzialności polityki makroostrożnościowej. W tym kontekście Szpunar (2012) argumentuje na rzecz uwzględnienia także ryzyka powstającego w sektorze niefinansowym, na przykład w postaci nadmiernego zadłużania się przedsiębiorstw w walutach obcych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wydaje się jednak, że kwestia ta jest już przedmiotem zainteresowania polityki makroostrożnościowej z perspektywy dostawców kredytu, a więc w ramach nadzoru nad sektorem bankowym.

7.5. Instrumenty polityki makroostrożnościowej

W nawiązaniu do czterech wymienionych wyżej rodzajów ryzyka systemowego również instrumenty polityki makroostrożnościowej można podzielić odpowiednio na cztery kategorie: instrumenty regulujące rozmiary i tempo wzrostu kredytu, instrumenty regulujące płynność środków finansowych, instrumenty regulujące wielkość kapitału własnego i/lub stopnia lewarowania oraz instrumenty regulujące transgraniczne przepływy kapitału.

7.5.1. Instrumenty regulujące rozmiary i tempo wzrostu kredytu

Dlaczego należy kontrolować tempo emisji kredytu i zapobiegać nadmiernej akcji kredytowej? Szybki wzrost kredytu w gospodarce może powodować destabilizację makroekonomiczną i podnosić zagregowany popyt ponad potencjalny poziom PKB, przy jednoczesnym finansowaniu nadmiernej konsumpcji oraz importu, a także tworzeniu napięć inflacyjnych. Najbardziej niebezpieczne są spekulacyjne bańki cen aktywów finansowane kredytem krótkoterminowym, ponieważ pęknięcie takiej bańki, czyli gwałtowny spadek cen aktywów, uniemożliwia spłatę kredytu i wywołuje często niewypłacalność kredytobiorców i straty banków. Inwestorzy zmuszeni są ograniczać pozostałe wydatki w celu zdobycia środków na spłatę kredytów (delewarowanie) i wyprzedają posiadane aktywa. Gdyby dotyczyło to tylko indywidualnych przedsiębiorców lub pojedynczych banków, negatywne skutki w przypadku pozostałych podmiotów gospodarczych byłyby znikome. Ponieważ postępują jednak tak wszyscy inwestorzy lub zdecydowana ich większość, dochodzi w rezultacie do kryzysu bankowego, gwałtownego wstrzymania kredytów i recesji w gospodarce.

Można zapobiec tym nieprawidłowościom poprzez nałożenie ograniczeń na wzrost kredytów. W tym celu stosuje się szereg różnorodnych instrumentów. Najbardziej popularne są limity na wskaźniki LTV, DTI, bezpośrednie limity kredytu lub wzrostu kredytu, limity na kredyty w walutach obcych (*fx*), a także marginesy na lewarowane zakupy akcji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Limit na wskaźnik LTV (stosunek wielkości kredytu do wielkości zabezpieczenia – *loan-to-value*) ustala maksymalną dopuszczalną relację wartości kredytu do wartości jego zabezpieczenia (*collateral*). W ten sposób limit LTV, stosowany indywidualnie na potrzeby pojedynczego banku (w skali mikro), określa wymagany wkład własny kredytobiorcy (*down payment*) i ogranicza wielkość kredytu dostępnego pod zastaw danego zabezpieczenia. Stosowany zaś w skali sektora (na szczeblu makro) niski limit LTV hamuje tempo wzrostu kredytu w całej gospodarce[18]. Warto jednak zauważyć, że nawet stały limit LTV może oddziaływać antycyklicznie na wzrost kredytów, ponieważ zwiększenie wartości zabezpieczenia (wycenianego według zasady *mark-to-market*) w okresie hossy wymusza też proporcjonalny wzrost wkładu własnego, co może okazac się niewykonalne (IMF 2011a, s. 25–26; Crowe i in. 2011, s. 19–21), chociaż istnieją też przypadki nie potwierdzające tej zależności[19]. Z kolei gdy limit LTV jest dodatkowo podnoszony w okresie bessy i obniżany w okresie hossy, staje się bardzo silnym i skutecznym instrumentem antycyklicznego regulowania tempa wzrostu akcji kredytowej. Aby dodatkowo wzmocnić oddziaływanie limitu LTV można go uzupełnić limitem na wskaźnik DTI oraz wymogiem ubezpieczenia kredytu u zewnętrznego ubezpieczyciela.

Limit na wskaźnik DTI (stosunek długu do dochodu – *debt-to-income* – lub bardziej precyzyjnie stosunek obciążenia z tytułu obsługi kredytu do dochodu w danym okresie) uzależnia wielkość kredytu od określonej miary dochodu kredytobiorcy, stanowiącej jego zdolność kredytową (płatniczą). W ten sposób, gdy stosowany jest indywidualnie (na szczeblu mikro) niższy limit DTI, wzrasta jakość aktywów banku, ponieważ maleje jednocześnie ryzyko wystąpienia trudności w obsłudze

[18] Istniejąca literatura zasadniczo potwierdza tezę o skuteczności limitu LTV w hamowaniu nadmiernego wzrostu kredytów. Na przykład Crowe i in, (2011) wykazują, że obniżenie LTV o 10 punktów procentowych wywołuje spadek ceny domów o 8–13%. Limity na LTV mają też wpływ na spowalnianie mechanizmu akceleratora finansowego dzięki ograniczeniu przenoszenia się skutków wzrostu dochodów na wzrost ceny domów (zob. IMF 2011a; Almeida, Campello i Liu 2006).

[19] Załóżmy, że wartość zabezpieczenia wynosi 100 PLN, a limit na LTV znajduje się na poziomie 80%. Pozwala to otrzymać kredyt do wysokości 80 PLN przy wkładzie własnym o wartości 20 PLN. Wzrost ceny zabezpieczenia o 50% (do poziomu 150 PLN) pozwoli wprawdzie wziąć większy kredyt (aż do poziomu 120 PLN), lecz będzie wymagało to też wzrostu udziału własnego do 30 PLN, a więc także o 50%. Ogólnie rzecz biorąc, zmniejsza się wówczas presja na wzrost kredytów, ponieważ przyrost dochodów w okresie hossy jest znacznie wolniejszy niż wzrost ceny aktywów. Chociaż większość badań empirycznych potwierdza, że stały limit na LTV hamował wzrost kredytów hipotecznych (zob. np. IMF 2011a, s. 25–26), pojawiają się też opinie odmienne. Gelain i in. (2013) wykazują, że w USA w latach 2000–2007 doszło do silnego wzrostu zadłużenia gospodarstw domowych z tytułu kredytów hipotecznych (z około 60% do około 100% dochodu do dyspozycji) mimo tego, że średni LTV utrzymywał się na względnie stabilnym poziomie (około 60%).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

zadłużenia przy danym dochodzie klientów banku (zmniejsza się ryzyko kredytowe). Banki stosują różne miary dochodu w celu wyliczenia DTI. Najczęściej bierze się pod uwagę dochód netto (po opodatkowaniu), lecz można też potrącać z dochodu także inne sztywne wydatki, takie jak podatki od nieruchomości, składki ubezpieczeniowe lub spłaty innych kredytów (np. w ramach kart kredytowych). Im bardziej rygorystyczne są zasady obliczania dochodu na potrzeby ustalenia wskaźnika DTI, tym trudniejsze stają się warunki uzyskania kredytu. Należy jednak zauważyć, że stosowany na skalę całego sektora określony limit DTI ma łagodny procykliczny wpływ, ponieważ zdolność kredytowa zmienia się wraz z poziomem dochodów w przebiegu cyklu koniunkturalnego. W szczególności wzrost dochodów w okresie hossy umożliwia zaciąganie większych kredytów, choć w wolniejszym tempie niż wzrost cen zabezpieczeń. Tę niedogodność można jednak wyeliminować poprzez antycykliczne zmiany wymaganego wskaźnika DTI (podnoszenie limitu w czasie bessy i obniżanie w czasie hossy), co zmniejsza ryzyko systemowe cyklicznych boomów kredytowych. Dodatkowo, gdy antycykliczny DTI jest stosowany łącznie z limitem LTV, stanowią one razem bardzo skuteczny instrument ograniczania akcji kredytowej (Gelain i in. 2013).

Pułapy (limity) kredytowe mogą dotyczyć albo wielkości akcji kredytowej, albo tempa wzrostu kredytu. Odnoszą się do kredytów udzielanych przez pojedynczy bank (limity mikro), do kredytów przeznaczonych dla określonego sektora (limity sektorowe) lub do całości akcji kredytowej dla gospodarki (limity makro). Pułap nałożony na cały agregat kredytowy (lub na aktywa krajowe netto) spowalnia akcję kredytową w całej gospodarce i przynosi podobne skutki jak zaostrzanie parametrów polityki pieniężnej[20]. Ustalony pułap kredytowy ma charakter antycykliczny, ponieważ nie zmienia się pod wpływem wzrostu cen aktywów (zabezpieczeń) w okresie hossy. W ten sposób pułapy te oddziałują na wymiar czasowy ryzyka systemowego. Z kolei limity sektorowe (związane np. z kredytami na zakup nieruchomości) ograniczają wspólną ekspozycję banków wobec danego sektora, hamują nadmierny wzrost cen określonych aktywów i powstrzymują tym samym narastanie bańki spekulacyjnej w danym sektorze. W ten sposób pułapy sektorowe kształtują sektorowy wymiar ryzyka systemowego.

Limity na poziom lub tempo wzrostu kredytów w walutach obcych – kredyty *fx* – są ważnym instrumentem ograniczającym ryzyko kursowe oraz ryzyko

[20] Pułapy kredytowe jako narzędzia administracyjne są skuteczniejsze od narzędzi parametrycznych, takich jak stopy procentowe banku centralnego, lecz mogą prowadzić do silniejszych zniekształceń bodźców i pociągać za sobą większe straty w zakresie efektywności.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kredytowe. Gdy klient jest „niezabezpieczony" (*unhedged*) przed ryzykiem kursowym (posiada dochody w walucie krajowej i nie stosuje zabezpieczeń), kredyt *fx* naraża go na ryzyko kursowe, a to z kolei wystawia bank na niebezpieczeństwo związane z ryzykiem kredytowym (w sytuacji gdy spadek kursu waluty krajowej powoduje wzrost obciążeń z tytułu spłaty kredytu, którym klient nie jest w stanie podołać). Kiedy zaś wiele banków udziela kredytów *fx* niezabezpieczonym klientom, pojawia się ryzyko systemowe, ponieważ deprecjacja waluty krajowej zwiększa ryzyko kredytowe we wszystkich bankach. Nałożenie limitów na kredyty *fx* lub wręcz ich zakazanie ogranicza skutecznie występowanie tego rodzaju niebezpieczeństw[21].

Omawiając instrumenty makroostrożnościowe regulujące wielkość kredytu, warto na koniec wspomnieć o instrumentach ograniczających zaciąganie kredytu przeznaczonego na lewarowane zakupy instrumentów udziałowych. Gdy banki inwestują za pośrednictwem biur brokerskich w udziałowe instrumenty finansowe, takie jak akcje, może dochodzić w okresach hossy do dokonywania nadmiernych inwestycji, zagrażających stabilności banku. Ryzyko to staje się szczególnie niebezpieczne w warunkach spekulacyjnego wzrostu cen akcji oraz inwestycji finansowanych długiem. W celu obniżenia go można stosować instrumenty ograniczające lewarowane zakupy akcji w postaci wymaganej wysokości wpłaty własnej. Taka wymagana wysokość należności – zwana marginesem wpłaty własnej (*stock margin requirements* – SMR) – to minimalny poziom wpłaty własnej w stosunku do ceny zakupu akcji. Wyższy poziom SMR ogranicza skalę lewarowanego zakupu aktywów oraz spowalnia wzrost cen i kredytów, co oznacza, że antycykliczne zmiany marginesów sprzyjają łagodzeniu cyklicznych napięć[22].

7.5.2. Instrumenty regulujące płynność aktywów

Regulacje bankowe nakładają na banki obowiązek utrzymywania tzw. buforów płynnościowych mających zapewnić niezbędną płynność w warunkach kryzysu (stresu) płynnościowego, czyli w sytuacji, gdy pojawiają się trudności z pozyskaniem

[21] Europejska Rada Ryzyka Systemowego uznała wysoki poziom zadłużenia sektora niefinansowego z tytułu kredytów walutowych za potencjalne zagrożenie dla stabilności finansowej na poziomie całej Unii Europejskiej. W październiku 2011 roku ESRB opublikowała listę siedmiu zaleceń, których celem jest przeciwdziałanie wzrostowi portfela kredytów walutowych udzielanych kredytobiorcom nieposiadającym zabezpieczenia przed ryzykiem kursu walutowego (NBP 2011, s. 52–55).

[22] Np. jeśli SMR wynosi 50%, to maksymalny stopień lewaru (*leverage ratio*) jest równy 2:1, co oznacza, że przy cenie zakupu o wartości 200 PLN, wpłata własna musi stanowić minimum 100 PLN, a wówczas dług może osiągnąć maksimum na poziomie 100 PLN.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

środków potrzebnych na wypłatę depozytów bądź spłatę zapadających zobowiązań. Kiedy większość banków nie spełnia wymogów płynnościowych, powstaje ryzyko systemowe prowadzące do gwałtownego ograniczenia akcji kredytowej w sytuacji stresu płynnościowego. Jednocześnie, ponieważ bufory płynnościowe wiążą część aktywów, ograniczają one także nadmierną akcję kredytową, zwłaszcza finansowaną środkami krótkookresowymi, hurtowymi. Ponieważ kryzys płynności trwa jednak z reguły dłużej niż kilka dni, bufory płynnościowe pozwalają jedynie „kupić" czas, aby bank dotknięty kryzysem płynności mógł odzyskać wymaganą płynność w „normalnym" trybie (np. upłynniając aktywa) lub zwrócić się do banku centralnego o udzielenie pomocy. W poszczególnych krajach stosowane są różne miary wymaganej płynności[23], lecz obowiązuje ogólna zasada, zgodnie z którą w skład buforów muszą wchodzić aktywa o wysokim stopniu płynności, takie jak łatwo zbywalne aktywa finansowe (np. obligacje skarbowe), rezerwy w banku centralnym, waluty wymienialne czy gotówka.

W ostatnich latach podjęto próbę harmonizacji wymogów płynnościowych w ramach najnowszego zestawu rekomendacji Bazylejskiego Komitetu Nadzoru Bankowego (*Basel Committee on Banking Supervision*), określanego jako Basel III. Komitet bazylejski zaleca w szczególności stosowanie normy (miary) płynności krótkoterminowej LCR (*Liquidity Coverage Ratio*). Norma ta stanowi, że bank powinien dysponować wystarczającym zasobem nieobciążonych aktywów płynnych wysokiej jakości, które mogą być bez przeszkód i natychmiastowo wymienione na gotówkę na rynku w drodze transakcji bilateralnej, tak aby w scenariuszu warunków skrajnych płynności można było zaspokoić potrzeby płynnościowe banku przez kolejne 30 dni kalendarzowych (tzw. „okres przetrwania"). Norma ta weszła w życie w państwach członkowskich Unii Europejskiej w 2015 roku.

Komitet bazylejski zadbał także o utrzymanie wymaganej płynności w dłuższym okresie, rekomendując stosowanie określonych limitów na niedopasowanie terminów zapadalności aktywów i pasywów (*maturity mismatch*). W warunkach kryzysu niemożność spłaty zobowiązań w danym terminie ze względu na brak płynności spowodowany niedopasowaniem terminów zapadalności zmusza bank do awaryjnej wyprzedaży aktywów, co obniża ich ceny i powoduje straty kapitałowe wśród innych banków (*externality*).

Taki wymóg płynnościowy odnoszący się do dłuższego okresu wprowadza norma (miara) płynności długoterminowej NSFR (*Net Stable Funding Ratio*),

[23] W Polsce stosuje się operacyjne miary płynności M1, M2, M3, M4.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

która ustanawia minimalny akceptowalny poziom stabilnego finansowania oparty na charakterystyce płynnościowej aktywów i działalności banku w horyzoncie jednego roku. Norma NSFR, będąca normą komplementarną wobec normy LCR, ma wymuszać odejście od niewłaściwego finansowania w krótkim okresie w kierunku bardziej stabilnego, długoterminowego finansowania aktywów i działalności banku.

Ważnym i skutecznym instrumentem zapewniającym wymaganą płynność jest stopa rezerwy obowiązkowej. Wprawdzie rezerwa obowiązkowa stanowi zasadniczo instrument polityki pieniężnej, kształtujący tempo wzrostu podaży pieniądza w całej gospodarce, lecz może być także z powodzeniem stosowana w polityce makroostrożnościowej. Chodzi bowiem przede wszystkim o to, że rezerwa obowiązkowa jest utrzymywana w postaci środków płynnych, co może ułatwiać przezwyciężanie kryzysów płynności w bankach i ma w związku z tym wpływ na sektorowy wymiar ryzyka systemowego. Ponadto, stopa rezerwy obowiązkowej wpływa również bezpośrednio na wielkość akcji kredytowej, dzięki czemu może być wykorzystywana w celu ograniczania cyklicznych wahań wielkości kredytu, a to pozwala regulować czasowy wymiar ryzyka systemowego.

7.5.3. Instrumenty regulujące wielkość kapitału banku

Instrumenty te mają na celu utrzymanie kapitału własnego banku na poziomie zapewniającym możliwość absorbowania strat wynikających z utraty wartości aktywów. Przyjmują one różne formy. Najczęściej stosowana jest minimalna dopuszczalna wartość relacji kapitału własnego do aktywów ważonych ryzykiem (zob. ramka 7.2), czyli współczynnik wypłacalności CAR, który może być stały lub zmienny w czasie. Gdy współczynnik CAR jest stały (tak jak wymagają tego zasady sformułowane przez Bazylejski Komitet Nadzoru Bankowego – w ramach tzw. Basel II), funkcjonuje on w istocie procyklicznie, ponieważ w okresie dobrej koniunktury współczynniki ryzyka na aktywach są niskie, a to pozwala zwiększać akcję kredytową przy danym kapitale własnym. Również odwrotnie, w okresie złej koniunktury współczynniki ryzyka wzrastają, co zmusza banki do ograniczania skali udzielanych kredytów przy ustalonym kapitale. W ten sposób sztywne współczynniki kapitałowe przyczyniają się do pogłębienia zjawisk kryzysowych w gospodarce.

Dlatego też w ramach polityki makroostrożnościowej zaleca się stosowanie zmiennych współczynników CAR oraz zmiennych wag ryzyka przypisanych poszczególnym rodzajom aktywów. Podnoszenie tych współczynników w okresie hossy hamuje wzrost akcji kredytowej w „dobrych czasach", podczas gdy obniżanie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ich w okresie bessy zmniejsza presję na pozyskiwanie kapitału przez banki w celu tworzenia rezerw na pokrycie strat w „złych czasach". Zmienne współczynniki kapitałowe działają więc antycyklicznie i w ten sposób adresują wymiar czasowy ryzyka systemowego. W podobny sposób funkcjonuje rekomendowany w ramach Basel III antycykliczny bufor kapitałowy, budowany w okresie hossy i zużywany w okresie bessy (tzw. *rainy fund*).

Ramka 7.2. Współczynnik wypłacalności

Jedną z podstawowych regulacji ostrożnościowych stanowi wymóg posiadania przez banki pewnego minimum własnego kapitału, pozwalającego na absorbowanie ewentualnych strat. Minimum kapitałowe jest ustalone jako określony przez krajowego nadzorcę i wyrażony w procentach stosunek kapitału własnego do aktywów ważonych ryzykiem, zgodnie z opisanym na początku rozdziału wzorem (7.1):

$$CAR = \frac{K_P + K_U}{\sum_{i=1}^{n} \left(w_i \cdot A_i \right)}, \tag{7.1}$$

gdzie K_P i K_U oznaczają odpowiednio kapitał podstawowy i kapitał uzupełniający banku. Kapitał podstawowy (*Tier 1 capital*) obejmuje kapitał akcyjny i sumę zakumulowanych zysków zatrzymanych w banku (*disclosed reserves*), podczas gdy na kapitał uzupełniający (*Tier 2 capital*) składają się nierozdysponowane zyski (*undisclosed reserves*), rezerwa rewaluacyjna, rezerwa ogólna oraz – niekiedy – dług podporządkowany (*subordinated debt*). Szczegółowa klasyfikacja składników kapitału w różnych krajach może być różna. Wagi ryzyka przypisane poszczególnym rodzajom aktywów odzwierciedlają różny poziom ryzyka poniesienia straty na rozmaitych formach aktywów. W Polsce stosuje się wagi ryzyka rekomendowane przez komitet bazylejski oraz dyrektywy i rozporządzenia bankowe UE (tabela 7.1).

Współczynnik wypłacalności opisany formułą (7.1) nie powinien być niższy niż 8%, co oznacza, że kapitał banku – czyli fundusze własne netto (podstawowe oraz uzupełniające) – musi wynosić minimum 8% aktywów banku i pozycji pozabilansowych ważonych ryzykiem, przy czym kapitał podstawowy – akcyjny – powinien sięgać minimum 4% aktywów. Konstrukcja współczynnika CAR jest pomyślana tak, aby bank posiadający bardziej ryzykowne aktywa musiał zgromadzić większy kapitał niż bank dysponujący bezpieczniejszymi zasobami.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 7.1. Wagi ryzyka przypisane kategoriom aktywów bankowych

Waga ryzyka (dane w %)	Kategorie aktywów
0%	aktywa niskiego ryzyka: gotówka, skarbowe papiery wartościowe i należności od narodowych rządów i banków centralnych, rządów i banków centralnych państw Unii Europejskiej i państw należących do OECD
20%	aktywa niskiego i średniego ryzyka: należności od banków krajowych i banków z siedzibą w państwach należących do OECD
50%	aktywa średniego ryzyka: przede wszystkim kredyty zabezpieczone hipotecznie
100%	aktywa wysokiego ryzyka: pozostałe kredyty, większość operacji pozabilansowych, inwestycje banku na rynku kapitałowym itp.

Źródło: opracowanie własne na podstawie danych KNF.

Dynamiczne (zmienne w czasie) wymogi (bufory) kapitałowe spełniają podwójną funkcję stabilizacyjną. Po pierwsze, ponieważ w okresie hossy banki zmuszone są do intensywniejszego zwiększania kapitału własnego (*equity*), którego pozyskanie jest zawsze kosztowniejsze niż zaciągnięcie długu (ze względu na to, że kapitał własny pozostaje obciążony większym ryzykiem, a z kolei odsetki od długu są z reguły odliczane od podatku – *tax-deductible*), to rośnie również koszt kredytów, co w konsekwencji hamuje ich wzrost. Po drugie, bufory kapitałowe podnoszą odporność systemu na szoki, ponieważ zwiększają zdolność ich absorbowania, a to obniża ryzyko wystąpienia systemowego niedoboru kredytu (*credit crunch*). W obu przypadkach dynamiczne bufory kapitałowe obniżają szkodliwą cykliczność zmian w zakresie wag ryzyka opartych na zewnętrznych ratingach, które są zawsze procykliczne (w ramach Basel II).

Ważnym instrumentem polityki makroostrożnościowej w przypadku omawianej kategorii są wielkości odpisów na utratę wartości aktywów (*provisions*). Tradycyjnie wielkości odpisów na zagrożone należności ustala się na podstawie historycznych danych dotyczących strat na aktywach danego banku. Gdy wielkość tych odpisów nie zmienia się w czasie, oddziałują one procyklicznie na akcję kredytową, ponieważ łagodzą wymogi kapitałowe w okresie hossy i zaostrzają je w okresie bessy. Warto zatem stosować zmienne wielkości odpisów w zależności od stanu koniunktury – podnosić je w czasie hossy w celu ograniczenia akcji kredytowej i obniżać w czasie bessy, aby ułatwić udzielanie kredytu – i w ten sposób ograniczać procykliczność. Najlepiej jest w tym celu wykorzystać stałą formułę, lecz można też decydować *ad hoc*.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Do instrumentów makroostrożnościowych mających wpływ na wielkość kapitału własnego należą także regulacje dotyczące ograniczeń w podziale zysku. Tradycyjnie celem tych regulacji jest wzmocnienie bazy kapitałowej banku (na poziomie mikro) dzięki zyskom zatrzymanym (*retained profits*), zwłaszcza w sytuacji, gdy bank nie spełnia wymogów adekwatności kapitałowej. Nałożenie silniejszych ograniczeń na podział zysku w czasie bessy działa antycyklicznie, ponieważ zmusza bank do powiększenia kapitału własnego i w ten sposób zwiększa możliwości udzielania kredytu. W podobny sposób ma funkcjonować tzw. dodatkowy bufor kapitałowy (*conservation capital buffer*), rekomendowany w ramach Basel III.

7.5.4. Instrumenty regulujące transgraniczne przepływy kapitału

Poszukując środków do finansowania własnej działalności, banki sięgają często do źródeł zagranicznych. Zaciągają pożyczki w bankach matkach, emitują obligacje na rynkach zagranicznych bądź finansują się na zagranicznych rynkach hurtowych. Przyczyną takiego stanu rzeczy jest na ogół niższy koszt pozyskiwania środków zagranicznych, czasem także niedostatek krajowych depozytów. Lecz im bardziej wielkość pasywów w walutach obcych przekracza wielkość aktywów w walutach obcych, tym większe staje się dodatkowe ryzyko dla banku, w tym przede wszystkim ryzyko kursowe wynikające ze zmian kursu walut oraz ryzyko płynności związane z możliwością nagłego wstrzymania napływu bądź możliwością nagłego odpływu kapitału. Osłabienie kursu waluty krajowej przekłada się na wzrost obciążeń z tytułu obsługi zadłużenia zagranicznego, co może powodować straty i zagrażać stabilności banku. Z kolei ryzyko nagłego odpływu kapitału rośnie, gdy banki krajowe finansują się krótkoterminowymi pożyczkami na zagranicznych rynkach pieniężnych. Gdy równocześnie wiele instytucji bankowych finansuje się na rynkach zagranicznych, powstaje ryzyko systemowe, ponieważ spadek kursu lub nagły odpływ kapitału (a najczęściej jedno i drugie jednocześnie) zagraża stabilności całego krajowego sektora bankowego. Badania empiryczne pokazują, że instrumenty stosowane w polityce zarządzania transgranicznymi przepływami kapitałów mogą skutecznie ograniczać ryzyko niestabilności makroekonomicznej i finansowej (IMF 2011e).

W celu ograniczenia systemowego ryzyka wynikającego z niezbilansowania aktywów i pasywów w walutach obcych krajowe władze nadzorcze mogą stosować limity na otwarte pozycje walutowe (*net open currency positions* – NOP). Ma to szczególne znaczenie w sytuacji, gdy pod wpływem szoku zewnętrznego wiele

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

banków jednocześnie sprzedaje lub kupuje waluty obce, co wywołuje gwałtowne wahania kursu i zmiany w zakresie wyceny aktywów i pasywów. Decyzje poszczególnych banków mają w ten sposób wpływ na bilanse innych banków, a w konsekwencji także na kondycję całego sektora (*negative externality*). Limity NOP obniżają więc ryzyko systemowe wynikające z możliwości upadku instytucji bardziej zadłużonych w walutach obcych.

Zamiast limitów NOP (lub równolegle z nimi) można stosować opłaty (*levies*) od otwartych pozycji walutowych w wysokości zależnej proporcjonalnie od ich wielkości. Rozwiązanie takie stosowały m.in. kraje Ameryki Łacińskiej i Korea Południowa[24]. Tradycyjnym instrumentem ograniczającym nadmierne zadłużanie się za granicą są nieoprocentowane odpisy rezerwowe (*unremunerated reserve requirement* – URR) – wymóg odprowadzania na specjalny rachunek rezerwy od pożyczek zaciągniętych za granicą. Takie wymogi odprowadzania nieoprocentowanych rezerw od środków zagranicznych (najczęściej w wysokości 10–20%) stosowane były m.in. w Chile po 1990 roku i w Kolumbii od 2007 roku. Instrument ten funkcjonuje podobnie jak podatek od napływu kapitału, choć nie ma charakteru fiskalnego, ponieważ rezerwa podlega zwrotowi w momencie zamknięcia pozycji walutowej. Trzeba jednak zaznaczyć, że skuteczność URR w warunkach bardzo niskich stóp procentowych jest ograniczona, ponieważ maleje wówczas koszt alternatywny utrzymywania rezerw.

Tradycyjnie stosowanym instrumentem ograniczania negatywnych skutków nagłego odpływu kapitału jest administracyjna kontrola transgranicznych przepływów kapitałowych (*capital controls* – CC, *capital flow management*). Najczęściej instrument ten dyskryminuje przepływy według miejsca rezydencji inwestorów – inwestorzy zagraniczni lub krajowi nie mogą dokonywać określonych transakcji transgranicznych. Kontrole przepływów kapitału są stosowane w różnych celach przez niektóre kraje o gospodarkach wschodzących lub rozwijające się, przy czym najczęściej chodzi o ograniczanie odpływu kapitału[25].

Choć administracyjna kontrola przepływów kapitału jest zasadniczo instrumentem polityki makroekonomicznej służącym utrzymaniu równowagi w bilansie płatniczym, może być także wykorzystywana jako instrument polityki makroostrożnościowej.

[24] Np. Korea Południowa wprowadziła w sierpniu 2011 roku obowiązkowe opłaty od zobowiązań zagranicznych (*non-core liabilities*): banki płacą 20 bps rocznie od zobowiązań w walutach obcych z terminem zapadalności poniżej 12 miesięcy na specjalny rachunek walutowy.

[25] Tymczasem np. Brazylia w 2010 roku nałożyła bezpośrednie ograniczenia w zakresie napływu kapitału, chcąc osłabić presję na aprecjację waluty krajowej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Niektóre badania wskazują, że kontrola przepływów kapitału ma pozytywny wpływ na ograniczenie boomów kredytowych i pożyczek w walutach obcych (Ostry i in. 2011) oraz że zwiększa stopień autonomii polityki pieniężnej i ogranicza wahania kursu (Magut, Reinhart i Rogoff 2010).

Kontrole kapitałowe mają jednak także istotne wady. Po pierwsze, raz wprowadzone, okazują się trudne do uchylenia[26]. Po drugie, istnienie kontroli kapitałowych podważa zaufanie inwestorów, powodując m.in. wstrzymanie napływu bezpośrednich inwestycji, podczas gdy obawa przed odpływem kapitału sprzyja ich utrzymywaniu. Po trzecie, długotrwale utrzymywanie kontroli obniża ich skuteczność, ponieważ uczestnicy rynku znajdują liczne sposoby na ich omijanie. W efekcie rosną koszty transakcyjne związane z operacjami międzynarodowymi[27].

Instrumentem, który potencjalnie może ograniczyć skalę transgranicznych przepływów kapitału, jest podatek od transakcji finansowych (*financial transaction tax* – FTT). Ostatnie lata przyniosły ponowny wzrost zainteresowania tym rozwiązaniem. W 2011 roku Komisja Europejska wystąpiła z propozycją opodatkowania stawką 0,1% transakcji kupna i sprzedaży akcji i obligacji oraz stawką 0,01% transakcji na instrumentach pochodnych[28]. Temat ten budzi jednak duże kontrowersje pod względem zarówno teoretycznym, jak i praktycznym[29]. Zwolennicy wprowadzenia podatku utrzymują, że FTT znajduje uzasadnienie ekonomiczne i społeczne, ponieważ sektor finansowy jest relatywnie mniej obciążony podatkami (usługi finansowe są zwolnione z VAT), a, ponosząc znaczną część odpowiedzialności za

[26] Na przykład, gdy Islandia zastosowała kontrolę przepływu kapitału w 2008 roku w celu zatrzymania odpływu kapitału po kryzysie bankowym, zakładano, że będzie ona obowiązywać przez kilka tygodni. Tymczasem jeszcze w 2015 roku ograniczenia te pozostawały w mocy.

[27] Mimo tych negatywnych doświadczeń Cypr wprowadził kontrolę przepływu kapitału w kwietniu 2013 roku jako pierwszy spośród krajów członkowskich UE. Wprawdzie kontrole te miały być „ograniczone i tymczasowe", lecz zostały uchylone dopiero w lutym 2014 roku, i to jedynie częściowo.

[28] Jedenaście państw członkowskich UE zgodziło się wprowadzić podatek od transakcji finansowych zgodnie z propozycją Komisji w ramach mechanizmu tzw. wzmocnionej współpracy (*enhanced cooperation*). Projekt stosownej dyrektywy został przedstawiony już w 2013 roku (Com/2013/76) ale do 2016 roku nie wszedł w życie.

[29] Dyskusja na ten temat trwa co najmniej od 1978 roku, kiedy amerykański ekonomista James Tobin zaproponował nałożenie niewielkiego podatku na transakcje wymiany walut, argumentując, że, „sypiąc piasek w tryby mechanizmu rynkowego", ograniczy się skalę transakcji czysto spekulacyjnych i przyczyni do ustabilizowania rynków. Pomysł Tobina nie doczekał się realizacji przede wszystkim ze względów praktycznych – uznano, że mógłby być skuteczny tylko wówczas, gdyby wszystkie, a przynajmniej zdecydowana większość państw wprowadziła taki podatek, co uniemożliwiłoby „migrację" transakcji walutowych za granicę do krajów niestosujących tego podatku. Co więcej, podatek Tobina krytykowano również na gruncie teoretycznym: zauważano bowiem, że, zwiększając koszty transakcyjne, miałby szkodliwy wpływ na efektywność mechanizmu rynkowego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

kryzys finansowy, powinien w większym stopniu partycypować w kosztach usuwania jego skutków. Nałożenie podatku osłabi tendencję do spekulacji finansowej, a dochody z tego tytułu zasilą napięty budżet europejski. Zwolennicy tej propozycji podkreślają, że ogromna większość transakcji finansowych dokonywanych na światowych rynkach nie tylko w żaden sposób nie służy realnej gospodarce, lecz staje się także źródłem niestabilności i szoków. Z kolei przeciwnicy takiego rozwiązania utrzymują, że skuteczne wprowadzenie FTT jest w praktyce niewykonalne ze względu na możliwość przenoszenia transakcji finansowych za granicę, i ostrzegają, że może to doprowadzić nie tyle do spadku, ile, wręcz przeciwnie, do wzrostu skali zaburzeń na rynkach finansowych[30].

7.5.5. Instrumenty oddziałujące na stronę aktywną i na stronę pasywną bilansu banku

Jak wskazano wyżej, choć ryzyko systemowe powstaje w różnych obszarach działalności bankowej, ostatecznie wszystkie kategorie ryzyka systemowego dotyczą albo aktywów, albo pasywów bilansu bankowego. Podobnie jest z instrumentami polityki makroostrożnościowej. Takie narzędzia jak LTV czy DTI, a także wskaźniki adekwatności kapitałowej (CAR) oddziałują na stronę aktywną bilansów banków, hamując wzrost udzielanych kredytów i dostosowując go do poziomu kapitału własnego. Operacyjne miary płynności mają także wpływ na stronę aktywów, narzucając wymogi ograniczające ryzyko płynności.

Ostatni kryzys finansowy pokazał jednak, że równie niebezpieczne mogą okazać się niedoskonałości narastające po stronie pasywów bilansu bankowego. Dotyczy to przede wszystkim nadmiernego finansowania akcji kredytowej krótkoterminowymi pożyczkami zaciąganymi na hurtowym rynku pieniężnym oraz pożyczkami zagranicznymi. W takiej sytuacji wskazane jest stosowanie instrumentów oddziałujących także na stronę pasywną, takich jak limity na dopasowanie struktury zapadalności aktywów i pasywów czy ograniczenia w finansowaniu zagranicznym.

[30] Wydaje się, że przy spełnieniu pewnych określonych warunków FTT może odegrać *per saldo* pozytywną rolę. Po pierwsze, stawka podatku musi być dostatecznie niska, aby, ograniczając skalę transakcji czysto spekulacyjnych o wysokiej częstotliwości, nie powodowała jednocześnie odpływu za granicę transakcji społecznie użytecznych. Po drugie, celem FTT powinno być jedynie ograniczenie skali transakcji spekulacyjnych, a nie maksymalizowanie dochodów budżetu czy redystrybucja dochodów – chodzi więc o podatek typu korekcyjnego, a nie fiskalnego. Po trzecie, FTT powinien obejmować możliwie najszerszą grupę podatników i szeroki zakres instrumentów finansowych, aby zapewnić dostatecznie szeroką bazę podatkową i uniknąć arbitrażu między instrumentami stanowiącymi bliskie substytuty (Rosati 2012).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 7.2. Związki między odmianami ryzyka systemowego i rodzajami instrumentów makroostrożnościowych (MP)

Instrumenty / Ryzyko		Rodzaje instrumentów MP			
		regulujące rozmiar i tempo wzrostu kredytu	regulujące płynność środków finansowych	regulujące wielkość kapitału własnego	regulujące transgraniczne przepływy kapitału
Odmiany ryzyka systemowego	ryzyko kredytowe	LTV, DTI, limity kredytowe, limity kredytów *fx*, marginesy na lewarowane inwestycje	bufory płynnościowe, Stopa rezerwy obowiązkowej	CAR, bufory kapitałowe	URR, CC
	ryzyko płynności		LCR, NSFR, bufory płynnościowe, operacyjne miary płynności, stopa rezerwy obowiązkowej		
	ryzyko lewarowania	marginesy na lewarowane inwestycje		CAR, bufory kapitałowe, wagi ryzyka, odpisy na utratę wartości aktywów, ograniczenia na podział zysku	opłaty (*levies*) od otwartych pozycji walutowych, URR, CC, FTT
	ryzyko zagranicznych przepływów kapitału				limity NOP, opłaty (*levies*) od otwartych pozycji walutowych, URR, CC, FTT

Źródło: opracowanie własne.

Tabela 7.2 zawiera podsumowanie związków występujących między różnymi odmianami ryzyka systemowego i rozmaitymi rodzajami instrumentów polityki makroostrożnościowej. Przyporządkowano w niej schematycznie poszczególne rodzaje narzędzi poszczególnym odmianom ryzyka. Jak wspomniano, niektóre instrumenty mogą absorbować kilka rodzajów ryzyka systemowego jednocześnie (np. bufory płynnościowe, CAR czy stopa rezerwy obowiązkowej). Należy zaznaczyć, że przedstawione w tabeli przyporządkowanie ma charakter ogólny, orientacyjny. Wybór i zastosowanie określonych instrumentów polityki

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

makroostrożnościowej zależy bowiem zawsze od przebiegu danej sytuacji i konkretnych warunków gospodarczych.

7.6. Zastosowanie polityki makroostrożnościowej

7.6.1. Jeden czy kilka instrumentów jednocześnie?

Badania nad funkcjonowaniem i skutkami polityki makroostrożnościowej w krajach rozwiniętych pokazują, że zastosowanie kilku instrumentów naraz najczęściej zwiększa ich skuteczność i ogranicza możliwości omijania niektórych regulacji (IMF 2012). Na przykład limity LTV i DTI są często wykorzystywane łącznie, ponieważ uzupełniają się wzajemnie przy obniżaniu cykliczności kredytu udzielanego na podstawie zabezpieczeń, absorbując to samo ryzyko kredytowe – LTV pod względem majątkowym, a DTI od strony dochodu. Ogólnie rzecz biorąc, gdy kredytowe instrumenty makroostrożnościowe są stosowane w celu ograniczenia ryzyka związanego z nadmierną akcją kredytową, można także wykorzystać równolegle płynnościowe instrumenty makroostrożnościowe do ograniczaniu ryzyka finansowania hurtowego oraz kapitałowe instrumenty makroostrożnościowe do zwiększenia buforów. Badania przeprowadzone przez IMF (IMF 2010) potwierdzają, że stosowanie wielu instrumentów jednocześnie jest powszechną praktyką, w szczególności w odniesieniu do instrumentów ograniczających kredyt i regulujących wielkość kapitału.

Trzeba jednak pamiętać, że jednoczesne stosowanie kilku instrumentów naraz zwiększa wymogi regulacyjne wobec banków, co oznacza podniesienie kosztów nadzoru i kosztów dostosowania po stronie banków. Ponadto, w warunkach wykorzystania wielu instrumentów trudniejsza staje się ich prawidłowa kalibracja (czyli dostosowanie „siły" instrumentu do skali nierównowagi) i ocena skuteczności. Mogą też wystąpić problemy w ramach komunikacji z rynkiem. Dlatego też należy stosować instrumenty nie tylko skuteczne, lecz także wiążące się z ponoszeniem mniejszych kosztów przez władze i banki.

7.6.2. Reguły czy dyskrecjonalność?

W przypadku każdej polityki publicznej ogłoszenie i stosowanie ustalonych reguł postępowania zwiększa przewidywalność, obniża premie za ryzyko w rynkowych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

stopach procentowych i zmniejsza koszty dostosowań po stronie uczestników rynku. W praktyce w polityce makroostrożnościowej dominuje jednak dyskrecjonalność. Wynika to stąd, że polityka makroostrożnościowa na obecnym etapie pozostaje wciąż oparta głównie na procesie uczenia się zarówno przez władze nadzorcze jak i uczestników rynku. Ponieważ prawidłowa kalibracja instrumentów zależy od rodzaju szoku, tempa narastania ryzyka i wielu innych parametrów, władze nadzorcze znajdują sposobność, aby wykorzystywać metodę prób i błędów w procesie nadzoru. Lecz dyskrecjonalność pociąga za sobą także dodatkowe koszty ze względu na rosnącą niepewność co do przyszłości. To powoduje wzrost premii za ryzyko oraz wzrost kosztów dostosowań organizacyjnych, a w rezultacie skala działalności gospodarczej ulega ograniczeniu. Dlatego też, aby obniżyć te koszty, należy stosować ustalone reguły tam, gdzie to możliwe, a działania dyskrecjonalne podejmować z odpowiednim wyprzedzeniem i unikając szoków. Każdorazowo potrzebny jest jednak precyzyjnie zdefiniowany zakres polityki makroostrożnościowej (*policy framework*) wraz z zasadami identyfikacji i monitorowania różnych rodzajów ryzyka oraz wykorzystania instrumentów makroostrożnościowych w taki sposób, aby obniżyć koszty koniecznej dyskrecjonalności.

7.6.3. Instrumenty stałe czy zmieniane w czasie?

Ponieważ narastanie ryzyka systemowego ma w wielu przypadkach charakter procykliczny, instrumenty makroostrożnościowe są często dostosowywane w taki sposób, aby stopniowo zwiększać ich antycykliczne oddziaływanie w miarę narastania boomu kredytowego. Dotyczy to zwłaszcza instrumentów regulujących zakres kredytu (LTV, DTI, wymogi dotyczące rezerw), a także, choć już w mniejszym stopniu, narzędzi regulujących wielkość kapitału.

Niektóre instrumenty makroostrożnościowe mają z natury charakter antycykliczny, np. zmienne odpisy na rezerwy czy dodatkowy bufor kapitałowy w ramach Basel III. Funkcjonują one wówczas jako „automatyczne stabilizatory", ponieważ ich konstrukcja sprawia, że rosną w czasie hossy i maleją w czasie bessy, by w ten sposób wygładzać cykl kredytowy. Większość instrumentów makroostrożnościowych nie dostosowuje się jednak automatycznie i trzeba je każdorazowo zmieniać, tak aby ograniczały szkodliwą procykliczność działalności banków. W ten sposób antycykliczne dostosowania instrumentów wzmacniają siłę ich oddziaływania. Jak wspomniano wyżej, o tyle, o ile jest to możliwe, dostosowania te powinny być dokonywane na podstawie uzasadnionych i przejrzystych reguł, aby banki były

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w stanie przygotować się na wdrożenie wymaganych modyfikacji. Należy natomiast unikać zbyt częstych i dyskrecjonalnie dokonywanych zmian, które mogą destabilizować system finansowy.

7.6.4. Polityka makroostrożnościowa a polityka makroekonomiczna (fiskalna, pieniężna, kursowa)

Polityka makroekonomiczna odnosi się do zagregowanych wielkości kredytu, pieniądza, popytu i wydatków. Przykładowo polityka pieniężna kształtuje cenę pieniądza rezerwowego i wpływa na popyt na kredyty w zakresie całej gospodarki. Dlatego też instrumenty polityki pieniężnej czy fiskalnej powinny być stosowane w sytuacji, gdy zagrożenie stabilności dotyczy całego sektora, ma charakter homogeniczny i jest odwracalne. Ponadto, instrumenty polityki pieniężnej (zmiana stóp procentowych, operacje otwartego rynku) mogą być zastosowane bardzo szybko, a także są przejrzyste i łatwe w komunikacji.

W przeciwieństwie do tego instrumenty makroostrożnościowe wymagają większych nakładów czasu i pracy na wdrożenie poszczególnych rozwiązań i pociągają za sobą większe koszty transakcyjne w przypadku banków. W porównaniu z polityką makroekonomiczną instrumenty makroostrożnościowe można jednak zrazem precyzyjniej adresować do określonych części sektora finansowego i kalibrować pod względem dostosowywania do określonych rodzajów ryzyka. Na przykład limity na wskaźniki LTV lub DTI, a także poziom opłat od zagranicznych zobowiązań (*levies*) można różnicować w zależności od wielkości kredytu lub jego przyrostu, miejsca położenia i wartości zabezpieczenia oraz typu zobowiązania (IMF 2011). Niższy limit LTV w zakresie kredytów na zakup droższych domów pozwala ograniczyć ryzyko kredytowe w bankach, a wyższy limit LTV w zakresie kredytów na zakup tańszych mieszkań może być uzasadniony ze względów społecznych.

Podejście indywidualnie adresowane (*targeted*) posiada jednak również pewne wady. Wymaga dostępu do bardziej szczegółowych danych, generuje wyższe koszty administracyjne i jest łatwiejsze do ominięcia. Nadmierny *targeting* może też prowadzić do zjawiska „ręcznego sterowania" (*micromanagement*), co zwiększa koszty związane z prowadzeniem polityki i obniża jej efektywność. Dlatego też dodatkowe korzyści wynikające z podejścia indywidualnie adresowanego należy zawsze zestawiać z jego potencjalnymi kosztami.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W praktyce wiele krajów stosuje instrumenty makroostrożnościowe wraz z określoną polityką pieniężną i/lub fiskalną[31]. Takie równoległe stosowanie instrumentów polityki staje się uzasadnione wtedy, gdy cykl kredytowy pokrywa się z cyklem koniunkturalnym; wówczas instrumenty makroostrożnościowe są stosowane jako część całościowej polityki stabilizacyjnej w celu zasadniczego hamowania popytu w czasie boomu i pobudzania go w warunkach recesji. Uważa się, że łączne stosowanie instrumentów polityki makroekonomicznej oraz instrumentów makroostrożnościowych podnosi skuteczność funkcjonowania obu rodzajów polityki (Unsal 2011). Badania przeprowadzone przez Międzynarodowy Fundusz Walutowy potwierdzają również, że instrumenty makroostrożnościowe wywierają statystycznie istotny wpływ na hamowanie nadmiernego wzrostu kredytu w okresach hossy, obniżanie stopnia lewarowania, zmniejszanie ryzyka płynności, lepsze dopasowanie terminów zapadalności zobowiązań do posiadanych aktywów oraz łagodzenie skali przepływów kapitału i wahań kursu walutowego (ramka 7.3).

Ramka 7.3. Wyniki badania panelowego MFW

Jedną z pierwszych całościowych analiz efektywności zastosowania różnych instrumentów makroostrożnościowych było badanie przeprowadzone przez Międzynarodowy Fundusz Walutowy na próbie 49 krajów w latach 2000–2010 (Lim i in. 2011). W analizie uwzględniono dziesięć instrumentów, by sprawdzić, czy przyczyniają się one do ograniczenia procykliczności rozmiarów kredytu, wielkości lewarowania oraz zmian w poziomie ryzyka systemowego związanego z płynnością i przepływami kapitału. W badaniu wykorzystano metodę regresji panelowej.

Uzyskane wyniki potwierdziły stosunkowo wysoką skuteczność instrumentów makroostrożnościowych. Okazało się, że większość analizowanych narzędzi wywiera statystycznie istotny wpływ na hamowanie nadmiernego wzrostu kredytu w okresach hossy, a także na obniżenie stopnia lewarowania, spadek ryzyka płynności, lepsze dopasowanie terminów zapadalności zobowiązań do posiadanych aktywów oraz na łagodzenie skali przepływów kapitału i wahań kursu walutowego. Podsumowanie wyników badania MFW przedstawia tabela 7.3.

[31] Na przykład Chiny, Hongkong SAR i Singapur stosowały podatki na obrót nieruchomościami równolegle do obniżonych współczynników LTV, tak aby zwiększyć ich wpływ na ograniczenie tempa wzrostu kredytów i tempa wzrostu cen nieruchomości. Z kolei w krajach CEE w okresie boomu kredytowego prowadzono restrykcyjną politykę pieniężną równolegle z ograniczeniami na kredyty w *fx* (IMF 2011).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 7.3. Skuteczność instrumentów makroostrożnościowych (MP) w badaniu MFW

Rodzaj instrumentu MP	Efekty zastosowania
Instrumenty regulujące rozmiar i tempo wzrostu kredytu	Instrumenty MP służące regulowaniu kredytu mogą być skuteczne w ograniczaniu ryzyka wynikającego z nadmiernego wzrostu akcji kredytowej i/lub wzrostu cen aktywów. Limity LTV i DTI powinny obowiązywać jako reguła, dostosowywane antycyklicznie lub adresowane do poszczególnych źródeł ryzyka. Można je uzupełniać wymogami rezerwy lub instrumentami kapitałowymi (np. zmienne w czasie wymogi odpisu na rezerwy), w sytuacji, gdy boom kredytowy staje się powszechny, a instrumentami służącymi regulowaniu kredytów w *fx*, jeśli ich wzrost staje się źródłem ryzyka. Podobnie skuteczne są limity kredytowe, rezerwy obowiązkowe i zmienne w czasie odpisy na kredyty z utratą wartości.
Instrumenty regulujące zjawisko lewarowania	Instrumenty MP regulujące wartość kapitału mogą być skuteczne w ograniczaniu ryzyka wynikającego z nadmiernego lewarowania. Najskuteczniejsze są bufory kapitałowe, które można zmieniać antycyklicznie poprzez zmiany współczynnika adekwatności kapitałowej (CAR), zmiany wag ryzyka na poszczególnych aktywach lub zmiany wielkości odpisów rezerw na zagrożone aktywa, co pozwala ograniczać nadmierny wzrost lewarowania. Gdy wzrost lewarowania jest spowodowany dążeniem banku do zwiększenia udzielanych kredytów, instrumenty kapitałowe można uzupełniać instrumentami kredytowymi, co pozwala ograniczać ryzyko już u jego źródeł.
Instrumenty regulujące płynność środków finansowych	Instrumenty MP regulujące płynność środków finansowych – takie jak zarządcze miary płynności, limity na niedopasowanie walutowe i limity na NOP w *fx* – mogą być wykorzystane w ograniczaniu systemowego ryzyka płynności, gdy wynika ono z finansowania w walutach obcych. Udział *core liabilities vs non-core liabilities* może okazać się użytecznym wskaźnikiem poziomu ryzyka związanego z finansowaniem hurtowym, w tym pochodzącym z zagranicy. Określony limit lub opłata od środków zagranicznych (*levy*) może zapobiegać narastaniu systemowego ryzyka utraty płynności.
Instrumenty regulujące transgraniczne przepływy kapitału	Gdy zwiększone ryzyko wiąże się z napływem kapitału, wszystkie trzy rodzaje instrumentów (kredytowe, płynnościowe i kapitałowe) mogą znaleźć zastosowanie. W badaniu wykorzystano wskaźnik zadłużenia zagranicznego jako zmienną mierzącą wspólne ryzyko wynikające z przepływów kapitału i zmienności kursu. Jedyną zmienną statystycznie istotną jest limit na NOP w *fx*: w przypadku danego poziomu zagranicznych aktywów, poziom zagranicznych zobowiązań kraju jest niższy o 15% w krajach, które stosują limity na NOP. Oznacza to, że instrumenty płynnościowe, takie jak limit na NOP w *fx*, okazują się skuteczne w ograniczaniu zależności całego sektora od finansowania zagranicznego. Instrumenty płynnościowe mogą być uzupełnione instrumentami kredytowymi, gdy chęć gwałtowanego zwiększenia akcji kredytowej skłania banki do pożyczania środków za granicą. W tym kontekście również instrumenty kapitałowe mogą zostać wykorzystane w celu ograniczenia wielkości i tempa wzrostu kredytu poprzez zwiększone bufory kapitałowe.

Źródło: opracowanie własne na podstawie Lim i in. (2011).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Badanie dowodzące istotnej komplementarności elementów polityki pieniężnej i polityki makroostrożnościowej przeprowadzili Brzoza-Brzezina, Kolasa i Makarski (2015). Autorzy zwrócili uwagę na to, że przebieg kryzysu finansowego w krajach strefy euro był zróżnicowany – w szczególności ekspansja kredytu bankowego i wynikający z niej wzrost skali zadłużenia sektora prywatnego w bankach okazał się znacznie większy w peryferyjnych krajach strefy (Grecja, Irlandia, Hiszpania, Portugalia) niż w państwach „centrum" (*core*). Na tej podstawie autorzy opracowali model makrofinansowy symulujący zachowanie się konsumentów w krajach peryferyjnych i państwach „centrum" w zależności od warunków dostępności kredytu na zakup nieruchomości, a następnie zbadali w jego ramach skuteczność polityki pieniężnej i polityki makroostrożnościowej (wykorzystując wskaźnik LTV) w przywracaniu stabilności makroekonomicznej i stabilności finansowej. Otrzymane wyniki nie tylko potwierdziły istotną rolę rynku nieruchomości w generowaniu cyklu *boom-bust*, lecz także wykazały, że polityka makroostrożnościowa może być ważnym instrumentem reagowania na szoki asymetryczne wywołane przez zmiany popytu na nieruchomości w poszczególnych krajach. W szczególności zaś autorzy pokazali, że o ile wspólna polityka pieniężna jest skuteczna w niwelowaniu wahań produkcji w całej strefie euro, o tyle okazuje się mniej efektywna przy stabilizowaniu cyklu kredytowego i przywracaniu równowagi zewnętrznej w poszczególnych krajach członkowskich. Z kolei polityka makroostrożnościowa jest mało skuteczna w korygowaniu cyklu koniunkturalnego, lecz pozostaje jednocześnie bardzo efektywna w przypadku stabilizowania cyklu kredytowego i przywracania równowagi w bilansie obrotów bieżących.

7.6.5. Ramy instytucjonalne polityki makroostrożnościowej

Polityka makroostrożnościowa musi być prowadzona w określonych ramach instytucjonalnych, zapewniających przejrzystość, przewidywalność i skuteczność działania. Powinny on sprzyjać umożliwieniu identyfikacji narastającego ryzyka, stworzeniu silnych bodźców do podejmowania szybkich i skutecznych działań w celu obniżenia tego ryzyka, dążeniu do wykorzystania najskuteczniejszych instrumentów i ułatwienia koordynacji różnych odmian polityki mającej wpływ na ryzyko systemowe. W ramach polityki makroostrożnościowej powinny się więc zawierać elementy takie jak mechanizm identyfikowania i monitorowania ryzyka systemowego, kryteria wyboru konkretnych celów do realizacji, kryteria wyboru

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

instrumentów makroostrożnościowych, procedury ich stosowania oraz metodologia oceny ich skuteczności.

Należy unikać zbyt złożonych i nadmiernie rozczłonkowanych struktur instytucjonalnych. Gdy odpowiedzialność za podejmowane decyzje jest rozłożona na zbyt wiele podmiotów, podziały kompetencyjne i rywalizacja między instytucjami przeszkadzają w identyfikacji i ocenie ryzyka systemowego oraz zwiększają koszty działania, obniżając jednocześnie skuteczność polityki makroostrożnościowej. Ponadto, zapewnienie silnych bodźców do szybkiego, wyprzedzającego rozwój wypadków działania wymaga jasnego określenia podmiotu nadrzędnego, wyposażonego w klarowny mandat i stosowne do tego prerogatywy decyzyjne, tak aby umożliwić mu przejęcie kontroli i odpowiedzialności za rezultaty osiągane w ramach stosowanej polityki makroostrożnościowej. Wreszcie, ponieważ wiele różnych instrumentów jest często wykorzystywanych jednocześnie, jasna strategia komunikowania się z rynkami pozostaje konieczna w celu zapewnienia przejrzystości i wiarygodności polityki.

Tworząc solidne ramy instytucjonalne, należy jasno określić relacje zachodzące między organem państwowym odpowiedzialnym za politykę makroostrożnościową z jednej strony a bankiem centralnym i rządem z drugiej strony. Niezależnie od tego, czy nadzór ostrożnościowy zostanie umiejscowiony w banku centralnym, czy w odrębnej niezależnej instytucji, bank centralny zawsze odgrywa ważną rolę w polityce makroostrożnościowej ze względu na posiadane informacje dotyczące banków, kompetencje w zakresie polityki pieniężnej i stabilności finansowej oraz odpowiedzialność z tytułu funkcji pożyczkodawcy ostatniej szansy. Wobec tego niezależny bank centralny ma podstawy do podejmowania działań w sytuacji narastającego ryzyka. Ponadto, silna pozycja banku centralnego w polityce makroostrożnościowej ułatwia niezbędną koordynację z polityką pieniężną, wpływającą na zagregowane poziomy popytu i kredytu – szerzej na ten temat piszą m.in. Goodhart (2011) oraz Tymoczko (2012).

Z kolei współdziałanie z rządem jest potrzebne w celu zapewnienia wsparcia dla polityki makroostrożnościowej ze strony polityki podatkowej i podejmowania niezbędnych inicjatyw legislacyjnych koniecznych w celu ograniczenia ryzyka systemowego (w tym tworzenia regulacji i systemu nadzoru w przypadku instytucji finansowych niebankowych – tzw. parabanków). Należy jednak pamiętać o tym, że ze względu na polityczną naturę władzy wykonawczej, silna rola rządu w polityce makroostrożnościowej może stwarzać ryzyko odwlekania niezbędnych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

decyzji, gdyby ich skutkiem miało stać się ograniczanie tempa wzrostu kredytów w tzw. „dobrych czasach".

7.6.6. Podstawy polityki makroostrożnościowej w Unii Europejskiej

W reakcji na kryzys finansowy z 2008 roku i rekomendacje raportu de Larosière'a (de Larosière 2009) państwa członkowskie UE podjęły prace nad stworzeniem ram całościowej polityki makroostrożnościowej. W kolejnych latach przyjęto w UE szereg rozwiązań instytucjonalnych mających zapobiegać narastaniu ryzyka systemowego. W 2010 roku utworzono Europejską Radę Ryzyka Systemowego (*European Systemic Risk Board* – ESRB) jako unijny organ koordynujący politykę makroostrożnościową państw członkowskich[32]. Nowy system regulacji makroostrożnościowych obowiązuje od 1 stycznia 2014 roku, po wejściu w życie dyrektywy CRD IV (*Capital Requirements Directive IV*) i rozporządzenia CRR (*Capital Requirement Regulation*), które przenoszą na grunt prawa europejskiego rekomendacje komitetu bazylejskiego Basel III. Nowe regulacje prawne przewidują, że polityka makroostrożnościowa w UE będzie prowadzona wspólnie przez ESRB i narodowe instytucje nadzorcze.

Na obszarze UE nie istnieje jednolity model instytucjonalny w zakresie nadzoru makroostrożnościowego. Za bezpośrednią kontrolę w poszczególnych państwach odpowiedzialne są narodowe organy nadzoru finansowego, które albo są umiejscowione w banku centralnym, albo stanowią odrębną instytucję (zob. tabela 7.4). Wśród ekspertów nie ma zgody co do *systemowej* wyższości jednej formy nadzoru nad drugą (piszemy na ten temat powyżej w punkcie 7.6.5 – zob. także Goodhard i Schoenmaker 1995). Jak pokazały doświadczenia wielu państw w okresie poprzedzającym kryzys finansowy, zarówno w ramach jednego, jak i drugiego systemu można prowadzić bardziej lub mniej skuteczną politykę nadzorczą[33].

Z kolei do obowiązków ESRB należy, między innymi, zbieranie informacji i danych sprawozdawczych z banków, identyfikacja ryzyka systemowego, monitorowanie i ocena polityki makroostrożnościowej prowadzonej przez poszczególne państwa, wydawanie opinii i (niewiążących) rekomendacji[34] oraz udzielanie

[32] Na podstawie Rozporządzenia EU 1092/2010.

[33] Państwa UE najbardziej dotknięte kryzysem bankowym miały nadzór usytuowany zarówno w banku centralnym (Irlandia, Hiszpania), jak i poza nim (Łotwa).

[34] Niewiążący charakter rekomendacji nie zwalnia państw członkowskich z obowiązku „zastosowania się lub wyjaśnienia odmowy" zgodnie z procedurą *comply or explain* (art. 16 i 17 rozporządzenia nr 1092 z 2010 roku).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wsparcia technicznego na prośbę państw członkowskich we wdrażaniu regulacji CRD i CRR (przy tworzeniu buforów anty-cyklicznych)[35].

Niezależnie od tego, czy nadzór zlokalizowany jest formalnie w banku centralnym, czy też poza nim, bank centralny pozostaje zawsze ważnym ogniwem procesu nadzorczego ze względu na swoje kompetencje w zakresie nadzoru mikroostrożnościowego nad bankami oraz na odpowiedzialność za politykę pieniężną. Wobec tego w skład Rady Ogólnej ESRB (*General Board*), oprócz szefów odrębnych krajowych instytucji nadzoru makroostrożnościowego, wchodzą także z urzędu prezesi banków centralnych spośród państw członkowskich UE.

Tabela 7.4. Usytuowanie nadzoru makroostrożnościowego nad bankami w krajach członkowskich UE – stan na dzień 1.01.2016 roku

Usytuowanie w banku centralnym	Usytuowanie w innej instytucji
Belgia	Austria
Bułgaria	Dania
Chorwacja	Finlandia
Cypr	Francja
Czechy	Luksemburg
Estonia	Łotwa
Grecja	Malta
Hiszpania	Niemcy
Holandia	Polska
Irlandia	Szwecja
Litwa	
Portugalia	
Rumunia	
Słowacja	
Słowenia	
Węgry	
Wielka Brytania	
Włochy	

Źródło: opracowanie własne na podstawie danych ESRB.

7.6.7. Ramy instytucjonalne polityki makroostrożnosciowej w Polsce

Cele, zakres oraz instrumenty polityki makroostrożnościowej w Polsce zostały określone w ustawie z dnia 5 sierpnia 2015 roku o nadzorze makroostrożnościowym nad systemem finansowym. Ustawa weszła w życie 1 października 2015 roku[36]. Zgodnie z jej zapisami, nadzór makroostrożnościowy nad systemem finansowym

[35] ESRB 2014.

[36] Ustawa z dnia 5 sierpnia 2015 r. o nadzorze makroostrożnościowym nad systemem finansowym i zarządzaniu kryzysowym w systemie finansowym (DzU 2015, poz. 1513).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

obejmuje identyfikację, ocenę i monitorowanie ryzyka systemowego powstającego w systemie finansowym lub jego otoczeniu oraz działanie na rzecz wyeliminowania lub ograniczania tego ryzyka z wykorzystaniem instrumentów makroostrożnościowych. Z kolei celem nadzoru makroostrożności jest przede wszystkim zwiększanie odporności systemu finansowego na wypadek materializacji ryzyka systemowego i wspieranie poprzez to długookresowego, zrównoważonego wzrostu gospodarczego kraju. Nadzór makroostrożnościowy realizowany jest przez Komitet Stabilności Finansowej, w skład którego wchodzą: Minister Finansów, Prezes Narodowego Banku Polskiego, Przewodniczący Komisji Nadzoru Finansowego oraz Prezes Zarządu Bankowego Funduszu Gwarancyjnego. Rozwiązania ustawowe przyjęte w Polsce są zgodne z rekomendacjami Unii Europejskiej, zawartymi w dyrektywie CRD i w rozporządzeniu CRR, które weszły w życie w 2014 roku.

Ustawa określa także katalog instrumentów polityki makroostrożnościowej, obejmujących przede wszystkim obowiązkowe wymogi kapitałowe w postaci tzw. buforów (bufor zabezpieczający, bufor antycykliczny, bufory instytucji ważnych systemowo, bufor ryzyka systemowego), ograniczenia wypłat z zysku oraz wymogi informacyjne i sprawozdawcze.

7.7. Transgraniczna transmisja szoków i potrzeba międzynarodowej koordynacji polityki makroostrożnościowej

Procesy integracji zachodzące na rynkach finansowych w skali globalnej w ostatnim dwudziestoleciu spowodowały, że narodowe systemy finansowe w poszczególnych gospodarkach stały się wzajemnie powiązane w skali międzynarodowej. Ta rosnąca współzależność jest wynikiem stopniowego znoszenia barier w transgranicznym przepływie kapitałów, któremu towarzyszył również skokowy postęp w rozwoju technologii informatycznych i telekomunikacyjnych. W rezultacie usługi finansowe są w coraz większym zakresie świadczone transgranicznie, co jeszcze bardziej uzależnia od siebie poszczególne gospodarki.

W warunkach silnej współzależności potrzebna jest intensywniejsza koordynacja polityki gospodarczej między poszczególnymi krajami, aby ograniczyć negatywne efekty oddziaływania polityki jednych państw na drugie (tzw. *spillover effects*)[37].

[37] Szerzej na ten temat piszemy w rozdziale 8.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Dotyczy to również polityki makroostrożnościowej. Po pierwsze, niezbędna staje się koordynacja międzynarodowa, ponieważ boomy kredytowe i spekulacyjne bańki cen aktywów są często finansowane przez napływ kapitału z zagranicy. W takiej sytuacji działania pojedynczego kraju mogą okazać się nieskuteczne. Po drugie, koordynacja polityki makroostrożnościowej jest wskazana również po to, aby ograniczać możliwości przeprowadzania arbitrażu regulacyjnego, czyli przenoszenia działalności bankowej do najmniej restrykcyjnych jurysdykcji. Wreszcie, koordynacja może okazać się korzystna wówczas, gdy pozwala ograniczać negatywne skutki transmisji szoków (*spillovers*) wywoływanych przez zróżnicowaną, indywidualną politykę gospodarczą poszczególnych krajów w warunkach globalnej integracji finansowej.

Wyliczając różne stopnie koordynacji polityki makroostrożnościowej, począwszy od najsłabszej, można wyróżnić następujące jej formy (CIEPR 2012):

1) wymiana informacji między krajami goszczącymi i krajami pochodzenia (*host-home*), na przykład na temat kondycji finansowej banku matki i banków córek;
2) wspólne reguły (*single rule book*), które ograniczają przestrzeń arbitrażu regulacyjnego między poszczególnymi jurysdykcjami narodowymi. Przykładem tego typu koordynacji są umowy o stosowaniu jednolitych standardów ostrożnościowych w bankowości w ramach Basel I, II i III;
3) wspólne reguły i nadzór, nie tylko pozwalające ograniczać arbitraż regulacyjny, lecz także umożliwiające internalizowanie efektów zewnętrznych (*externalities*) poprzez działania nadzorcze;
4) wspólne reguły, nadzór oraz mechanizm restrukturyzacji i likwidacji banków oraz wspólny system wzajemnych ubezpieczeń (unia bankowa).

W przypadku niezastosowania koordynacji polityki makroostrożnościowej transmisję szoków można ograniczyć poprzez wprowadzenie jednostronnej kontroli transgranicznego przepływu kapitałów. W ramach UE takie działania są zabronione, m.in. dlatego, że koszt powrotu do posegmentowanych rynków narodowych byłby bardzo duży. Innym sposobem jest nałożenie limitu na wielkość krajowych depozytów, jakie filie mogą utrzymywać w banku matce (lub w innych filiach) w celu ograniczenia ryzyka związanego z wykorzystaniem lokalnego banku do finansowania banku matki (i odwrotnie), co pomaga utrzymać stabilność w kraju goszczącym. Wreszcie można też nałożyć limity na krótkoterminowe pożyczki zaciągane za granicą (*non-core liabilities*), chociaż takie rozwiązanie segmentuje rynki finansowe i ogranicza dostęp państwa do zagranicznego kapitału.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

7.8. Zakończenie

W niniejszym rozdziale wyjaśniliśmy mechanizm powstawania ryzyka systemowego w sektorze bankowym oraz opisaliśmy pożądany kształt polityki makroostrożnościowej, której celem jest ograniczanie tego zjawiska. O potrzebie jej stosowania świadczy kryzys finansowy z lat 2008–2010. Należy jednak mieć świadomość, że polityka ta nie może być czymś w rodzaju *panaceum* – nawet najlepiej uprawiana polityka makroostrożnościowa nie jest w stanie całkowicie zapobiec kryzysom finansowym. Dlatego też niezbędna okazuje się koordynacja polityki makroostrożnościowej i pozostałych rodzajów polityki makroekonomicznej, także w skali międzynarodowej. Jak argumentowaliśmy w rozdziale 5, konieczne staje się powołanie instytucji „pożyczkodawcy ostatniej szansy" – tę funkcję z reguły pełni bank centralny – aby łagodzić przejściowe braki płynności w sektorze bankowym. Potrzebny jest również przejrzysty mechanizm służący likwidacji banków niewypłacalnych wraz z regułami dotyczącymi sposobu alokacji strat. Ponadto, ponieważ polityka makroostrożnościowa nie znajduje zastosowania w próżni, istotne stają się też jakość i sposób powiązania jej z polityką pieniężną i fiskalną – bez tego rodzaju koordynacji system finansowy nie będzie prawidłowo funkcjonował. Z drugiej strony, polityka makroostrożnościowa może okazać się mało skuteczna w warunkach możliwości przeprowadzenia arbitrażu regulacyjnego i występowania nacisków politycznych.

Polityka makroostrożnościowa – jak każda działalność publiczna – pociąga za sobą koszty związane z gromadzeniem informacji i koniecznością przeprowadzania odpowiednich regulacji przez banki oraz inne instytucje finansowe, co może zakłócać funkcjonowanie mechanizmu rynkowego i obniżać efektywność gospodarowania. Mamy wówczas do czynienia z wymiennością, jaka zachodzi na poziomie efektywności i stabilności. Na przykład wyższe wymogi kapitałowe i płynnościowe wzmacniają stabilność systemu finansowego, lecz jednocześnie zwiększają koszt kredytu i mogą obniżać tempo wzrostu gospodarczego. W takiej sytuacji rozstrzyganie o kształcie polityki makroostrożnościowej może więc wymagać jakościowej wyceny (*value judgement*) związanych z nią pośrednich kosztów i korzyści w ramach różnych wariantów polityki i z uwzględnieniem wielu rozmaitych założeń.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Literatura cytowana

Acemoglu D., Ozdaglar A., Tahbaz-Salehi A., *Systemic Risk and Stability in Financial Networks*, "NBER Working Paper" 2013, no. 18727, http://www.nber.org/papers/w18727 (12.02.2016).

Acharya V.V., *A theory of systemic risk and design of prudential regulation*, "Journal of Financial Stability" 2009, no. 5.

Acharya V.V., Pedersen L.H., Philippon T., Richardson M.P., *Measuring Systemic Risk*, "CEPR Discussion Papers" 2012, no. 8824.

Acharya V.V., Engle R., Richardson M., *Capital shortfall: a new approach to ranking and regulating systemic risks*, American Economic Association Meeting, 7 January, 2012, mimeo.

Adrian T., Brunnermeier M.K., *COVAR*, "NBER Working Paper" 2011, no. 17454.

Aiyar S., Calomiris C.W., Wielądek T., *Does Macro-Prudential Policy Leak: Evidence from a UK Policy Experiment*, "NBER Working Paper" 2013, no. 17822.

Allen F., Gale D., *Understanding financial crises*, Oxford University Press, Oxford 2007.

Almeida H., Campello M., Liu C., *The Financial Accelerator:Evidence from International Housing Markets*, "Review of Finance" 2006, no. 10.

Bernanke B., Gertler M., *Monetary policy and asset price volatility*, "Economic Review" 1999, issue Q IV.

Bernanke B., Gertler M., Gilchrist S., *The financial accelerator in a quantitative business cycle framework*, w: *Handbook of Macroeconomics*, red. J.B. Tylor, M. Woodford, wydanie 1, t. 1, Elsevier, North Holland 1999.

Billio M., Getmansky M., Lo A.W., Pelizzon L., *Measuring Systemic Risk in the Finance and Insurance Sectors,* "MIT Sloan School Working Paper" 2010, no. 4774–10.

Blanchard O.J., Dell'Ariccia G., Mauro P., *Rethinking Macroeconomic Policy*, IMF Staff Position Note, SPN/10/03, 2010.

Borio C., *Towards a Macroprudential Framework for Financial Supervision and Regulation?*, "BIS Working Paper" 2003, no. 128.

Borio C., *Implementing a Macroprudential Framework: Blending Boldness and Realism*", "Capitalism and Society" 2011, vol. 6, issue 1, article 1.

Brownlees C.T., Engle R.F., *Volatility, Correlation and Tails for Systemic Risk Measurement*, "Working Paper", 24 May, 2001, NYU Stern School of Management.

Brunnermeier M. i in., *The Fundamental Principals of Financial Regulation,* 11th Geneva Report on the World Economy, Geneva 2009.

Brunnermeier M.K., Pedersen L.H., *Market Liquidity and Funding Liquidity*, "The Review of Financial Studies" 2009, no. 22.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Brzezina-Brzoza M., Kolasa M., Makarski K., *Macroprudential policy and imbalances in the euro area*, "Journal of International Money and Finance" 2015, vol. 51.

Calomiris C., *Reducing the Risks of the New Macro-Prudential Policy Regime*, Manhattan Institute for Policy Research and Shadow Open Market Committee, New York 2013.

Cecchetti S., *The future of financial intermediation and regulation*, "Current Issues in Economics and Finance" 1999, no. 5(8).

CIEPR, *Banks and Cross-Border capital Flows: Policy Challenges and Regulatory Responses*, Brookings, September 2012.

Clement P., *The term 'macroprudential': origins and evolution*, "BIS Quarterly Review", March 2010.

Crockett A., *Marrying the Micro- and Macro-prudential Dimensions of Financial Stability*, remarks before the 11th International Conference of Banking Supervisors, September 20–21, Basel 2000.

Crowe C., Dell'Ariccia G., Igan D., Rabanal P., *Policies for Macrofinancial Stability: Options to Deal with Real Estate Booms*, IMF Staff Discussion Note, SDN/11/02, 2011.

De Bandt O., Hartmann Ph., *Systemic Risk. A Survey*, "Working Paper" 2000, no. 35, European Central Bank.

De Larosière J., *Report of the High-Level Group on Financial Supervision in the EU*, chaired by J. de Larosière, Bruxelles, 25 February, 2009.

Diamond D.W., Dybvig Ph.H., *Bank Runs, Deposit Insurance, and Liquidity*, "Journal of Political Economy" 1983, no. 91(3).

Elliott D.J., Feldberg G., Lehnert A., *The History of Cyclical Macroprudential Policy in the United States*, Finance and Economics Discussion Series, Divisions of Research and Statistics and Monetary Affairs, Federal Reserve Board, Washington 2013.

Elliott D., *An Overview of Macroprudential Policy and Countercyclical Capital Requirements*, 10 March 2011, The Brookings Institution, Washington.

Engle R.F., Manganelli S., *CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles*, "Journal of Business and Economic Statistics" 2004, no. 23(4).

ESRB, *Flagship Report on Macro-prudential Policy in the Banking Sector*, March 2014, Frankfurt.

Galati G., Moessner R., *Macroprudential Policy—A Literature Review*, "BIS Working Papers" 2011, no. 337.

Goodhart C.A.E., *The Macro-Prudential Authority: Powers, Scope and Accountability*, "OECD Journal: Financial Market Trends" 2011, vol. 2011, issue 2.

Goodhart C.A.E., Schoenmaker D., *Should the Functions of Monetary Policy and Banking Supervision be Separated?*, "Oxford Economic Papers" 1995, no. 47.

Gray D., Jobst A.A., *Tail Dependence Measures of Systemic Risk Using Equity Options Data. Implications for Financial Stability*, "IMF Working Paper" 2009, no. 1354.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Grey D., Jobst A.A., *Systemic Contingent Claims Analysis – A Model Approach to Systemic Risk*", w: *Managing Risk in the Financial System*, red. J.R. La Brosse, R. Olivares-Caminal, D. Singh, Edward Elgar, London 2011.

IMF, *Financial Stability and Macroprudential Policy Survey*, Washington 2010a.

IMF, *United States: Publication of Financial Sector Assessment Program Documentation – Technical Note on Stress Testing*, "IMF Country Report" 2010b, no. 10/244, http://www.imf.org/external/pubs/ft/scr/2010/cr10244.pdf (2.11.2015).

IMF, *Macroprudential Policy: An Organizing Framework*, 14 March, 2011a, Washington.

IMF, *United Kingdom: Stress Testing the Banking Sector Technical Note*, "IMF Country Report" 2011c, no. 11/222, http://www.imf.org/external/pubs/ft/scr/2011/cr11227.pdf (2.11.2015).

IMF, *Germany: Technical Note on Stress Testing*, "IMF Country Report" 2011d, no. 11/371, http://www.imf.org/external/pubs/ft/scr/2011/cr11371.pdf (2.11.2015).

IMF, *Recent Experiences in Managing Capital Inflows – Cross-Cutting Themes and Possible Policy Framework*, 14 February, 2011e, https://www.imf.org/external/np/pp/eng/2011/02a.pdf (3.11.2015).

IMF, *Key Aspects of Macroprudential Policy*, 10 June, 2013.

Jobst A.A., Grey D., *Systemic Contingent Claims Analysis – Estimating Market-Implied Systemic Risk*, "IMF Working Paper" 2013, no. 13/54.

Kindleberger C.P., *Manias, panics, and crashes: a history of financial crises*, Wiley, New York 1996.

Lim C. i in., *Macroprudential Policy: What Instruments and How to Use Them? Lessons from Country Experiences*, "IMF Working Paper" 2011, no. 11/238.

Magud N.E., Reinhart C.M., Rogoff K.S., *Capital Controls: Myth and Reality – A Portfolio Balance Approach*, "NBER Working Paper" 2011, no. 16805, http://www.nber.org/papers/w16805 (30.04.2014).

Minsky H.P., *The Financial Instability Hypothesis*, "The Jerome Levy Economics Institute Working Paper" 1992, no. 74.

Ostry J.D. i in., *Managing Capital Inflows: What Tools to Use?*, IMF Discussion Note, SDN/11/06, 2011.

Raport o stabilności systemu finansowego. Grudzień 2011 r., Narodowy Bank Polski, Warszawa 2011.

Rochet Jean-Charles R., Tirole J., *Interbank Lending and Systemic Risk*, "Journal of Money, Credit and Banking" 1996, no. 28(4).

Rosati D.K., *Bat na spekulację*, „Gazeta Wyborcza", 30 stycznia 2012.

Rosati D.K., *Polityka makroostrożnościowa i stabilność sektora bankowego*, „Bank i Kredyt" 2014, nr 4.

Segoviano M.A., *The Consistent Information Multivariate Density Optimizing Methodology*, "Discussion Paper" 2006, no. 557, Financial Markets Group, London School of Economics.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Segoviano M.A., Charles Goodhart C., *Banking Stability Measures*, "IMF Working Paper" 2009, no. 09/4.

Szpunar P., *Rola polityki makroostrożnościowej w zapobieganiu kryzysom finansowym*, „Materiały i Studia" 2012, nr 278, Narodowy Bank Polski, Warszawa 2012.

The Role of Macroprudential Policy, A Discussion Paper, Bank of England, November 2009.

Tressel T., Verdier T., *Optimal Prudential Regulation of Banks and the Political Economy of Supervision*, "IMF Working Paper" 2014, no. 14/90.

Tymoczko D., *Wybrane aspekty funkcjonowania polskiego systemu finansowego*, „Zeszyty Naukowe Kolegium Gospodarki Światowej" 2012, nr 34.

Unsal F.D., *Capital Flows and Financial Stability: Monetary Policy and Macroprudential Responses*, "IMF Working Paper" 2011, no. 11/189.

Viñals J., *Seeing Both the Forest and the Trees – Supervising Systemic Risk*, Opening Remarks at the 11th Annual International Seminar on Policy Challenges for the Financial Sector, Washington, 2 June, 2011, http://www.imf.org/external/np/speeches/2011/060211.htm (3.11.2015).

Zhou C., *Are Banks Too Big to Fail? Measuring Systemic Importance of Financial Institutions*, "International Journal of Central Banking", December 2010.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozdział 8

KOORDYNACJA POLITYKI GOSPODARCZEJ W SKALI MIĘDZYNARODOWEJ

8.1. Globalizacja

Globalizacja to rosnąca współzależność w skali międzynarodowej. Ostatnie dwie dekady XX i pierwsza dekada XXI wieku przyniosły niezwykle silny wzrost intensywności oraz zakresu powiązań handlowych, produkcyjnych i finansowych zarówno między poszczególnymi krajami, jak i w skali globalnej. Sprzyjały temu zmiany instytucjonalne, takie jak stopniowe usuwanie przeszkód i barier w przepływie dóbr, usług i kapitału oraz rozwój technologii informatycznych i telekomunikacyjnych. Regionalna integracja gospodarcza nasila zaś te tendencje poprzez harmonizację przepisów prawa i standardów technicznych w skali regionalnej i międzynarodowej, a także poprzez poszerzanie zakresu kompetencji instytucji międzynarodowych i ponadnarodowych. Globalizacja sprawia, że poszczególne kraje stają się od siebie coraz bardziej uzależnione, a próby izolowania się od wpływów zewnętrznych okazują się nie tylko mało skuteczne, lecz prowadzą także do stagnacji i zacofania. Politycy podejmujący decyzje w zakresie polityki gospodarczej w poszczególnych krajach nie mogą i na ogół nie chcą ignorować wpływu otoczenia zewnętrznego na gospodarkę ich państwa i funkcjonowanie rynków. W rezultacie wewnętrzna sytuacja gospodarcza poszczególnych krajów w coraz większej mierze zależy od tego, co dzieje się na rynkach międzynarodowych w ich bliższym i dalszym sąsiedztwie, oraz od tego, czy i w jakim stopniu życie gospodarcze i społeczne danego państwa regulują umowy międzynarodowe, zarówno bilateralne, jak i te zawierane w ramach wielostronnych porozumień.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

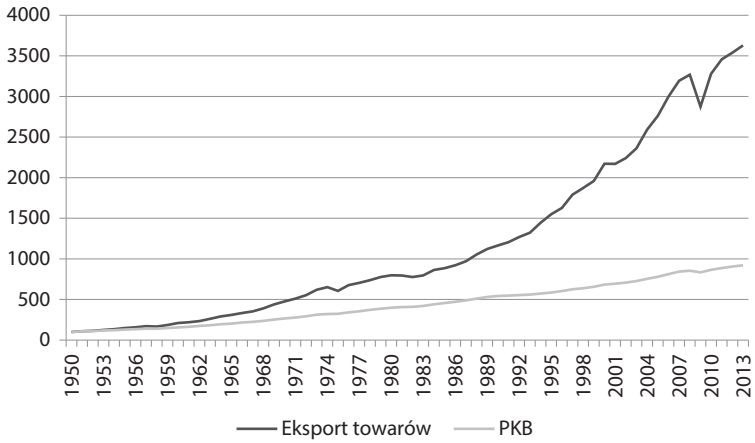
Wszystkie przywołane zjawiska wiążą się z szeregiem ograniczeń nakładanych na krajową politykę gospodarczą, a jednocześnie stawiają też przed nią nowe wyzwania. Z jednej strony zakres swobody krajowej polityki w kształtowaniu procesów gospodarczych kurczy się wraz ze wzrostem siły i skali oddziaływania powiązań, jakie występują między gospodarką krajową a zagranicą. Decyzje podejmowane w ramach polityki pieniężnej czy fiskalnej muszą uwzględniać w coraz większym stopniu sytuację panującą na międzynarodowych rynkach finansowych oraz reakcje zagranicznych inwestorów i skalę transgranicznych przepływów kapitału. Z kolei ustalenia przeprowadzane w zakresie polityki handlowej stają się w coraz większym stopniu zależne od międzynarodowych porozumień liberalizujących wymianę handlową. Z drugiej zaś strony dzięki współpracy państw pojawiają się nowe możliwości radzenia sobie ze zjawiskami o zasięgu ponadkrajowym i globalnym, takimi jak masowe migracje, międzynarodowa przestępczość, rosnące zanieczyszczenie środowiska naturalnego czy ocieplenie klimatu. Współzależność oznacza więc ograniczenie swobody polityki – czyli tradycyjnie pojmowanej suwerenności, lecz niesie też określone korzyści w związku z rozwojem handlu oraz inwestycji w skali międzynarodowej, a także za sprawą wywoływania presji na koordynację działań poszczególnych państw.

Miarą globalizacji w zakresie handlu może być stopień otwarcia poszczególnych gospodarek krajowych na wymianę międzynarodową, czyli stosunek eksportu lub importu (lub ich sumy) do PKB. Wykres 8.1. przedstawia wzrost światowego eksportu towarowego na tle wzrostu światowego PKB w latach 1950–2013 (w cenach stałych). W okresie tym eksport rósł średnio w tempie 5,6% rocznie, podczas gdy tempo wzrostu PKB wynosiło 3,5% rocznie. W rezultacie, o ile światowy PKB wzrósł w tym czasie około 9-krotnie, o tyle wielkość światowego eksportu zwiększyła się ponad 35-krotnie, a stosunek eksportu do PKB wzrósł z 6,4% w 1950 roku do 25,1% w 2013 roku. Obecnie poszczególne kraje są więc znacznie silniej powiązane przez handel niż kilkadziesiąt lat temu.

Jeszcze szybszy stał się postęp globalizacji w zakresie przepływów finansowych. Jej miarą może być stosunek sumy transgranicznych przepływów kapitału do PKB lub stosunek zagranicznych aktywów i pasywów w bilansach krajowych instytucji finansowych do PKB. Wykres 6.4 (zamieszczony w rozdziale 6) ukazuje zmiany w zakresie relacji sumy aktywów zagranicznych i zobowiązań zagranicznych do wielkości PKB w wybranych krajach w latach 1980–2011. W większości państw rozwiniętych, takich jak Francja, Niemcy, USA i Wielka Brytania, relacja ta wzrosła wówczas średnio cztero- lub pięciokrotnie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 8.1. Wzrost światowego eksportu towarów i światowego PKB w latach 1950–2013 (1950=100)



Źródło: opracowanie własne na podstawie danych WTO.

Postęp globalizacji nie oznacza jednak, że gospodarka światowa staje się w pełni i równomiernie zintegrowana. Część krajów, zwłaszcza tych najsłabiej rozwiniętych, tylko w niewielkim stopniu uczestniczy w handlu i przepływach kapitałowych. Co więcej, granice państwowe nadal stanowią istotną barierę w wymianie międzynarodowej, oddzielając od siebie obszary o różnych systemach prawnych, instytucjonalnych i walutowych. Sprawia to, że obrót gospodarczy w obrębie danego kraju pozostaje z reguły wciąż o wiele wyższy niż obrót transgraniczny, pochodzący z transakcji zawieranych między krajami. W teorii handlu zjawisko to nosi nazwę „efektu granicy" (*border effect* – McCallum 1995). Jego odpowiednikiem w dziedzinie przepływów finansowych jest zwiększona skłonność państwowych instytucji finansowych do inwestowania w krajowe papiery wartościowe, czyli efekt „własnego podwórka" (tzw. *home bias* – zob. Tesar i Werner 1992). Wszystko to sprawia, że globalizacja wcale nie jest w pełni „globalna".

W skali międzynarodowej bliskość geograficzna, pokrewieństwo kulturowe i związki historyczne pozostają nadal ważnymi czynnikami określającymi intensywność powiązań gospodarczych między krajami. To za ich sprawą dochodzi do powstania i rozwoju regionalnych ugrupowań integracyjnych, w skład których wchodzą kraje sąsiadujące ze sobą lub historycznie powiązane gospodarczo i gdzie handel oraz inwestycje są znacznie bardziej zliberalizowane niż w odniesieniu do krajów trzecich. Najbardziej zaawansowanym regionalnym ugrupowaniem

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

integracyjnym na świecie jest Unia Europejska, w ramach której funkcjonuje jednolity rynek i wspólna waluta. Inne ugrupowania integracyjne mają charakter głównie stref wolnego handlu o różnym zakresie liberalizacji wymiany. Do najważniejszych z nich należą: Północno-Amerykańska Strefa Wolnego Handlu (NAFTA), Mercosur, czyli unia celna obejmująca część państw Ameryki Łacińskiej, oraz ugrupowanie ASEAN, zrzeszające wybrane państwa Azji Południowo-Wschodniej.

Usuwając bariery ograniczające handel oraz inwestycje w obrębie poszczególnych regionów, ugrupowania te często utrzymują wiele utrudnień w wymianie gospodarczej z innymi regionami. Eliminowanie tych przeszkód staje się zaś następnie przedmiotem żmudnych negocjacji między blokami handlowymi. Na przykład porozumienie TTP (*Trans-Pacific Trade Partnership*), podpisane w 2016 roku przez Stany Zjednoczone i dwanaście krajów z regionu Pacyfiku, było negocjowane przez siedem lat. Z kolei negocjacje nad porozumieniem TTIP (*Trans-Atlantic Trade and Investment Partnership*) między Stanami Zjednoczonymi a Unią Europejską rozpoczęły się w 2013 roku i jeszcze w 2016 były wciąż dalekie od zakończenia.

8.2. Potrzeba międzynarodowej koordynacji polityki

8.2.1. Międzynarodowe dobra publiczne

Międzynarodowa koordynacja polityki gospodarczej polega na świadomym kształtowaniu polityki narodowej przez poszczególne państwa, mającym na celu uwzględnienie zachodzących między nimi współzależności (Wallich 1984). Jednym z istotnych następstw rozwoju handlu międzynarodowego i globalizacji jest obiektywna potrzeba wspólnego działania i koordynacji polityki gospodarczej w skali międzynarodowej. Rosnąca współzależność gospodarcza między suwerennymi państwami sprawia, że skutki polityki prowadzonej przez jedno państwo wywierają wpływ na sytuację gospodarczą w innych krajach. Oddziaływanie to występuje w postaci efektów zewnętrznych (*externalities*), które nie są na ogół uwzględniane w krajowym rachunku korzyści i kosztów związanych z podjęciem określonej decyzji przez rząd pojedynczego państwa. Ponadto, współzależność poszczególnych gospodarek stwarza zapotrzebowanie na pewne wspólne dobra publiczne o charakterze międzynarodowym, takie jak harmonizacja przepisów w skali międzynarodowej czy stabilność finansowa w warunkach otwartej gospodarki.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Okoliczności te skłaniają do tworzenia reguł współpracy i porozumień między krajami w celu ograniczania negatywnych skutków uzależnienia gospodarek krajowych od zjawisk, nad którymi polityka krajowa nie sprawuje kontroli, oraz zapewnienia dostępu do międzynarodowych dóbr publicznych. Porozumienia te doprowadziły do utworzenia specjalnych instytucji międzynarodowych, których zadaniem jest koordynacja polityki krajowej i ograniczanie różnych rodzajów ryzyka, jakie niesie globalizacja. Najważniejszą rolę spośród wszystkich tych instytucji odgrywa Międzynarodowy Fundusz Walutowy (IMF), Światowa Organizacja Handlu (WTO), Organizacja Współpracy i Rozwoju Gospodarczego (OECD), a także zinstytucjonalizowana współpraca w ramach grup G7, G8 i G20, w skład których wchodzą największe gospodarki świata.

Uzasadnienie celowości koordynacji polityki gospodarczej w skali międzynarodowej jest zbliżone do uzasadnienia w przypadku interwencji państwa w ramach pojedynczej gospodarki (Frenkel i in. 1988). Jak pamiętamy z pierwszego rozdziału, w pewnych szczególnych sytuacjach sam mechanizm rynkowy nie jest w stanie zapewnić optymalnych rozwiązań. W ramach gospodarki globalnej ta niesprawność rynku – podobnie jak w przypadku pojedynczego państwa – dotyczy dwóch zasadniczych kwestii. Pierwszą z nich jest potrzeba dostarczania dóbr publicznych. W przypadku pojedynczego kraju do dóbr takich należą dobra *lokalne*, takie jak edukacja, porządek publiczny czy obrona narodowa. Jak też już wiadomo, teoria ekonomii uczy, że w normalnych warunkach rynkowych podaż takich dóbr ukształtuje się poniżej społecznego optimum lub w ogóle się nie pojawi. Stąd wynika więc potrzeba dostarczania tych dóbr przez władzę publiczną w ramach polityki gospodarczej.

Z kolei na poziomie międzynarodowym istotne stają się dobra publiczne o zasięgu międzynarodowym i globalnym, takie jak ograniczanie emisji gazów cieplarnianych i ochrona klimatu, ochrona biologicznych zasobów morskich, zapewnienie stabilności finansowej w skali międzynarodowej, a także jednolite standardy prawne w wielu dziedzinach współpracy międzynarodowej, regulujące m.in. przepływy bezpośrednich inwestycji zagranicznych, nadzór bankowy czy rachunkowość korporacyjną. Dobra te mają charakter *globalnych* dóbr publicznych ze względu na wywoływane przez nie rozległe pozytywne efekty zewnętrzne, daleko wykraczające poza granice poszczególnych państw. Z uwagi na ich ponadnarodowy zasięg, dostarczanie tych dóbr na poziomie pojedynczego kraju jest mało skuteczne. Na przykład walka z ociepleniem klimatycznym nie przyniesie oczekiwanych rezultatów, jeśli ograniczy się do wysiłków podejmowanych przez jedno czy też kilka państw. Występowanie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

efektów zewnętrznych związanych z działalnością przemysłu i transportu (emisja gazów cieplarnianych) oraz możliwość „jazdy na gapę" (*free riding*) sprawiają, że tylko skoordynowana akcja wszystkich lub przynajmniej większości państw może doprowadzić do skutecznego obniżenia emisji gazów cieplarnianych w skali globalnej i zahamowania tendencji sprzyjających ocieplaniu się klimatu.

Innym przykładem globalnego dobra publicznego jest stabilność kursów walutowych. System płynnego kursu nie zapewnia stabilności ze względu na charakterystyczne dla rynku finansowego cechy, takie jak niepełna informacja, zwiększone ryzyko oraz związana z tym skłonność do paniki i zachowań stadnych (*herd behaviour*)[1]. Usztywnienie kursu danej waluty wobec innej jednostki pieniężnej jest sposobem ograniczania niepewności związanej z wahaniami kursu i ryzykiem kursowym. Nie zapewnia to jednak gwarancji stabilności, ponieważ na kurs mają też wpływ decyzje innych krajów. Można wykazać, że gdy mamy do czynienia z N niezależnych walut, może funkcjonować tylko $(N - 1)$ niezależnie ustalanych bilateralnych kursów walutowych. Kurs co najmniej jednej waluty będzie w tej sytuacji zależny od pozostałych kursów[2]. Możliwym rozwiązaniem w przypadku tej asymetrii pozostaje wspólne uzgodnienie kursów walutowych, Alternatywą jest rezygnacja z kontroli kursu. Koordynacja staje się więc potrzebna wówczas, gdy stabilność kursu jest czymś pożądanym i traktowanym jako wartość sama w sobie[3].

8.2.2. Efekty „rozlewania" (*spillovers*)

Drugim celem koordynacji jest ograniczanie ujemnych efektów zewnętrznych (*negative externalities*). Ich przykładem są negatywne konsekwencje, jakie indywidualne decyzje poszczególnych państw mogą nieść dla innych krajów na skutek występowania powiązań gospodarczych w skali międzynarodowej i mechanizmu przenoszenia efektów polityki gospodarczej danego kraju na inne państwa (*spillover effects* – zob. Frenkel i in. 1988). Na przykład luźna polityka pieniężna prowadzona przez duży kraj, taki jak Stany Zjednoczone, wywołuje osłabienie kursu waluty

[1] Płynne kursy walutowe w warunkach rosnącej współzależności poszczególnych gospodarek, spowodowanej globalizacją oraz integracją, okazują się dziś mniej skutecznym mechanizmem łagodzenia szoków, niż miało to miejsce w latach 70. i 80., ponieważ ich wahania są nasilone przez spekulacyjne przepływy kapitałów. Więcej na ten temat pisaliśmy w rozdziale 6.

[2] Załóżmy, że dysponujemy trzema walutami: złotym, euro i dolarem. Ustalenie kursu euro wobec dolara i kursu złotego wobec euro jednoznacznie wyznacza kurs złotego wobec dolara. Podobnie dzieje się w przypadku większej liczby walut.

[3] Jest to też argument na rzecz utworzenia unii walutowej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

tego kraju i aprecjację walut innych krajów, co może hamować ich eksport oraz spowalniać wzrost gospodarczy. Z kolei podniesienie stóp procentowych w USA pociąga za sobą odpływ kapitału z innych państw i wywołuje presję na deprecjację ich walut, co zmusza te kraje do zaostrzenia własnej polityki pieniężnej. W takiej sytuacji międzynarodowa koordynacja może zapobiegać wyniszczającej polityce konkurencyjnych dewaluacji, zwanej także polityką „zubożania sąsiada" (*beggar-thy--neighbour policy*). W ten sposób kraje usiłują zwiększyć własny dobrobyt kosztem dobrobytu sąsiadów. Przykładem koordynacji mającej na celu uniknięcie konkurencyjnych dewaluacji były obowiązujące w latach 70. i 80. okresowe porozumienia zawarte między najbardziej uprzemysłowionymi krajami świata (grupa G7), mające na celu stabilizację kursu dolara[4].

Jeszcze innym przykładem występowania negatywnych efektów zewnętrznych jest zaoferowanie przez dany kraj ulg podatkowych dla zagranicznych inwestorów w nadziei, że przyciągnie to do kraju kapitał zagraniczny i pozwoli przyspieszyć wzrost gospodarczy. Powoduje to jednak odpływ inwestycji z innych krajów i często zmusza je również do stosowania konkurencyjnych regulacji podatkowych. Zbyt agresywna rywalizacja między krajami w tym zakresie może prowadzić do powszechnego ograniczenia korzyści płynących z zagranicznych inwestycji, globalnego spadku dochodów podatkowych i pogorszenia warunków rozwoju rodzimych przedsiębiorstw. Jest to kolejny przykład polityki „zubożania sąsiada". Podobny efekt mogą przynieść indywidualne działania poszczególnych krajów mające na celu wspieranie własnego eksportu i ochronę przed importem. Potrafią one prowokować do podejmowania działań retorsyjnych ze strony innych krajów, co ostatecznie prowadzi do ograniczenia skali handlu międzynarodowego i obniżenia globalnego dobrobytu (Tobin 1988). W takiej sytuacji ponadnarodowa koordynacja polityki gospodarczej, w postaci wielostronnych porozumień zakazujących stosowania instrumentów naruszających konkurencję w handlu międzynarodowym (takich jak sprzedaż dumpingowa czy subwencjonowanie eksportu), przynosi korzyści wszystkim uczestnikom. Kolejnym przykładem występowania ujemnych efektów zewnętrznych jest stosowanie zbyt łagodnych reguł nadzoru bankowego w danym kraju w celu pozyskania dodatkowych inwestycji w sektorze bankowym, co w warunkach swobody przepływu kapitału zwiększa ryzyko kryzysu bankowego w tym oraz w innych państwach. Koordynacja polityki w postaci ustalenia

[4] Przykładami międzynarodowej koordynacji polityki kursowej były omówione w rozdziale 6 porozumienia Plaza Agreement i Louvre Agreement.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jednolitych reguł mikroostrożnościowych i takich samych wymogów nadzorczych może w takiej sytuacji wydatnie ograniczyć ryzyko kryzysu.

Szczególnym przykładem międzynarodowej koordynacji polityki gospodarczej są reguły fiskalne obowiązujące w państwach członkowskich UE. Ich przyjęcie było podyktowane koniecznością zapewnienia wysokiej wiarygodności i stabilności wspólnej waluty europejskiej. Jak wiemy z niedawnego doświadczenia, nieprzestrzeganie tych reguł doprowadziło w latach 2010–2012 do kryzysu zadłużenia suwerennego w niektórych krajach członkowskich strefy euro.

Teoretycznego uzasadnienia odnośnie do koordynacji narodowej polityki gospodarczej dostarcza strategiczna teoria gier. Najbardziej znaną ilustracją problemu jest tak zwany dylemat więźnia (*Prisoner's Dilemma*). Stanowi on przykład strategicznej współzależności, jaka występuje między dwoma podmiotami podejmującymi samodzielnie decyzje i która sprawia, że obydwu uczestnikom gry opłaca się porozumieć w celu uzgodnienia wyboru optymalnego zachowania (zob. ramka 8.1).

Ramka 8.1. Dylemat więźnia

Załóżmy, że dwóch więźniów – nazwijmy ich A i B – jest podejrzanych o popełnienie konkretnego przestępstwa. Żaden z nich nie przyznaje się do winy. Przyjmijmy również, że jeśli A oskarży B o popełnienie przestępstwa, to A zostanie uwolniony, a B skazany na 10 lat (i odwrotnie, gdy to B oskarży A), jeśli z kolei obaj wzajemnie się oskarżą, każdy z nich zostanie uznany za winnego i skazany na 5 lat, wreszcie gdy żaden z nich ani nie przyzna się do winy, ani nie oskarży drugiego więźnia, obaj zostaną skazani tylko na rok więzienia. Jak widać, los każdego z więźniów zależy nie tylko od jego własnego zachowania, lecz od od sposobu postępowania drugiego więźnia. W szczególności, sytuacja każdego z nich pozostaje uzależniona od tego, czy oskarżą się wzajemnie, czy też zdecydują się na współpracę. Takie położenie nazywa się strategiczną współzależnością. Wszystkie możliwe sytuacje decyzyjne wraz z macierzą wypłat zostały przedstawione w tabeli 8.1[5].

Wybór optymalnej strategii zależy od postępowania drugiego z więźniów. Przy braku możliwości porozumiewania się strategią preferowaną przez każdego z nich będzie wzajemne oskarżanie się, co prowadzi jednak do skazania ich obu na 5 lat. Jest to rozwiązanie niekooperacyjne, zwane też równowagą w sensie Nasha (*Nash*

[5] Na podstawie Tucker (1980) oraz Benassy-Quéré i in. (2010).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

equilibrium). Z tabeli 8.1 wyraźnie wynika jednak, że nie jest to rozwiązanie optymalne, ponieważ lepszą strategią pozostaje współpraca obu więźniów, zapewniająca wyrok rocznego pozbawienia wolności w przypadku każdego z nich. Stanowi ona rozwiązanie kooperacyjne. Przykład ten pokazuje, że decentralizacja decyzji – lub, inaczej mówiąc, w pełni „suwerenne" ich podejmowanie – może okazać się nieoptymalne i że koordynacja polityki w postaci wspólnego uzgadniania decyzji pozwala osiągać lepsze rezultaty.

Tabela 8.1. Dylemat więźnia

Strategia więźnia A	Strategia więźnia B	
	oskarżenie więźnia A	współpraca z więźniem A
Oskarżenie więźnia B	(–5; –5)	(0; –10)
Współpraca z więźniem B	(–10; 0)	(–1; –1)

Źródło: opracowanie własne.

Axelrod (1984) wykazał, że porozumienie między graczami nie musi być ani formalne, ani dosłowne. W sytuacji, gdy podejmowane decyzje są wielokrotnie powtarzane, gracze zaczną sami wybierać rozwiązanie kooperacyjne, ponieważ odstępstwo od niego pociąga za sobą nieuchronnie większe koszty. W systematycznie powtarzanych sytuacjach decyzyjnych nastąpi więc proces uczenia się, skłaniający graczy do współpracy i zapewniający równowagę kooperacyjną. Większość decyzji z zakresu polityki gospodarczej nie ma jednak charakteru powtarzalnego, a nawet jeśli taki charakter posiada, to z reguły towarzyszą temu zmieniające się warunki i okoliczności. W takich sytuacjach dylemat więźnia musi być rozwiązywany na drodze negocjacji i porozumień. Współcześnie stanowią one podstawę międzynarodowej współpracy suwerennych państw.

W praktyce często spotyka się przypadki występowania strategicznej współzależności między państwami, w których indywidualnie podejmowane decyzje mają charakter suboptymalny, co uzasadnia podejmowanie międzynarodowej współpracy. Dobrym przykładem może być zjawisko konkurencyjnej dewaluacji, której celem jest przerzucenie kosztów walki z recesją na sąsiednie kraje poprzez obniżenie kursu własnej waluty. Jest to klasyczny przykład polityki „zubażania sąsiada". Jeśli w ten sposób zachowują się wszystkie państwa, wówczas żadne z nich nie odniesie korzyści. Z taką właśnie sytuacją mieliśmy do czynienia w czasie Wielkiego

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Kryzysu z lat 30., kiedy poszczególne kraje bezskutecznie próbowały „wyeksportować" własne bezrobocie przy pomocy dewaluacji.

Ramka 8.2. Model symetrycznej koordynacji polityki*

Jest to model obejmujący dwa kraje o otwartych gospodarkach, powiązane sztywnym kursem walutowym, oraz resztę świata[6]. Każdy kraj dysponuje tylko jednym instrumentem polityki – polityką fiskalną, ponieważ polityka pieniężna w reżimie sztywnego kursu jest nieautonomiczna (zmiany podaży pieniądza są podporządkowane celowi kursowemu). W warunkach wystąpienia szoku zewnętrznego każdy kraj stara się zminimalizować odchylenia od równowagi wewnętrznej i zewnętrznej. Stan równowagi (nierównowagi) wewnętrznej w danym kraju mierzony jest wielkością różnicy między produktem rzeczywistym a produktem potencjalnym Y (jak pamiętamy, różnicę tę nazywa się luką podażową – *output gap*), z kolei stan równowagi (nierównowagi) zewnętrznej mierzony jest wielkością salda bilansu płatniczego B. Odpowiednie stany równowagi w przypadku drugiego kraju oznaczymy odpowiednio jako Y^* i B^*. Zarówno luka podażowa, jak i saldo bilansu płatniczego są funkcjami własnej polityki fiskalnej g oraz polityki fiskalnej sąsiada g^* (przy czym wzrost g i g^* oznacza zwiększenie wydatków rządowych).

$$Y = \alpha g + \beta g^* - s, \tag{8.1}$$

$$Y^* = \alpha g^* + \beta g - s, \tag{8.2}$$

gdzie $\alpha > \beta > 0$, podczas gdy s reprezentuje symetryczny zewnętrzny szok popytowy (negatywny impuls od reszty świata). Saldo bilansu płatniczego w każdym z krajów kształtuje się odpowiednio według równań:

$$B = \delta(g^* - g) - s, \tag{8.3}$$

$$B^* = \delta(g - g^*) - s, \tag{8.4}$$

gdzie $\delta > 0$. Gdy $s > 0$, popyt zewnętrzny maleje, co wywołuje spadek produkcji i pogorszenie bilansu płatniczego w obu krajach. W takiej sytuacji każde z państw może albo czekać, aż sąsiad zastosuje stymulację fiskalną w celu podtrzymania

[6] Zob. Benassy-Quere (2010), s. 122–123.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

koniunktury, albo samemu przeprowadzić taką stymulację. W przypadku drugiego z rozwiązań kraj wybierze taką politykę fiskalną g, aby zminimalizować określoną funkcję straty L:

$$L = wY^2 + B^2, \tag{8.5}$$

$$L^* = w^*Y^{*2} + B^{*2}, \tag{8.6}$$

gdzie w, $w^* > 0$ stanowią wagi przypisane jednostce luki popytowej względem jednostki salda bilansu płatniczego (gdy równowaga wewnętrzna jest bardziej istotna niż zewnętrzna, wówczas $w > 1$). Funkcje L i L^* mierzą łączne odchylenia od równowagi wewnętrznej i zewnętrznej[7].

Budując funkcję Lagrange'a dla równań (8.5) i (8.6) przy ograniczeniach opisanych wzorami (8.1)–(8.4), a następnie różniczkując i przyrównując do zera, otrzymujemy warunki pierwszego rzędu dla L, L^* = minimum. Pozwalają one określić optymalne decyzje fiskalne każdego z krajów. Gdy państwa będą podejmować decyzje niezależnie od siebie, optymalna polityka fiskalna w przypadku każdego z nich zostanie określona następująco (jest to równowaga niekooperacyjna w sensie Nasha):

$$g = \frac{\left(\delta^2 - w\alpha\beta\right)g^* + \left(w\alpha - \delta\right)s}{w\alpha^2 + \delta^2}, \tag{8.7}$$

$$g^* = \frac{\left(\delta^2 - w^*\alpha\beta\right)g + \left(w^*\alpha - \delta\right)s}{w^*\alpha^2 + \delta^2}. \tag{8.8}$$

Wielkość g i g^*, czyli reakcja każdego z państwa na egzogeniczny szok popytowy, gdzie $s > 0$, zależy od względnych wag przypisanych poziomowi produkcji krajowej względem równowagi zewnętrznej. Drugie wyrażenie umieszczone w nawiasie w liczniku ułamka (8.7) pokazuje, że jeśli $w\alpha > \delta$, pierwszy kraj zareaguje na szok zwiększeniem wydatków budżetowych w celu przywrócenia równowagi wewnętrznej, choć odbędzie się to również kosztem pogłębienia nierównowagi zewnętrznej (zob. równanie 8.3). Z kolei jeśli $w\alpha < \delta$, kraj zareaguje na szok zaostrzeniem polityki fiskalnej w celu przywrócenia równowagi zewnętrznej kosztem pogłębienia nierównowagi wewnętrznej. Podobnie zachowa się drugi kraj (w zależności od

[7] Jak łatwo zauważyć, funkcje L i L^* przyjmują wartość 0 (minimum) odpowiednio dla Y, $B = 0$, i Y^*, $B^* = 0$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

relacji $w^*\alpha$ i δ). Należy jednak zaznaczyć, że reakcja każdego z nich będzie zależeć też od reakcji sąsiada. Pierwsze wyrażenie znajdujące w nawiasie w liczniku ułamka (8.7) pokazuje, że reakcja polityki fiskalnej w pierwszym kraju zależy od wielkości g^* w drugim z państw oraz od znaku i wielkości wyrażenia ($\delta^2 - w\alpha\beta$). Podobnie, reakcja polityki fiskalnej w drugim kraju zależy od wielkości g w pierwszym kraju oraz od znaku i wielkości wyrażenia ($\delta^2 - w^*\alpha\beta$). Oznacza to, że, poszukując optymalnej strategii politycznej, każdy kraj powinien uwzględniać politykę drugiego państwa.

Przy założeniu, że $w = w^*$, można wykazać, że oba kraje prowadzą dokładnie taką samą politykę (ponieważ α, β i δ w przypadku obu państw są jednakowe), w związku z czym $g = g^*$. Przekształcając wcześniejsze równania (8.7) i (8.8), otrzymujemy:

$$g = g^* = \frac{(w\alpha - \delta)s}{w\alpha(\alpha - \beta)}. \tag{8.9}$$

Podstawiając $g = g^*$ do równań (8.1) – (8.4), uzyskujemy:

$$Y = Y^* = \frac{-\delta s}{w\alpha}, \tag{8.10}$$

$$B = B^* = -s, \tag{8.11}$$

oraz:

$$L = (1 + \delta^2/w\alpha^2)s^2. \tag{8.12}$$

Ostatnie równanie pokazuje wielkość funkcji straty w punkcie równowagi w sensie Nasha, czyli w warunkach braku koordynacji i współpracy. Jak widać, L jest nie tylko zawsze większe od zera, lecz również zawsze większe niż s^2, ponieważ wyrażenie zawarte w nawiasie w równaniu (8.12) pozostaje zawsze większe od jedności. Łatwo wykazać jednak, że gdyby kraje podjęły współpracę i uzgodniły taką politykę fiskalną, jaka pozwoliłaby im zachować równowagę wewnętrzną, a więc kiedy $Y = Y^* = 0$, wówczas w przypadku $g = g^*$ wartość funkcji straty wyniosłaby $L = L^* = s^2$. Ponieważ $s^2 < (1 + \delta^2/w\alpha^2)\, s$, to funkcja straty w warunkach koordynacji polityki przyjmuje mniejszą wartość niż wobec jej braku.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

8.3. Granice koordynacji

8.3.1. Zawodność koordynacji

W praktyce koordynacja polityki narodowej napotyka na rozmaite przeszkody i nigdy nie ma gwarancji, że w każdej sytuacji okaże się ona lepszym rozwiązaniem niż jej brak. W pewnych warunkach koordynacja może nie ułatwiać prowadzenia polityki gospodarczej, a niekiedy staje się nawet szkodliwa. Ekonomiści wysuwają szereg argumentów wskazujących na ograniczenia i potencjalnie negatywne skutki tego zjawiska. Jednym z najbardziej znanych krytyków koordynacji polityki w skali międzynarodowej jest amerykański ekonomista Martin Feldstein, były szef zespołu doradców ekonomicznych prezydenta Ronalda Reagana. Wśród wskazywanych przez niego zagrożeń związanych z koordynacją do najważniejszych należą: możliwość „strategicznego oszukiwania" przez poszczególne kraje, prawdopodobieństwo wystąpienia negatywnego wpływu koordynacji na mechanizm konkurencji rynkowej, ryzyko związane z częściową, odcinkową i niepełną koordynacją polityki oraz koszty samej koordynacji (Feldstein 1988).

Kraje mogą nie wywiązywać się z obowiązku współpracy, gdy pojawia się pokusa nadużycia (*moral hazard*) lub zjawisko „pasażera na gapę" (*free riding*). Sytuacja taka ma miejsce wówczas, gdy indywidualne korzyści wynikające z naruszenia wzajemnych ustaleń stają się większe niż oczekiwane zyski płynące z przestrzegania podjętych zobowiązań. Dobrym przykładem zastosowania niekorzystnej struktury bodźców jest system przestrzegania dyscypliny budżetowej w ramach Unii Gospodarczej i Walutowej. Zgodnie z zapisami Traktatu o Funkcjonowaniu UE oraz Paktu Stabilności i Wzrostu, deficyt budżetowy w państwach członkowskich nie może przekraczać 3% PKB, a dług publiczny 60% PKB. Niektóre kraje naruszały jednak te zobowiązania, ponieważ wiedziały, że ukaranie ich poprzez nałożenie sankcji finansowych jest mało prawdopodobne. Zarazem zdawały sobie sprawę, że zadłużanie się ponad ustalony limit nie jest czymś szczególnie groźnym, gdyż w ramach unii walutowej zaciąganie pożyczek przez pojedynczy kraj staje się tańsze[8]. Ponadto, kraje wysoko zadłużone mogą liczyć na pomoc pozostałych państw członkowskich, ponieważ bankructwo jednego z nich (*sovereign default*) pociągałoby za sobą duże straty kapitałowe wśród pozostałych członków unii walutowej.

[8] Wyjaśniamy to zjawisko w dalszej części rozdziału.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Co więcej, niektóre kraje ukrywały dane dotyczące prawdziwych rozmiarów deficytu i długu. W obu przypadkach koordynacja polityki narodowej okazała się nieskuteczna, a nadmierne zadłużenie doprowadziło UGW do głębokiego kryzysu. Zawodność koordynacji (*coordination failure*) może więc wynikać z silnie zniekształconej struktury bodźców lub z niedoskonałości instrumentów kontrolnych.

Trzeba też zauważyć, że międzynarodowa koordynacja polityki zogniskowana wokół niewłaściwie dobranych celów może przynieść rezultaty gorsze niż niezależna polityka poszczególnych krajów. Taka sytuacja ma miejsce wówczas, gdy cele są wzajemnie sprzeczne bądź zależności makroekonomiczne zostają źle rozpoznane. Przykładowo koordynacja polityki w zakresie bardzo ambitnego ograniczania emisji dwutlenku węgla (CO_2) w ramach UE może doprowadzić do znacznego obniżenia konkurencyjności krajów członkowskich wobec krajów trzecich, nie przynosząc jednocześnie zauważalnej poprawy w zakresie klimatu ze względu na brak współpracy ze strony krajów trzecich.

Podkreśla się także, że częściowa, niepełna koordynacja może prowadzić do gorszych rezultatów niż brak koordynacji. Na przykład próba koordynowania polityki pieniężnej w skali międzynarodowej bez jednoczesnego skutecznego koordynowania polityki fiskalnej może doprowadzić do poważnych zaburzeń w postaci wystąpienia nierównowagi zewnętrznej w poszczególnych krajach. Z taką sytuacją mieliśmy do czynienia w strefie euro, gdzie wspólna polityka pieniężna przyczyniła się do konwergencji rynkowych stóp procentowych, lecz rozbieżność polityki fiskalnej sprawiła, że w niektórych państwach wystąpił wysoki deficyt budżetowy, a w innych nadwyżka. Gdyby kraje te pozostały przy narodowej walucie, prawdopodobnie nie doszłoby do wzrostu zadłużenia na tak dużą skalę. Dodatkowym skutkiem wynikającym z braku koordynacji polityki fiskalnej w ramach unii walutowej może być powstanie wysokich deficytów płatniczych w jednych krajach i jednocześnie dużych nadwyżek wśród innych państw.

Trzeba też pamiętać, że koordynacja polityki krajowej pociąga za sobą określone koszty. Negocjacje wymagają odpowiedniego nakładu czasu i pieniędzy, a uzgodnione kompromisy muszą uzyskać poparcie polityczne w poszczególnych krajach. Co więcej, koszty te z reguły rosną wraz ze zwiększaniem liczby państw uczestniczących w negocjacjach i w miarę wzrostu liczby obszarów negocjacyjnych. Może się więc okazać, że w przypadku kompleksowych i wielostronnych negocjacji rezygnacja z koordynowania polityki i utrzymanie w tym zakresie niezależności jest w przypadku danego państwa bardziej opłacalne niż akceptacja marnego kompromisu.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

8.3.2. Koordynacja polityki w krajach rozwiniętych

Historia gospodarcza ostatnich kilku dziesięcioleci dostarcza wielu przykładów międzynarodowej koordynacji polityki. Zaburzenia na rynkach walutowych w latach 70. i rozpad systemu z Bretton Woods zmusiły kraje rozwinięte do podjęcia współpracy w celu ograniczenia gwałtownych wahań kursów walutowych. Forum tych ustaleń stanowiły spotkania na szczycie państw G7. Przykładem nieudanej koordynacji mogą być ustalenia szczytu w Bonn z lipca 1978 roku, kiedy kraje nadwyżkowe (Niemcy, Japonia) zobowiązały się do prowadzenia bardziej ekspansywnej polityki fiskalnej w celu przywrócenia równowagi międzynarodowej. Działania te zakończyły się fiaskiem głównie dlatego, że ich pozytywne skutki zniweczył drugi kryzys naftowy, do jakiego doszło kilka miesięcy później, co z kolei wywołało gwałtowny wzrost inflacji. Bardziej udane, choć jedynie na krótką metę, okazały się próby stabilizowania kursów walutowych podjęte w ramach porozumień zawartych w Nowego Jorku w października 1985 roku (*Plaza Agreement*) oraz w Paryżu w styczniu 1987 roku (*Louvres Agreement*).

Bardziej skuteczne były próby koordynowania polityki pieniężnej. Porozumienia zawarte między bankami centralnymi krajów G7 umożliwiały szybką reakcję w odpowiedzi na nieoczekiwane kryzysy, takie jak załamanie giełdowe na Wall Street z października 1987 roku, atak na World Trade Center we wrześniu 2001 roku oraz kryzys finansowy, do którego doszło we wrześniu 2008 roku. Po tym ostatnim kryzysie forum koordynacji polityki uległo rozszerzeniu z G7 i G8 do G20, choć efektywność podejmowanych działań w ramach tej poszerzonej formuły współpracy jest nadal dość ograniczona, głównie z powodu występowania dużej rozbieżności interesów między krajami rozwiniętymi i krajami rozwijającymi się.

Ramka 8.3. Grupa G20

Grupa Dwudziestu, czyli G20 (*The Group of Twenty*), stanowi formę regularnej międzynarodowej współpracy gospodarczej rządów i banków centralnych 20 największych gospodarek świata. W skład grupy G20 wchodzi 19 krajów (Argentyna, Australia, Brazylia, Kanada, Chiny, Francja, Niemcy, Indie, Indonezja, Włochy, Japonia, Meksyk, Federacja Rosyjska, Arabia Saudyjska, Republika Południowej Afryki, Korea Południowa, Turcja, Zjednoczone Królestwo i Stany Zjednoczone)

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

oraz Unia Europejska (reprezentowana przez Przewodniczącego Komisji Europejskiej i Prezesa Europejskiego Banku Centralnego)[9].

Grupa G20 powstała w 1999 roku jako forum spotkań ministrów finansów mające na celu omawianie kwestii związanych z polityką gospodarczą i jej wpływem na międzynarodową stabilność finansową. W 2008 roku w następstwie kryzysu finansowego, który ujawnił skalę i siłę występujących powiązań między poszczególnymi krajami, państwa zrzeszone w G20 postanowiły zintensyfikować koordynację polityki finansowej i zdecydowały o przeniesieniu poziomu współpracy na szczebel szefów państw i rządów. W latach 2008–2011 szczyty państw G20 odbywały się dwa razy do roku, a następnie ograniczono ich częstotliwość do jednego w roku. Gospodarzami kolejnych szczytów są poszczególne państwa grupy G20 na zasadzie rotacyjnej.

Lista gospodarek należących do G20 przedstawiona jest w tabeli 8.2. Kraje z tego ugrupowania reprezentują łącznie 85% światowej produkcji brutto, 80% światowego handlu międzynarodowego, a zamieszkuje je ponad 65% ludności całego świata. W rzeczywistości grupa G20 nie zrzesza dwudziestu największych gospodarek świata – przykładowo nie należy do niej Hiszpania (14 miejsce na liście największych gospodarek według MFW), Holandia (17 miejsce) i Szwajcaria (20 miejsce). Zamiast nich do G20 zostały zaproszone państwa o nieco mniejszych gospodarkach, reprezentujące jednak grupę krajów „wschodzących" (*emerging*) i rozwijających się z innych kontynentów niż Europa i Ameryka Północna. Znalazły się tu więc takie państwa jak Argentyna (24 miejsce) i Republika Południowej Afryki (33 miejsce – na podstawie danych za 2014 rok). Polska zajmuje 23. miejsce wśród największych gospodarek świata i nie należy do grupy G20.

Grupa G20 funkcjonuje równolegle z grupą G7, obejmującą siedem najbardziej rozwiniętych demokratycznych krajów świata. Forum współpracy G7 zostało utworzone w 1975 roku po kryzysie walutowym, który doprowadził do rozpadu systemu z Bretton Woods. W 1997 roku grupa G7, zasilona przez Rosję, przekształciła się w grupę G8[10].

[9] W spotkaniach grupy G20 oprócz stałych członków biorą też udział szefowie najważniejszych organizacji międzynarodowych, takich jak ONZ, Międzynarodowy Funduszu Walutowy, Bank Światowy, Światowa Organizacja Handlu (WTO) i OECD.

[10] W marcu 2014 roku Rosja została wykluczona z grupy G8 w następstwie zaanektowania Krymu.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tabela 8.2. Grupa G20 (dane za 2014 rok)

	Handel zagraniczny	PKB (nominalny, dane w mln USD)	PKB (PPP, dane w mln USD)	PKB/ mieszkańca (nominalny, dane w mln USD)	PKB/ mieszkańca (PPP, dane w mln USD)	HDI	Ludność (w tys.)
Argentyna	142,370	543,061	951,001	12,735	22,302	0,808	42,961
Australia	496,700	1,442,722	1,099,771	61,066	46,550	0,933	23,599
Brazylia	484,600	2,346,583	3,275,799	11,573	16,155	0,744	202,768
Kanada	947,200	1,785,387	1,595,975	50,304	44,967	0,902	35,467
Chiny	4,201,000	10,356,508	18,088,054	7,572	13,224	0,719	1,367,520
Francja	1,212,300	2,833,687	2,591,170	44,332	40,538	0,884	63,951
Niemcy	2,866,600	3,874,437	3,748,094	47,774	46,216	0,911	80,940
Indie	850,600	2,051,228	7,411,093	1,608	5,808	0,586	1,259,695
Indonezja	346,100	888,648	2,685,893	3,524	10,651	0,684	251,490
Włochy	948,600	2,147,744	2,135,359	35,335	35,131	0,872	59,960
Japonia	1,522,400	4,602,367	4,767,157	36,222	37,519	0,890	127,061
Korea Płd.	1,170,900	1,410,383	1,783,950	27,970	35,379	0,891	50,437
Meksyk	813,500	1,291,062	2,148,884	10,784	17,950	0,756	119,581
Rosja	844,200	1,860,598	3,576,841	12,718	24,449	0,778	146,300
Arabia Saudyjska	521,600	746,248	1,609,628	24,252	52,311	0,788	30,624
Republika Płd. Afryki	200,100	350,082	707,097	6,483	13,094	0,658	53,699
Turcja	417,000	798,332	1,514,859	10,381	19,698	0,759	77,324
Zjednoczone Królestwo	1,189,400	2,950,039	2,569,218	45,729	39,826	0,892	64,511
Stany Zjednoczone	3,944,000	17,348,075	17,348,075	54,370	54,370	0,914	318,523
Unia Europejska	4,485,000	18,527,116	18,640,411	36,645	36,869	0,876	505,570

Źródło: opracowanie własne na podstawie IMF 2015.

8.4. Teoria unii gospodarczej

Jeszcze sto lat temu praktycznie wszystkie kompetencje decyzyjne z zakresu polityki gospodarczej były skoncentrowane na szczeblu rządów centralnych państw narodowych. Od tamtej pory ten model rządzenia uległ dość zasadniczej zmianie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

– głównie ze względu na wpływ dwóch tendencji. Z jednej strony, zgodnie z zasadą subsydiarności, część kompetencji decyzyjnych była stopniowo przekazywana niższym szczeblom władzy krajowej w związku z dążeniem do wykorzystania przewagi informacyjnej, jaką dysponują władze regionalne, gdy przychodzi decydować o sprawach lokalnych. Zjawisko to było omówione w rozdziale 3 w kontekście dyskusji o finansach samorządów terytorialnych i federalizmie fiskalnym. Z drugiej strony część kompetencji została przekazana na szczebel transnarodowy lub ponadnarodowy w celu czerpania korzyści z międzynarodowej integracji gospodarczej i finansowej. W tym kontekście pojawia się pytanie o granice koordynacji polityki narodowej. Czy można w jakiś racjonalny sposób określić pożądany zakres przedmiotowy i terytorialny takiej integracji?

Odpowiedzi na to pytanie dostarcza teoria unii gospodarczych (*theory of economic unions*), która wskazuje na racjonalne przesłanki koordynacji polityki narodowej w postaci utworzenia międzynarodowych organizacji – unii gospodarczych – wyposażonych w określone kompetencje decyzyjne. Jak już stwierdziliśmy, koordynacja polityki – czyli porozumienie poszczególnych państw w sprawie *wspólnego* podejmowania decyzji w określonym zakresie – ma na celu ograniczanie negatywnych efektów zewnętrznych wynikających z polityki pojedynczych krajów (*spillover effects*) oraz dostarczanie pożądanych międzynarodowych dóbr publicznych (Alesina i in. 2001, 2005). Państwa, które wspólnie utworzą unię gospodarczą, będą oczekiwać określonych korzyści związanych z jej powstaniem (dostarczanie międzynarodowych dóbr publicznych i ograniczanie negatywnych efektów zewnętrznych). Koordynacja polityki pociąga za sobą również pewne koszty – zarówno w postaci kosztów negocjacji i uzgodnień, jak i kosztów przyjęcia rozwiązań suboptymalnych, czyli wynikających z zawieranych wspólnie kompromisów i odbiegających od wariantów polityki najlepszych z perspektywy pojedynczego państwa.

Czy można określić optymalne rozmiary unii gospodarczej? Inaczej mówiąc, czy możliwe jest ustalenie właściwej liczby państw członkowskich w przypadku danej organizacji gospodarczej? Aby odpowiedzieć na to pytanie, załóżmy, że na świecie występuje ogółem m krajów, które mają zróżnicowane preferencje, a część z nich rozważa utworzenie unii gospodarczej. Przyjmijmy też w celu uproszczenia, że poszczególne państwa są do siebie zbliżone pod względem rozmiaru gospodarki, lecz różnice w zakresie preferencji stają się tym większe, im dalej położone są od siebie poszczególne kraje, co oznacza, że sąsiadujące ze sobą państwa wykazują najbardziej zbliżone preferencje.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Na korzyści wynikające z integracji składają się międzynarodowe dobra publiczne dostarczane przez unię – takie jak harmonizacja prawa, jednolity rynek czy stabilność finansowa – oraz możliwość uniknięcia określonych negatywnych efektów zewnętrznych (*spillovers*). Ponieważ założyliśmy, że kraje są do siebie zbliżone pod względem rozmiaru gospodarki, korzyści będą rosnąć liniowo wraz ze wzrostem liczby krajów członkowskich n. Oznaczmy sumę korzyści wynikających z przyjęcia kraju i do unii jako $B_i = B$ dla wszystkich $i = 2, 3, \ldots, n$, gdzie n jest liczbą krajów członkowskich unii, przy czym $n \leq m$. Z kolei koszty związane z integracją obejmują straty związane z koniecznością częściowej rezygnacji z własnych preferencji przez poszczególne kraje oraz koszt niezbędnych negocjacji, których celem jest znalezienie wspólnego, kompromisowego rozwiązania. Koszty te stają się tym większe, im bardziej zróżnicowane są preferencje indywidualne oraz im więcej krajów wchodzi w skład unii. Jeśli uporządkujemy poszczególne państwa od najbliższych do tych położonych najdalej, to dołączanie do unii kolejnych krajów oznacza rosnący koszt związany zarówno z częściową rezygnacją z własnych preferencji, jak i z negocjowaniem wspólnych rozwiązań w coraz bardziej zróżnicowanej grupie państw. W rezultacie koszty C rosną wykładniczo wraz ze wzrostem liczby krajów członkowskich n. Oznaczmy sumę korzyści z integracji jako ΣB i zapiszmy:

$$\Sigma B = (n-1)B. \tag{8.13}$$

Z kolei sumę kosztów związanych z integracją wyraźmy w postaci równania:

$$C = Gn^{\alpha}, \tag{8.14}$$

gdzie G reprezentuje koszt utworzenia unii składającej się z pierwszych dwóch krajów (dla $n = 2$), a współczynnik $\alpha > 1$.

Aby ustalić optymalną liczbę krajów tworzących unię gospodarczą, należy zbudować funkcję użyteczności stanowiącą różnicę między korzyściami ΣB i kosztami C, a następnie znaleźć pochodną tej funkcji i przyrównać ją do zera. Funkcja użyteczności U przyjmuje postać:

$$U = \Sigma B - C = (n-1)B - Gn^{\alpha} = max. \tag{8.15}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Pochodna funkcji U wynosi:

$$U' = B - G\alpha \check{n}^{\alpha-1} = 0, \tag{8.16}$$

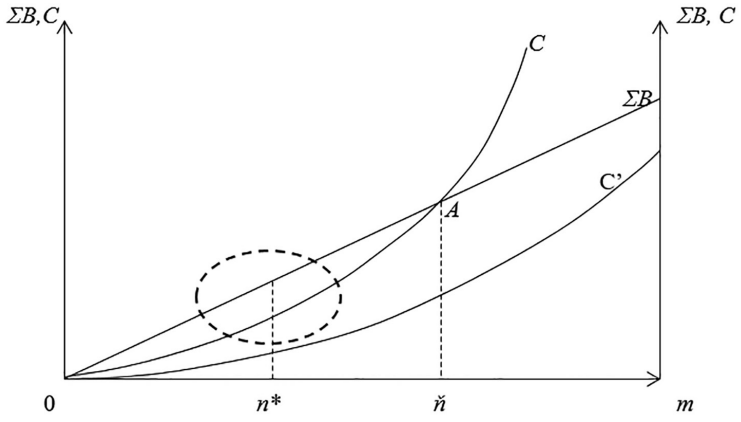
gdzie $\check{n}$ jest poszukiwaną optymalną liczbą krajów członkowskich wchodzących w skład unii gospodarczej. Rozwiązując wzór (8.16) dla $\check{n}$, otrzymujemy ostatecznie:

$$\check{n} = \sqrt[\alpha-1]{\frac{B}{G\alpha}}, \tag{8.17}$$

gdzie wykładnik pierwiastka $(\alpha - 1) > 0$. Równanie (8.17) pokazuje, że optymalna liczba państw członkowskich w unii $\check{n}$ staje się tym większa, im większa jest wartość B, oraz tym mniejsza, im większe są wartości G i α.

Zagadnienie dotyczące optymalnej liczby krajów członkowskich tworzących unię gospodarczą można przedstawić graficznie (rysunek 8.1). Na osi poziomej odłożono liczbę krajów ($i = 1, 2, \ldots, m$), a na osi pionowej korzyści B i koszty C. Krzywa C przedstawia koszt integracji jako funkcję liczby krajów członkowskich, a prosta ΣB odzwierciedla korzyść z integracji także jako funkcję liczby państw członkowskich. Oznacza to, że opłaca się przyjmować do unii nowe państwa aż do momentu wyrównania się kosztów i korzyści w punkcie A. Punkt ten wyznacza maksymalną liczbę krajów członkowskich $\check{n}$.

Rysunek 8.1. Maksymalna i optymalna liczba krajów członkowskich wchodzących w skład unii gospodarczej



Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Położenie punktu *A* zależy od nachylenia oraz kształtu krzywych kosztu i korzyści. Przy założeniu, że kraje są jednakowe pod względem wielkości i poziomu PKB/mieszkańca, nachylenie krzywej *ΣB* zależy od stopnia uzgodnionej koordynacji polityki narodowej. Im większy jest zakres koordynacji, tym wyższa staje się podaż międzynarodowych dóbr publicznych, a więc tym bardziej strome jest nachylenie linii *ΣB* i tym bardziej wzrasta wartość *ň* (przy danym położeniu krzywej *C*). Gdy uchylimy założenie o jednakowej wielkości krajów, krzywa *ΣB* staje się linią łamaną, lecz nadal rosnącą. Z kolei nachylenie krzywej *C* zależy od średniego kosztu integracji *G* oraz od poziomu współczynnika *α*. Im wyższa wartość *G* oraz im większe *α*, tym szybciej wznosi się krzywa *C* i tym niższa staje się wartość *ň* (przy danym położeniu krzywej *ΣB*). Łączna suma korzyści wynikających z integracji w punkcie równowagi *A* równa jest polu zawierającemu się między krzywą *ΣB* i krzywą *C*.

Niektóre unie gospodarcze mogą przynosić korzyści wszystkim krajom świata. Koszt takiej integracji ilustruje krzywa *C'*, która znajduje się poniżej krzywej *ΣB* na całej swej długości, a więc odzwierciedla wartość wydatków ponoszonych przez wszystkie *m* kraje świata. Przykładem opłacalnej koordynacji polityki na szczeblu globalnym jest polityka klimatyczna, ponieważ korzyści z tego tytułu czerpią wszystkie kraje. Niewiele jest jednak unii o charakterze globalnym, przeważają bowiem przede wszystkim ugrupowania regionalne. Z rysunku 8.1 można wywnioskować, że integracja regionalna obejmująca ograniczoną liczbę krajów znajdujących się w danym regionie powinna być bardziej popularna ze względu proporcjonalnie największe korzyści zapewniane ograniczonej liczbie krajów uczestniczących, co pokazuje obszar zaznaczony linią przerywaną. Optymalną liczbą państw członkowskich w tej sytuacji jest n^*, ponieważ dla tej wartości *n* odległość między linia *ΣB* i linią *C* jest największa. Liczba państw wchodzących w skład takich ugrupowań jest oczywiście niższa niż *ň*. Przyjęty układ krzywych na rysunku 8.1 pokazuje, że zwiększanie liczby krajów członkowskich do poziomu $n^* < ň$ zapewnia rosnące korzyści netto (*ΣB* – *C*), podczas gdy zwiększanie liczby państw członkowskich powyżej liczby n^* aż do liczby *ň* przynosi już malejące korzyści. Natomiast dalszy wzrost liczby państw członkowskich powyżej liczby *ň* przynosi straty netto.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

8.5. Teoria unii walutowej

8.5.1. Podstawy teorii optymalnych obszarów walutowych

Szczególną odmianą unii gospodarczych jest unia walutowa. Na gruncie teoretycznym ekonomiczne argumenty na rzecz rezygnacji z własnej waluty i przyjęcia wspólnej waluty zostały sformułowane na początku lat 60. w ramach teorii tzw. optymalnych obszarów walutowych (*Optimum Currency Areas* – OCA). Podwaliny „klasycznej" teorii OCA stworzył kanadyjski ekonomista Robert Mundell, który za swoje prace, między innymi w tej dziedzinie, otrzymał w 1999 roku nagrodę Nobla (zob. m.in. Mundell 1961, 1969). W krótkim artykule opublikowanym w 1961 roku Mundell zauważył, że krajowi opłaca się zrezygnować z własnej waluty, gdy korzyści w postaci wyeliminowania kosztów transakcyjnych związanych z koniecznością wymiany walut w operacjach międzynarodowych są większe niż koszty wynikające z utraty samodzielności w zakresie polityki pieniężnej i kursowej w przypadku rezygnacji z własnej waluty. Obszar, na którym opłaca się posługiwać wspólną walutą (albo walutami narodowymi powiązanymi ze sobą sztywnymi kursami walutowymi), Mundell nazwał optymalnym obszarem walutowym (zob. ramka 8.4).

Ekonomista ten wskazał jednocześnie na dwa ograniczenia integracji walutowej. Po pierwsze, gdy gospodarki krajów partnerskich są strukturalnie zróżnicowane (heterogeniczne), stają się podatne na wstrząsy asymetryczne[11]. Wówczas, gdy rynki pracy w tych krajach nie zapewniają dostatecznie wysokiej mobilności siły roboczej i elastyczności płac, państwa te powinny zachować raczej własną walutę i niezależną politykę pieniężną, aby móc skutecznie reagować na występujące szoki. Po drugie, poszczególne kraje mogą mieć różne preferencje co do (względnego) poziomu inflacji i bezrobocia – w szczególności zaś mogą się różnić pod względem stopnia tolerancji dla wyższej inflacji. To zróżnicowanie preferencji często utrudnia lub wręcz uniemożliwia prowadzenie jednolitej polityki pieniężnej, wspólnej dla tych krajów. W takich warunkach nie opłaca się tworzyć unii walutowej, lecz lepiej jest pozostać raczej przy własnej walucie. Ponadto, waluty poszczególnych

[11] Wstrząs – lub szok – asymetryczny to wstrząs popytowy lub podażowy, który uderza w daną gospodarkę, lecz nie ma wpływu na inne gospodarki wchodzące w skład unii walutowej. Więcej na temat szoków asymetrycznych piszemy w punkcie 8.5.3.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

państw nie powinny być w takiej sytuacji trwale związane z walutami partnerów. Wręcz przeciwnie, kraje strukturalnie heterogeniczne oraz te różniące się preferencjami powinny stosować elastyczne systemy kursowe. Zapewniają one bowiem możliwość dokonania zmian kursu waluty, tak aby móc dostosować gospodarkę do szoków zewnętrznych bez ryzyka recesji i spadku zatrudnienia.

Ramka 8.4. Optymalny obszar walutowy według Roberta Mundella

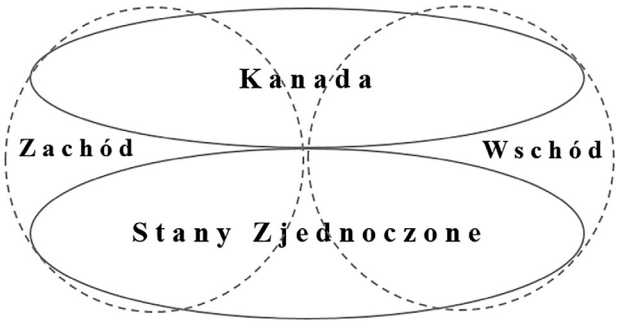
Robert Mundell (1961) definiuje obszar walutowy jako terytorium, na którym obowiązują stałe kursy walutowe. Może tu funkcjonować albo tylko jedna waluta, albo wiele walut narodowych powiązanych ze sobą stałymi kursami. Przykładem obszaru walutowego o wielu walutach były kraje Unii Europejskiej przed powstaniem unii walutowej, kiedy waluty narodowe były powiązane sztywnymi kursami w ramach Europejskiego Systemu Walutowego (*European Monetary System*). Zwiększenie (lub ograniczanie) łącznej podaży pieniądza w takim obszarze wymaga z reguły skoordynowanej akcji wszystkich banków centralnych, co bywa jednak politycznie trudne do przeprowadzenia, ponieważ poszczególne kraje mogą wykazywać różne preferencje w stosunku do inflacji i bezrobocia. W takiej sytuacji – gdy kursy są stałe, a zarazem problemy związane z niezbędną koordynacją polityki są trudne lub wręcz niemożliwe do przezwyciężenia – lepiej byłoby posiadać więc zwyczajnie jedną walutę i jeden bank centralny.

Załóżmy, że w ramach tak zdefiniowanego obszaru walutowego funkcjonują dwa kraje A i B powiązane ze sobą sztywnymi kursami, i że w w każdym z nich występuje początkowo pełne zatrudnienie i zrównoważony bilans płatniczy. Przyjmijmy następnie, że z jakiegoś powodu dochodzi do przesunięcia popytu z kraju B do kraju A. Jeśli kraj A chce uniknąć niepożądanego wzrostu cen, powinien ograniczyć podaż pieniądza, z kolei jeśli kraj B stara się uniknąć recesji i bezrobocia, powinien zwiększyć podaż pieniądza. Te różnokierunkowe zmiany podaży pieniądza spowodują presję na aprecjację waluty kraju A i deprecjację waluty kraju B. W ramach systemu sztywnego kursu zastosowanie jakichkolwiek zmian jest jednak niemożliwe. Jeśli ceny w kraju A nie wzrosną dzięki restrykcyjnej polityce pieniężnej, popyt przesunie się na produkty kraju A, a cały ciężar dostosowania spadnie na kraj B, w którym musi wzrosnąć bezrobocie. Jeżeli kraj A nie zaostrzy natomiast polityki pieniężnej i dopuści do wzrostu cen, popyt powróci do kraju B, który dzięki temu utrzyma pełne zatrudnienie, choć kosztem wzrostu inflacji w kraju A.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Tak więc albo będziemy mieli do czynienia z inflacją w kraju A, albo z bezrobociem w kraju B. W takiej sytuacji uzasadnione staje się odejście od systemu kursu sztywnego i wprowadzenie zmiennego kursu walutowego. Wówczas dewaluacja (deprecjacja) kursu w kraju B lub rewaluacja (aprecjacja) kursu w kraju A zmienia relacje cen (*terms of trade*) w wymianie między krajami i pozwala przywrócić równowagę w obu państwach. Oznacza to, że w takiej sytuacji kraje A i B nie powinny tworzyć wspólnego obszaru walutowego.

Rysunek 8.2. Przykład Mundella: obszary walutowe nie muszą pokrywać się z granicami państw



Źródło: opracowanie własne.

Mundell wskazuje, że granice obszarów walutowych wcale nie muszą pokrywać się z granicami krajów. Aby lepiej wyjaśnić tę różnicę, ekonomista rozpatruje sytuację dwóch krajów (USA i Kanady), które posiadają odrębne waluty narodowe. Mundell zakłada, że każde z tych państw składa się z dwóch zasadniczych regionów ekonomicznych – Wschodu i Zachodu – w których wytwarza się odpowiednio samochody i drewno. Sytuacja ta została przedstawiona na rysunku 8.2, gdzie linie ciągłe stanowią granice państw, a linie przerywane to granice regionów. Załóżmy następnie, że kursy walutowe są zmienne i że z jakiegoś powodu spada popyt na samochody. Bezpośrednim skutkiem takiego stanu rzeczy jest wzrost bezrobocia na Wschodzie i wzrost inflacji na Zachodzie w obu krajach. Banki centralne każdego z państw mogą w tej sytuacji zwiększyć podaż pieniądza, co pozwoli ograniczyć bezrobocie na Wschodzie, lecz kosztem dalszego wzrostu inflacji na Zachodzie. Z drugiej strony mogą też zmniejszyć podaż pieniądza, dzięki czemu uda się ograniczyć inflację na Zachodzie, ale kosztem dalszego wzrostu bezrobocia na Wschodzie. Podobnie jak w poprzednim przypadku nie da się w takiej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

sytuacji wyeliminować zarówno bezrobocia, jak i inflacji jednocześnie. Co ważniejsze, również zmiana kursu walutowego nie przywróci naruszonej równowagi między regionami, ponieważ jest ona w stanie przywracać równowagę jedynie między państwami, a nie między regionami. Oznacza, to, że w danym przypadku zarówno USA, jak i Kanada nie są obszarami walutowymi opisanymi w definicji Mundella. Możliwym rozwiązaniem tego problemu byłoby utworzenie dwóch obszarów walutowych – jeden obejmowałby wschodnie regiony USA i Kanady, a drugi zachodnie regiony obu państw (rysunek 8.2). Wówczas deprecjacja dolara „wschodniego" względem „zachodniego" pozwoliłaby przywrócić równowagę w obu regionach i w obu krajach.

Główny wniosek wynikający z teorii Mundella jest więc taki, że optymalny obszar walutowy wyznaczają rzeczywiste powiązania ekonomiczno-handlowe na danym terytorium gospodarczym, a nie polityczno-administracyjne granice państw. To spostrzeżenie torowało drogę do uzasadnienia dla utworzenia unii walutowej obejmującej większą liczbę państw ściśle powiązanych gospodarczo.

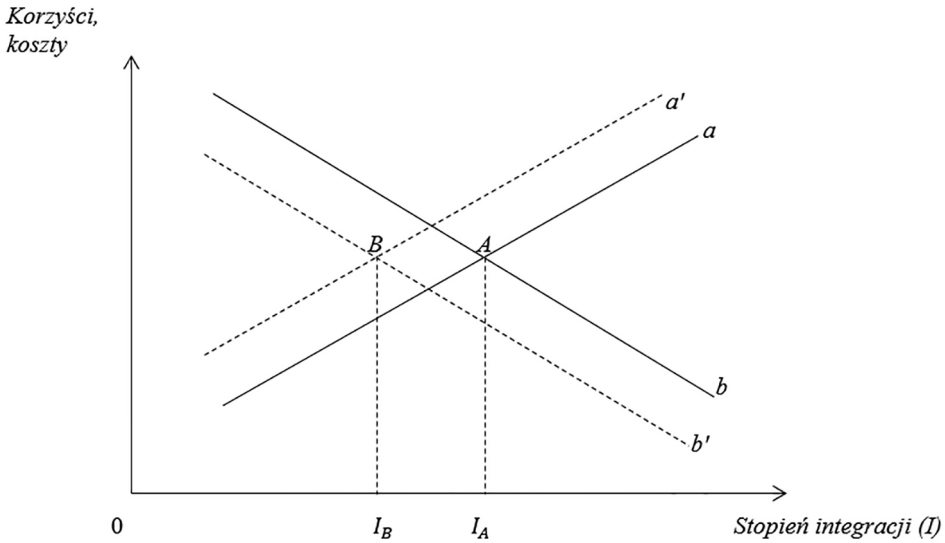
Rysunek 8.3 przedstawia istotę klasycznej teorii Mundella. Na osi poziomej odłożono stopień integracji gospodarki danego kraju z gospodarką innego państwa, z którym dany kraj chce utworzyć unię walutową. Alternatywnie oś pozioma może mierzyć stopień integracji gospodarki danego kraju z gospodarką istniejącej unii walutowej, do której dany kraj chce dołączyć. Stopień integracji może być mierzony udziałem sumy eksportu dóbr i usług danego kraju do drugiego państwa (lub do unii walutowej) oraz importu dóbr i usług danego kraju z drugiego państwa (lub z unii walutowej) w całości wymiany zagranicznej danego kraju. Alternatywnie stopień integracji daje się też zmierzyć udziałem sumy eksportu i importu w PKB danego kraju. Na osi pionowej odłożono koszty i korzyści wynikające z posiadania własnej waluty jako funkcję stopnia integracji danego kraju z unią walutową. Linia *a* jest linią kosztu posiadania własnej waluty jako funkcji stopnia integracji – ma ona charakter rosnący, ponieważ w miarę wzrostu stopnia integracji wzrasta też skala wzajemnej wymiany i rosną koszty transakcyjne związane z koniecznością wymiany walut. Linia *b* stanowi zaś linię korzyści wynikających z posiadania własnej waluty jako funkcji stopnia integracji – ma ona charakter opadający, ponieważ w miarę wzrostu stopnia integracji maleje korzyść z posiadania odrębnej polityki pieniężnej i kursowej (rośnie synchronizacja cyklu koniunkturalnego i spada prawdopodobieństwo szoków asymetrycznych).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Linie *a* i *b* przecinają się w punkcie *A*, przy stopniu integracji równym I_A. Gdy stopień integracji jest mniejszy niż I_A, korzyści wynikające z posiadania własnej waluty są większe niż koszty z tym związane, a krajowi opłaca się wówczas utrzymywać własną walutę. Gdy stopień integracji jest większy niż I_A, korzyści płynące z posiadania własnej waluty są mniejsze niż koszty z tym związane, a krajowi opłaca się z kolei dołączyć do unii walutowej.

Istotny wkład do „klasycznej" teorii OCA wnieśli w kolejnych latach ekonomiści amerykańscy, między innymi Ronald McKinnon i Peter Kenen. Pierwszy z nich (1963) zauważył, że im bardziej gospodarka jest otwarta na wymianę z otoczeniem międzynarodowym, tym mniej sztywne są płace i ceny w reakcji na zmiany kursu oraz tym bardziej uzasadnione staje się zastosowanie sztywnego kursu walutowego między krajem i zagranicą (lub przyjęcie wspólnej waluty). McKinnon wskazał również, że im bardziej otwarte są gospodarki względem siebie, tym bardziej zacieśnia się przestrzeń sprzyjająca występowaniu szoków asymetrycznych i tym większa staje się synchronizacja zmian produkcji w tych krajach. Ekonomista zwrócił też uwagę na to, że w małej otwartej gospodarce płynny kurs stanowi źródło niestabilności oraz że słabością takiej gospodarki jest ograniczona płynność rynku finansowego.

Rysunek 8.3. Korzyści i koszty związane z posiadaniem własnej waluty



Źródło: opracowanie własne.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Rozwijając teorię Mundella, Kenen (1969) badał warunki występowania szoków asymetrycznych (*asymmetric shocks, idiosyncratic shocks*) i wykazał, że kraje o bardziej zdywersyfikowanej strukturze produkcji i eksportu – zarówno pod względem towarowym, jak i geograficznym – są lepszymi kandydatami do zawarcia unii walutowej, ponieważ szoki asymetryczne w poszczególnych sektorach równoważą się i częściowo znoszą nawzajem. Kenen uznał to za argument przemawiający na rzecz stosowania sztywnych kursów między krajami rozwiniętymi, o bardziej zróżnicowanych strukturach produkcji i wymiany. Z kolei nadmierne uzależnienie od jednego rynku lub od danego rodzaju eksportu zwiększa potencjał szoków asymetrycznych, co nakazuje krajom gospodarczo monokulturowym utrzymywanie zmiennych kursów walutowych, aby kompensować wstrząsy w *terms of trade*. Dlatego też Kenen zalecał krajom rozwijającym się stosowanie elastycznych reżimów kursowych. Wskazał on także, że im wyższy jest udział handlu wewnątrzgałęziowego (*intra-industry trade*) w wymianie prowadzonej między krajami partnerskimi, tym silniejsze stają się powiązania produkcyjne, a także wzrasta synchronizacja przebiegu cyklu koniunkturalnego w tych krajach oraz maleje potencjał szoków asymetrycznych.

8.5.2. Dalszy rozwój teorii optymalnych obszarów walutowych

Robert Mundell po kilku latach rozwinął i zmodyfikował swoją teorię, uzupełniając ją o analizę sytuacji na rynku aktywów finansowych i uwzględniając występowanie dynamicznych oczekiwań. W pracy opublikowanej w 1973 roku (Mundell 1973) wziął pod uwagę fakt, że sektor prywatny utrzymuje część swoich zasobów w postaci aktywów finansowych, takich jak pieniądz lub papiery wartościowe emitowane przez własny rząd, i wskazał, że aktywa te straciłyby na wartości, gdyby kraj będący ich emitentem został dotknięty szokiem asymetrycznym (aż do bankructwa w skrajnym przypadku). Chcąc zdywersyfikować ryzyko, sektor prywatny danego kraju będzie więc dążył do zdywersyfikowania struktury swojego portfela aktywów, uzupełniając go o aktywa emitowane przez inne kraje i nominowane w innych walutach. Lecz czynnikiem ograniczającym pożądaną dywersyfikację portfela jest ryzyko zmian kursu walutowego, które zaburzają wycenę aktywów. Skłania ono, zwłaszcza pośredników finansowych (banki), do bilansowania aktywów i pasywów w jednej walucie. W ten jednak sposób dążenie do obniżenia ryzyka utraty wartości aktywów z tytułu szoku asymetrycznego stoi w konflikcie z dążeniem do obniżenia ryzyka kursowego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Na tej podstawie Mundell wywnioskował, że im bardziej dany kraj jest podatny na szoki asymetryczne, tym bardziej uzasadnione staje się wejście tego kraju do obszaru wspólnej waluty. Wynika to stąd, że sektor prywatny w tym kraju może się wówczas zabezpieczyć przed ryzykiem utraty wartości aktywów z tytułu (bardziej prawdopodobnego) szoku asymetrycznego, skupując aktywa emitowane przez pozostałe kraje i nie ponosząc z tego tytułu ryzyka kursowego w ramach unii walutowej. W takich warunkach utrzymywanie własnej waluty utrudnia lub uniemożliwia pożądaną dywersyfikację portfela aktywów i zwiększa premie za ryzyko kursowe w krajowych stopach procentowych. Wspólna waluta pozwala więc sektorowi prywatnemu w sposób bardziej zrównoważony rozłożyć ryzyko na posiadanych aktywach (*private risk sharing mechanism*). Warto zauważyć, że wniosek ten jest w pewnym stopniu sprzeczny z twierdzeniem, który Mundell sformułował w swojej „klasycznej" teorii z 1961 roku, kiedy dowodził, że kraje podatne na szoki asymetryczne powinny zachować swoją odrębną walutę[12].

Obserwacja Mundella w zakresie dywersyfikacji ryzyka i jego wpływu na poziom stóp procentowych okazała się szczególnie istotna przy analizie celowości włączenia do unii walutowej krajów o małych otwartych gospodarkach, których waluty narodowe obciążone były dużym ryzykiem kursowym. Od tego momentu ekonomiści zaczęli zdawać sobie sprawę, że ważną korzyścią wynikającą z wejścia do unii walutowej może być spadek stóp procentowych na skutek eliminacji ryzyka kursowego (De Grauwe 2000). Dotyczy to w szczególności gospodarek „doganiających" (*catching up*), znajdujących się na niższym poziomie rozwoju gospodarczego niż kraje rozwinięte.

Teoria optymalnych obszarów walutowych przestała być intelektualną ciekawostką i stała się przedmiotem intensywnych badań, gdy w 1989 roku ukazał się raport przewodniczącego Komisji Europejskiej Jacques'a Delorsa o potrzebie utworzenia w Europie unii walutowej[13], stanowiący później podstawę traktatu z Maastricht. W wyniku tych badań w okresie następnej dekady powstała nowa, rozwinięta i wzbogacona teoria optymalnych obszarów walutowych, którą można nazwać – w odróżnieniu od teorii „klasycznej" – teorią współczesną. Za jedną z jej istotnych cech uznano uchylenie założenia o statycznych oczekiwaniach

[12] Omówienie i porównanie obu teorii Mundella (Mundell I i Mundell II) można znaleźć w: McKinnon (2004).

[13] *Report on the Economic and Monetary Union in the European Communities*, Committee for the Economic and Monetary Union, Brussels, April 1989.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

uczestników rynku i uwzględnienie wpływu utworzenia unii walutowej na rynki finansowe i kapitałowe (temat ten omawia m.in. Borowski 2012).

Eichengreen i Bayoumi (1993) badali warunki występowania szoków asymetrycznych w Europie na tle gospodarki Stanów Zjednoczonych. Stwierdzili, że wprawdzie warunki optymalności unii walutowej nie są w Unii Europejskiej spełnione w stopniu takim, w jakim ma to miejsce w USA, ale mimo to potencjał szoków asymetrycznych w przypadku krajów kontynentalnej Europy jest niewielki. Frankel i Rose (1998) na podstawie danych empirycznych potwierdzili hipotezę McKinnona mówiącą o tym, że wraz z pogłębianiem integracji nasila się synchronizacja cyklu koniunkturalnego między krajami, i na tej podstawie sformułowali tezę o endogeniczności kryteriów optymalnego obszaru walutowego. Zgodnie z tym twierdzeniem, nawet jeśli dany kraj w momencie wejścia do unii walutowej nie spełnia klasycznych kryteriów członkostwa, takich jak odpowiednio wysoki poziom integracji, to na skutek włączenia się w obszar wspólnej waluty wzajemna wymiana ulega intensyfikacji i poziom integracji się podnosi. W rezultacie dany kraj spełnia kryterium członkostwa *ex post*. Frankel i Rose (1998) pokazali też, że związki handlowe między krajami stają się tym silniejsze, im wyższa jest korelacja w zakresie działalności gospodarczej krajów, i odwrotnie (zob. też Frankel 1999).

Pojawiły się też opinie przeciwstawne. Na przykład amerykański ekonomista Paul Krugman dowodził – odwrotnie niż Frankel i Rose – że im wyższy jest stopień integracji, tym niższa staje się synchronizacja cyklu koniunkturalnego, ponieważ dzięki integracji znikają bariery w handlu i następuje regionalna koncentracja produkcji (zgodnie z teorią aglomeracji)[14], która pogłębia różnice między strukturami gospodarczymi poszczególnych krajów. Przykładem może być produkcja samochodów, która w warunkach integracji ma tendencję do koncentrowania się wśród niewielkiej liczby krajów, głównie tam, gdzie koszty wytwarzania są niższe. Zgodnie z tą teorią wzrost stopnia integracji zwiększa różnice strukturalne między krajami i podnosi potencjał szoków asymetrycznych (Krugman 1993, Krugman i Venables 1993). Wyniki badań empirycznych dość jednoznacznie wskazywały jednak na to, że integracja zwiększa konwergencję, głównie dzięki rozwojowi handlu wewnątrzgałęziowego, a hipoteza Krugmana nie znalazła w tym przypadku

[14] Teoria aglomeracji, rozwinięta przez Paula Krugmana w ramach tzw. Nowej Geografii Ekonomicznej (*New Economic Geography*), przewiduje, że produkcja będzie lokować się w pobliżu głównych centrów popytu, a popyt umiejscowi się w pobliżu głównych centrów produkcji (Krugman 1991). Aglomeracja wynika głównie z rosnących przychodów ze skali produkcji oraz dodatnich efektów zewnętrznych związanych z lokalizacją działalności produkcyjnej i konsumpcji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

żadnego potwierdzenia (European Commission 1990). W kierunku silniejszej konwergencji na skutek integracji miał także oddziaływać wzrost roli usług w tworzeniu PKB, których dostarczanie z natury rzeczy ma charakter zdecentralizowany (De Grauwe 2000, s. 37).

Teoria unii walutowej przeszła głęboką ewolucję w okresie ostatniego półwiecza. Głównym impulsem do dokonywania kolejnych jej modyfikacji były przemiany zachodzące w gospodarce światowej w ostatnich kilku dekadach, w tym postępująca globalizacja oraz integracja rynków, zwłaszcza finansowych. O ile w latach 60. tylko nieliczne, najbardziej rozwinięte kraje posiadały w pełni wymienialne waluty, a przepływy kapitału były związane głównie z finansowaniem transakcji eksportu oraz importu dóbr i usług, o tyle obecnie waluty krajowe większości państw są w pełni wymienialne, a jedynie około 3% przepływów finansowych na świecie ma związek z zakupami dóbr i usług – cała reszta to przepływy czysto finansowe, z reguły o charakterze spekulacyjnym.

Poszerzeniu zakresu wymienialności walut narodowych i wzrostowi skali transgranicznych przepływów kapitału towarzyszyły dwie ważne konsekwencje odnośnie do bilansu kosztów i korzyści wynikających z wejścia do unii walutowej. Po pierwsze, koszty transakcyjne związane z wymianą walut wzrosły na skutek szybkiego rozwoju handlu międzynarodowego i rynków finansowych. Posiadanie własnej waluty okazało się w związku z tym zasadniczo bardziej kosztowne, co zwiększyło atrakcyjność wejścia do unii walutowej. Po drugie, autonomia poszczególnych krajów w prowadzeniu niezależnej polityki pieniężnej uległa znacznemu ograniczeniu na skutek swobody przepływów kapitałowych, których efektem są egzogeniczne zmiany w podaży pieniądza krajowego. Utrata autonomii w polityce pieniężnej na skutek wejścia kraju do unii walutowej stała się w związku z tym zasadniczo mniej kosztowna. Po trzecie, w wyniku poszerzenia zakresu wymienialności walut narodowych oraz w obliczu zwiększenia się skali transgranicznych przepływów kapitału, wiele krajów odeszło od stosowania systemu kursów stałych i wprowadziło kursy płynne. W tej sytuacji utrata autonomii w zakresie polityki kursowej w związku z przyłączeniem się danego państwa do unii walutowej staje się mniej kosztowna, ponieważ w warunkach kursu płynnego kraje w istocie nie prowadzą samodzielnej polityki kursowej.

Co więcej, oprócz tradycyjnych argumentów przemawiających na rzecz integracji walutowej pojawiły się nowe przesłanki, takie jak spadek stóp procentowych, dostęp do większych zasobów kapitału i wzrost płynności rynku w wyniku połączenia krajowych rynków kapitałowych oraz wzmocnienie stabilności makroekonomicznej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

krajów mniej rozwiniętych i mniej zatem stabilnych po wejściu do obszaru stabilnej i wiarygodnej waluty. W tym kontekście Alesina i Barro (2002) wskazują, że przyjęcie przez dany kraj waluty kraju o niskiej oraz stabilnej inflacji pozwala krajowi przyjmującemu zaimportować „wiarygodność" monetarną i obniżyć w ten sposób premie za ryzyko na własnym rynku pieniężnym i kapitałowym. Ostatnie lata pokazały także, że uczestnictwo w unii walutowej może być czynnikiem umacniającym międzynarodową pozycję polityczną kraju uczestniczącego[15].

Ogólnie rzecz biorąc, rozwój teorii optymalnych obszarów walutowych w okresie ostatnich dwudziestu lat doprowadził do istotnej modyfikacji oryginalnej koncepcji Roberta Mundella. Zmiany, jakie dokonały się w tej dziedzinie, poszły w kilku kierunkach. Po pierwsze, na skutek szybkiego postępu globalizacji oraz integracji międzynarodowej, a także związanego z tym wzrostu współzależności gospodarek krajowych, szoki asymetryczne nie stanowią już takiego zagrożenia w przypadku stabilności, jakiego obawiał się Mundell, przynajmniej w odniesieniu do krajów rozwiniętych i w ramach regionalnych ugrupowań integracyjnych. Po drugie, swoboda transgranicznych przepływów kapitałowych oraz integracja rynków finansowych silnie ograniczają możliwość skutecznego oddziaływania na koniunkturę wewnętrzną – produkcję i zatrudnienie – przy pomocy instrumentów polityki pieniężnej i kursowej, zwłaszcza w małych otwartych gospodarkach (Alesina i Barro 2002; Alesina, Barro i Tenreyro 2002). Po trzecie wreszcie, różnice na poziomie preferencji wobec inflacji i bezrobocia okazały się mniej istotne, niż zakładał Mundell, czego pośrednim dowodem jest zaakceptowanie przez wszystkie kraje UGW stabilności cen jako głównego celu polityki pieniężnej.

8.5.3. Wstrząsy asymetryczne

Wstrząsy egzogeniczne mogą mieć charakter symetryczny lub asymetryczny. Pierwsze z nich dotykają kraje unii walutowej w jednakowy sposób, a zatem wymagają takiej samej reakcji w zakresie polityki gospodarczej. Przykładowo wstrząsem symetrycznym jest spadek popytu we wszystkich krajach unii walutowej na skutek cyklicznej recesji. W takiej sytuacji pożądaną reakcją ze strony wszystkich państw

[15] W ramach Unii Europejskiej państwa członkowskie strefy euro tworzą grupę bardziej zintegrowaną i współzależną. Uzgadniają one wiele wspólnych decyzji podczas spotkań szefów państw i rządów w ramach szczytów państw strefy euro (*Eurosummits*) posiedzeń ministrów finansów w ramach tzw. Eurogrupy (*Eurogroup*). Wypracowane wówczas wspólne stanowisko jest na ogół później przyjmowane przez pozostałe kraje członkowskie UE.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

jest poluzowanie polityki pieniężnej poprzez obniżenie stóp procentowych lub poluzowanie polityki fiskalnej poprzez zwiększenie wydatków rządowych (lub obniżenie podatków). W warunkach szoków symetrycznych nie jest więc potrzebna odrębna narodowa polityka pieniężna – wystarczy jedna wspólna polityka monetarna. Istnieje więc w tym przypadku uzasadnienie dla uwspólnotowienia polityki gospodarczej. Z kolei wstrząsy asymetryczne oddziałują na kraje unii walutowej w sposób zasadniczo zróżnicowany. Na jedne państwa wywierają dodatni wpływ, na inne zaś ujemny, lub też dotykają tylko jednego kraju bez wpływu na pozostałe.

W ramach teorii unii walutowej zakłada się, że wstrząsy symetryczne mogą być łagodzone przy pomocy zmian wspólnej polityki pieniężnej, ponieważ oddziałują one jednakowo – ekspansywnie lub restrykcyjnie – na wszystkie kraje członkowskie unii walutowej. Jeśli chodzi natomiast o wstrząsy asymetryczne, wspólna polityka pieniężna nie jest w stanie oddziaływać w sposób zróżnicowany na poszczególne kraje. Dlatego łagodzenie ich skutków powinno być zasadniczo zadaniem polityki fiskalnej, ponieważ poszczególne kraje zachowują nad nią pełną kontrolę i mogą reagować w sposób zróżnicowany, w zależności od konkretnej sytuacji. Poniżej zajmiemy się wstrząsami asymetrycznymi, które stanowią główne zagrożenie wobec stabilizacji makroekonomicznej w unii walutowej.

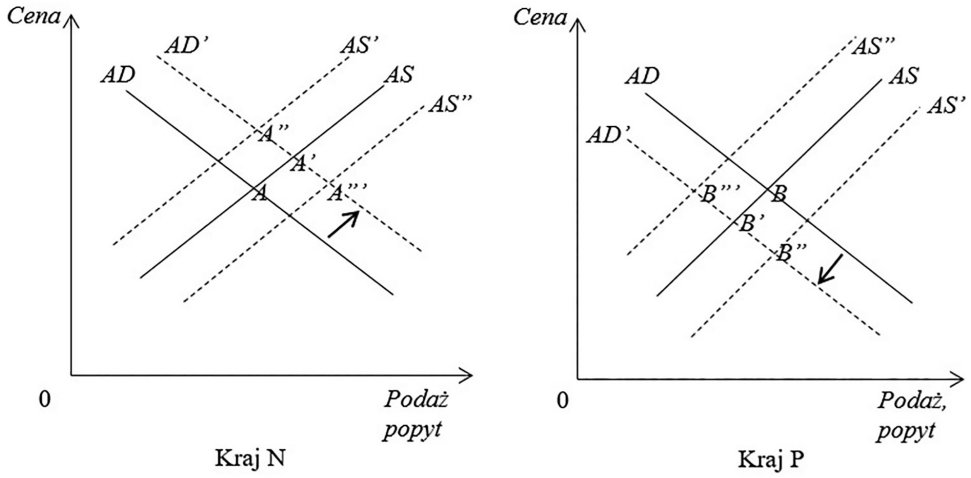
Pojęcie asymetrycznych wstrząsów – lub szoków – popytowych (*asymmetric shocks, idiosyncratic shocks*) pojawia się już w pierwszej wersji teorii optymalnych obszarów walutowych. Możliwość ich wystąpienia oraz konieczność niwelowania związanych z tym skutków była dla Mundella argumentem przemawiającym na rzecz utrzymywania płynnych kursów walutowych. Rozwijając teorię Mundella, Kenen (1969) badał warunki wystąpienia wstrząsów asymetrycznych i wykazał, że kraje o bardziej zdywersyfikowanej strukturze gospodarczej są lepszymi kandydatami do unii walutowej, ponieważ wstrząsy asymetryczne w poszczególnych sektorach są często różnokierunkowe, a ich skutki mają tendencję do równoważenia się i częściowo znoszą się nawzajem. To samo odnosi się do struktury geograficznej handlu zagranicznego – nadmierna zależność od jednego rynku zwiększa potencjał wstrząsów asymetrycznych, a bardziej zróżnicowana struktura terytorialna handlu pozwala na równoważenie szoków mających źródło w poszczególnych rynkach. Kenen uznał to zresztą za argument przemawiający na rzecz wprowadzenia sztywnych kursów między krajami rozwiniętymi, o bardziej zróżnicowanych strukturach produkcji i wymiany, z kolei krajom rozwijającym się zalecał bardziej elastyczne reżimy kursowe. Kenen wskazał także, że im wyższy jest udział handlu wewnątrzgałęziowego (*intra-industry trade*) w wymianie prowadzonej między

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

krajami partnerskimi, tym silniejsze stają się powiązania produkcyjne, a także wzrasta synchronizacja przebiegu cyklu koniunkturalnego w tych krajach i maleje potencjał wstrząsów asymetrycznych.

Mechanizm i skutki oddziaływania wstrząsu asymetrycznego można ukazać za pomocą prostego modelu AS-AD, przedstawionego na rysunku 8.4 (zob. m.in. Borowski 2003, 2011). Przyjmijmy, że dwa kraje – kraj N (na przykład Niemcy) i kraj P (na przykład Polska) – prowadzą między sobą wymianę i że w obu państwach wielkości produkcji i popytu zapewniające równowagę rynkową są wyznaczone przez punkty przecięcia krzywych łącznej podaży *AS* i łącznego popytu *AD*. Załóżmy obecnie, że z jakiegoś powodu popyt w kraju N na import z kraju P spada, podczas gdy rośnie popyt na produkty kraju N. Oznacza to, że kraj P jest dotknięty ujemnym wstrząsem popytowym, który ma charakter asymetryczny. Krzywa popytu w kraju P przesuwa się lewo, a krzywa popytu w kraju N przemieszcza się w prawo.

Rysunek 8.4. Efekty wstrząsu asymetrycznego



Źródło: opracowanie własne.

Jakie będą skutki wstrząsu asymetrycznego z perspektywy obu krajów? Przesunięcie popytu w kraju N z dóbr importowanych z kraju P na dobra krajowe wywołuje spadek popytu na produkty kraju P oraz wzrost produkcji w kraju N. Ceny, produkcja i zatrudnienie w kraju P wówczas spadną (krzywa *AS* przesuwa się w dół do *AS*'), podczas gdy w kraju N nastąpi ich wzrost (krzywa *AS* przesuwa się w górę do *AS*'). Nowymi punktami równowagi będą odpowiednio punkty *D*' i *C*'. Ponieważ

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ogólny poziom wydatków krajowych w obu krajach nie ulega jednak zmianie, w handlu wzajemnym pojawi się nierównowaga, co uwidoczni się w nadwyżce bilansu obrotów bieżących w kraju N i w deficycie rachunku bieżącego w kraju P (salda rachunku bieżącego nie zostały uwzględnione na rysunku)[16]. W kraju P wystąpi deficyt bieżący, ponieważ krajowa produkcja spadnie przy danym poziomie wydatków, podczas gdy w kraju N wystąpi nadwyżka, ponieważ przy danym poziomie wydatków produkcja rośnie ze względu na przesunięcie popytu z dóbr importowanych na dobra krajowe. W rezultacie w kraju N pojawi się inflacja, a w kraju P bezrobocie.

Nierównowaga płatnicza wywoła presję na waluty obu krajów – presję aprecjacyjną na walutę kraju N i presję deprecjacyjną na walutę kraju P. W systemie kursów płynnych nastąpią dostosowania w poziomie kursu, lecz w systemie kursów stałych nierównowaga spowoduje odpływ rezerw walutowych z kraju P do kraju N. Taki stan rzeczy nie może jednak trwać zbyt długo i wymaga zmiany polityki.

Jakie są możliwe reakcje w zakresie polityki w obu krajach wobec wstrząsu asymetrycznego? Gdy kraj P posiada własną walutę, może poluzować politykę pieniężną. Obniżenie stóp procentowych spowoduje wzrost popytu krajowego, produkcji i zatrudnienia, wobec czego krzywa *AD* przesunie się w prawo, na przykład do początkowego położenia. Z kolei podniesienie stóp procentowych w kraju N ograniczy popyt i zahamuje inflację, a krzywa popytu przesunie się w lewo. W obu państwach sytuacja powraca do równowagi w punktach *A* i *B*. Kraje mogą też wykorzystać politykę kursową. Gdy obowiązuje stały kurs walutowy, dewaluacja waluty kraju P wobec waluty kraju N zmniejsza popyt w N i zwiększa popyt w P. W efekcie funkcje podaży przesuwają się odpowiednio w górę i w dół (do położenia *AS'*, a nowymi punktami równowagi są odpowiednio *A"* i *B"* (pamiętajmy, że realna deprecjacja waluty P działa jak obniżka płac realnych, która zwiększa kosztową konkurencyjność kraju P).

Jeśli jednak kraje N i P wchodzą w skład unii walutowej, nie mogą wykorzystać wspólnej polityki pieniężnej w celu zniwelowania skutków wstrząsu asymetrycznego, ponieważ przywrócenie równowagi wymagałoby podniesienia stóp procentowych w kraju N i obniżenia ich w kraju P. Analogicznie nie można wykorzystać też polityki kursowej, ponieważ kraj N wymaga aprecjacji, a kraj P deprecjacji kursu. Dlatego też, zgodnie z sugestią Mundella, kraje wchodzące do unii walutowej muszą

[16] Przypomnijmy, że saldo bilansu obrotów bieżących (*current account*) jest równe różnicy między krajową produkcją *Y* i krajową absorpcją *A* (obejmującą wszystkie wydatki), czyli $CA = Y - A$.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

dysponować alternatywnymi instrumentami pozwalającymi chronić gospodarkę przed skutkami wstrząsu asymetrycznego. Do narzędzi tych zalicza się dużą mobilność siły roboczej, wysoką elastyczność płac oraz swobodę w polityce fiskalnej.

Duża mobilność siły roboczej pozwala uniknąć bezrobocia w następstwie wstrząsu asymetrycznego. Pracownicy, którzy utracili zatrudnienie na skutek spadku popytu na produkty kraju P, mogą migrować do kraju N, gdzie popyt na pracę rośnie. W takiej sytuacji krzywe podaży ulegają odpowiednio przesunięciu w prawo (kraj N) i w lewo (kraj P) do położenia *AS*". Nowymi punktami równowagi są *A*"' i *B*"'[17].

Duża elastyczność wynagrodzeń pozwala również uniknąć bezrobocia pojawiającego się w następstwie wstrząsu asymetrycznego. Obniżka płac nominalnych (i realnych) w kraju P poprawia jego konkurencyjność względem kraju N, i przy danych cenach produkcja rośnie. Krzywe podaży ulegają przesunięciu odpowiednio w lewo (kraj N) i w prawo (kraj P), tak jak w przypadku dewaluacji kursu (do AS'), a nowymi punktami równowagi są *A*" i *B*".

Wreszcie też ekspansja fiskalna w kraju P pozwala zwiększyć popyt na jego produkty. Dzięki temu krzywa popytu przesuwa się w prawo, w stronę początkowego położenia *AD* oraz początkowego punktu równowagi *B*. Podobny efekt ma symetryczne zacieśnienie polityki fiskalnej w kraju N.

Źródłem wstrząsów asymetrycznych są zasadniczo wydarzenia zewnętrzne, mające charakter egzogeniczny względem gospodarki. Czasem mogą one jednak wynikać z krajowej polityki fiskalnej, np. z nałożenia podatków lub wzrostu wydatków, a także zmian instytucjonalnych, takich jak zmiany prawa gospodarczego, zmiany podatkowe czy przepisy regulujące dostęp do rynku krajowego (Bąk i Maciejewski 2015). Z kolei krajowa polityka pieniężna lub kursowa nie może być oczywiście źródłem szoku w ramach unii walutowej, ponieważ jest zastąpiona wspólną polityką. Dlatego też uważa się, że wejście do unii walutowej zasadniczo obniża potencjał występowania szoków asymetrycznych. Nie tylko zanika odrębna polityka pieniężno-kursowa, lecz także ma miejsce dalsza harmonizacja prawa, rozwija się wzajemna wymiana gospodarcza, a w rezultacie synchronizacja przebiegu cyklu koniunkturalnego w poszczególnych krajach członkowskich unii walutowej ulega nasileniu. To ostatnie zjawisko – pogłębianie integracji na skutek przyjęcia wspólnej waluty – nazywa się endogenicznością konwergencji gospodarczej.

[17] Na znaczenie mobilności siły roboczej w neutralizowaniu skutków szoków asymetrycznych wskazywali przed Mundellem także Meade (1957) i Scitovski (1958).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Szczególną odmianą asymetrycznego wstrząsu popytowego jest szok inflacyjny. Wywołuje on dwa przeciwstawne efekty. Pierwszy z nich oddziałuje za pośrednictwem kanału realnych stóp procentowych: spadek realnych stóp procentowych na skutek wzrostu inflacji wywołuje dalszy wzrost popytu oraz inflacji. Jest to więc efekt destabilizujący. Drugi efekt oddziałuje za pośrednictwem kanału realnego kursu walutowego. Inflacja wywołuje aprecjację kursu realnego, co pogarsza konkurencyjność i pociąga za sobą spadek eksportu, popytu oraz inflacji. Jest to więc efekt stabilizujący. Ostateczny efekt wstrząsu inflacyjnego zależy więc od względnej siły obu efektów składowych.

Skutki wstrząsu asymetrycznego mogą być łagodzone poprzez dywersyfikację aktywów (zgodnie z „drugą" teorią Mundella), czyli otwarcie rynku kapitałowego na transakcje międzynarodowe, dochody od czynników produkcji otrzymywane z zagranicy oraz możliwość zaciągania pożyczek na krajowym rynku kredytowym. Dywersyfikacja aktywów poprzez transgraniczne transakcje zakupu i sprzedaży aktywów na otwartym rynku kapitałowym pozwala sektorowi prywatnemu na bardziej zrównoważone rozłożenie ryzyka związanego ze skutkami wstrząsów asymetrycznych. Otwarty rynek kapitałowy sprzyja budowaniu bezpieczniejszych portfeli aktywów w sektorze prywatnym. Wstrząs asymetryczny, który uderza w dany kraj, wywołuje spadek wartości krajowych aktywów, lecz nie ma wpływu na aktywa zagraniczne. Pozwala to chronić wartość całego portfela i łagodzi efekt majątkowego wstrząsu. Drugim narzędziem służącym obniżaniu ryzyka są dochody od inwestycji zagranicznych lub dochody z tytułu wynagrodzeń otrzymywanych z zagranicy. Nie ulegają one obniżeniu na skutek wstrząsu asymetrycznego i łagodzą łączne zmiany poziomu dochodów w kraju. Wreszcie trzecim mechanizmem jest rynek kredytów. Możliwość zaciągania kredytu, który amortyzuje zmiany dochodów wywołane przez wstrząs asymetryczny, pozwala wygładzać konsumpcję w czasie. Zarówno otwarty rynek kapitałowy, jak i rynek kredytowy stanowią ważne rynkowe mechanizmy niwelowania skutków szoku asymetrycznego w unii walutowej. Określamy je mianem prywatnych mechanizmów podziału ryzyka (*private risk sharing*). W sektorze publicznym podobną rolę odgrywają bezzwrotne lub zwrotne transfery fiskalne dokonywane między krajami bądź regionami (*public risk sharing*), co wymaga jednak stworzenia w pewnym zakresie wspólnego budżetu (von Hagen 2000, Werning i Fahri 2012).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

8.6. Potrzeba koordynacji polityki gospodarczej w unii walutowej

8.6.1. Transmisja impulsów polityki fiskalnej w unii walutowej: kanał handlu i kanał stóp procentowych (*fiscal policy spillovers*)

Jak wspomniano wyżej, wstrząsy egzogeniczne mogą mieć charakter symetryczny lub asymetryczny. W przypadku unii walutowej główny instrument reagowania na szoki symetryczne stanowi wspólna polityka pieniężna. Jest ona narzędziem polityki stabilizacyjnej stosowanej w trakcie przebiegu cyklu koniunkturalnego w skali całej unii walutowej. Z kolei głównym instrumentem reagowania na szoki asymetryczne ma pozostawać polityka fiskalna, realizowana przez poszczególne kraje unii. Jak zaznaczyliśmy jednak już wcześniej, indywidualna polityka fiskalna poszczególnych państw w określony sposób wpływa też na sytuację pozostałych krajów, co uzasadnia potrzebę koordynowania polityki.

Jak już wiadomo, literatura ekonomiczna notuje dwa główne powody, które wskazują na korzyści wynikające z zastosowania międzynarodowej koordynacji polityki krajowej. Pierwszym z nich jest potrzeba dostarczania określonych międzynarodowych (lub globalnych) dóbr publicznych, takich jak stabilność finansowa czy harmonizacja prawa. Drugim powodem jest zaś to, że polityka poszczególnych krajów może generować negatywne efekty zewnętrzne w stosunku do innych państw (*spillovers*). W przypadku unii walutowej wskazuje się na inne jeszcze argumenty przemawiające na rzecz koordynacji, takie jak ryzyko „pasażera na gapę" czy pokusa nadużycia, które omawiamy w następnym punkcie. Warto również zauważyć, że koordynacja polityki gospodarczej w unii walutowej zwiększa wiarygodność polityki narodowej poszczególnych państw, ponieważ eliminuje ryzyko wystąpienia konfliktu między określonymi krajami oraz ogranicza skutki rozproszenia efektów tej polityki w przypadku braku ich koordynacji.

Aby zrozumieć, na czym polegają efekty zewnętrzne w ramach unii walutowej, rozpatrzymy sytuację, w której jeden lub kilka krajów członkowskich podejmuje decyzję o zastosowaniu polityki fiskalnej w celu eliminacji skutków negatywnego szoku egzogenicznego. Wówczas, niezależnie od tego czy mamy do czynienia z szokiem symetrycznym czy asymetrycznym, ekspansja fiskalna w jednym kraju zwiększa – poprzez wzrost importu tego kraju – także popyt w innych krajach unii. Im większa część ekspansji fiskalnej danego kraju trafia do innych państw, tym mniej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

skuteczna staje się polityka fiskalna tego kraju w pobudzaniu własnej gospodarki. Ekspansja fiskalna wiąże się zarazem z kosztem w postaci wzrostu deficytu budżetowego finansowanego długiem. W takiej sytuacji może wystąpić konflikt interesów między krajami członkowskimi unii walutowej. Państwo stosujące ekspansję fiskalną jest zainteresowane tym, aby dodatkowy popyt zasilał w jak największym stopniu do jego własną gospodarkę, podczas gdy pozostałe kraje unii zabiegają o to, żeby popyt ten „wyciekał" za granicę i w jak największym stopniu trafiał do ich gospodarek (efekt handlowy).

Zauważmy następnie, że ekspansja fiskalna w jednym kraju unii – wobec braku akomodacji pieniężnej – powoduje w efekcie wzrost stóp procentowych w całej unii. Pozostałe kraje ponoszą więc dodatkowe koszty związane z ekspansją fiskalną jednego kraju. Koszty te stają się tym wyższe, im bardziej rośnie skala wzrostu wydatków publicznych jednego kraju w relacji do wydatków publicznych w całej unii. Mamy więc w takiej sytuacji do czynienia z dwoma efektami zewnętrznymi – efektem handlowym i efektem stóp procentowych. O ile efekt handlowy wynikający z ekspansji fiskalnej w jednym kraju unii jest korzystny dla pozostałych państw członkowskich, o tyle efekt stóp procentowych okazuje się w tym przypadku niekorzystny.

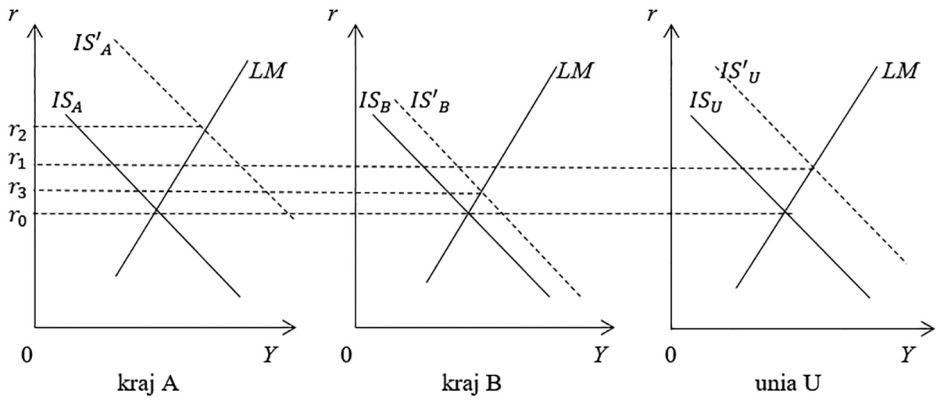
Załóżmy, że unię walutową tworzą dwa kraje o bardzo podobnych gospodarkach, A i B, oraz że początkowo w każdym z nich budżet jest zrównoważony, a stopa bezrobocia zbliżona jest do stopy naturalnej. Oba kraje są powiązane ze sobą handlem zagranicznym oraz wspólną polityką pieniężną, co oznacza, że stopy procentowe są w nich jednakowe. Przyjmijmy następnie, że, chcąc ograniczyć bezrobocie, kraj A decyduje się zwiększyć wydatki rządowe finansowane deficytem. Skutki tej polityki wygodnie jest zilustrować za pomocą krzywych *IS* i *LM*. Trzy wykresy przedstawione na rysunku 8.5 reprezentują odpowiednio kraje A i B oraz całą unię U. Początkowy stan równowagi wyznaczają punkty przecięcia się krzywych *IS* z krzywymi *LM*, opisujące wspólną dla całej unii stopę procentową równowagi na poziomie r_0.

W pierwszym etapie wzrost zagregowanego popytu w kraju A przesuwa krzywą IS_A w prawo (do położenia IS'_A) – w efekcie zwiększa się poziom produkcji i zatrudnienia w kraju A. Ekspansja fiskalna w tym kraju wywołuje jednak także efekty zewnętrzne oddziałujące na funkcjonowanie kraju B i całej unii. Ponieważ państwa są powiązane przez handel, wzrost wydatków rządowych w kraju A wywołuje także wzrost eksportu z kraju B do kraju A. Dzięki temu krzywa IS_B w kraju B przesuwa się do położenia IS'_B, choć zmiana położenia jest w tym przypadku

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

oczywiście mniejsza niż w kraju A. Kraj B zyskuje więc dzięki efektowi handlowemu. Wreszcie, krzywa IS_U w przypadku całej unii również przesuwa się w prawo do IS'_U. Przemieszczenie się krzywej IS_U stanowi wypadkową przesunięcia krzywych *IS* w kraju A oraz w kraju B, ponieważ jest jednocześnie mniejsze niż w kraju A, lecz większe niż w kraju B.

Rysunek 8.5. Efekty zewnętrzne polityki fiskalnej (*fiscal spillovers*) stosowanej w ramach unii walutowej: efekt handlowy i efekt stóp procentowych



Źródło: opracowanie własne.

Czy to oznacza, że kraj B zawsze będzie odnosił korzyści z ekspansji fiskalnej kraju A? Niekoniecznie. Pamiętajmy, że oba kraje tworzą unię walutową, w ramach której stopy procentowe w obu państwach są zawsze jednakowe. Zakładając, że zastosowana będzie neutralna polityka pieniężna (brak akomodacji), nowy poziom stóp procentowych w całej unii zostanie wyznaczony w punkcie przecięcia się nowej krzywej IS'_U z krzywą *LM*. Jak widać, wzrost wydatków rządowych w kraju A wywoła więc jednocześnie wzrost stóp procentowych w całej unii walutowej z początkowego poziomu r_0 do poziomu r_1. Wzrost ten w przypadku kraju A okazuje się jednak wyraźnie niższy niż potencjalny wzrost stóp procentowych, jaki nastąpiłby, gdyby kraj A nie należał do unii walutowej – byłby to poziom r_2. Kraj A zyskuje zatem dzięki temu, że skutki ekspansji fiskalnej w zakresie stóp procentowych rozkładają się na wszystkie państwa członkowskie, dzięki czemu faktyczny wzrost stóp procentowych jest niższy. Jednocześnie wzrost stóp procentowych w kraju B jest wyraźnie wyższy niż wzrost stóp procentowych, jaki nastąpiłby w sytuacji, gdyby kraj B nie był członkiem unii walutowej – byłby to poziom r_3. Kraj B notuje zatem straty,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ponieważ stopy procentowe rosną w większym stopniu, niż gdyby kraj B nie należał do unii walutowej – poziom stóp procentowych wzrasta bowiem z poziomu r_0 do r_1 zamiast z r_0 do r_3, przy czym $r_1 > r_3$.

W bardzo podobny sposób można opisać efekty zewnętrzne wynikające z ekspansywnej polityki fiskalnej, będącej reakcją na wstrząs symetryczny dotykający jednakowo wszystkie kraje unii (zob. ramka 8.5).

Ramka 8.5. Polityka fiskalna i polityka pieniężna w unii walutowej w warunkach szoku symetrycznego

Przyjmijmy, że unię walutową tworzą dwa identyczne kraje A i B, które są powiązane ze sobą handlem zagranicznym oraz wspólną polityką pieniężną (stopy procentowe pozostają jednakowe). Załóżmy następnie, że oba kraje są dotknięte symetrycznym szokiem zewnętrznym w postaci spadku popytu na produkty całej unii. Skutki tego zjawiska oraz możliwe reakcje w zakresie polityki zostały ukazane za pomocą krzywych *IS* i *LM* na rysunku 8.6, który jest repliką rysunku 8.5[18]. W pierwszym etapie spadek popytu przesuwa w jednakowym stopniu krzywe *IS* w krajach A i B – a zatem też w całej unii – w lewo w dół, wywołując spadek produkcji i obniżenie stóp procentowych (z r_0 do r_1). Załóżmy teraz, że kraj A zareaguje na szok popytowy odpowiednim wzrostem wydatków rządowych. Wówczas krzywa *IS* w tym kraju powróci do pierwotnego położenia[19]. Jednocześnie ekspansja fiskalna w kraju A wywołuje wzrost importu z kraju B, co oznacza, że krzywa *IS* w kraju B także przesunie się w prawo (choć o mniejszą wielkość niż w kraju A). Również zagregowana krzywa *IS* odnośnie do całej unii przesunie się w prawo, w stopniu pośrednim między wielkościami, jakie odnotowano w przypadku A i B, a stopy procentowe wzrosną z r_1 do r_2 (przejście z *E'* do *E"*). Ekspansywna polityka fiskalna w kraju A powoduje zatem w drugim etapie przesunięcie krzywych *IS* w obu krajach, lecz o różną wielkość. Zmiana położenia krzywej *IS* odnośnie do całej unii stanowi wypadkową przesunięcia indywidualnych krzywych *IS* w obu krajach.

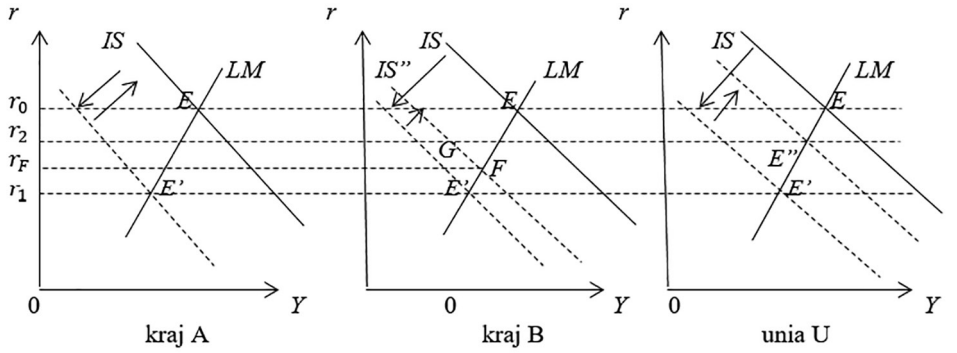
[18] Benassy-Quéré i in. (2010), s. 228–231.

[19] Zauważmy jednak, że jeśli część dodatkowego popytu rządowego w kraju A skieruje się na import z kraju B, przesunięcie krzywej *IS* w kraju A może okazać się niewystarczające, aby powrócić do początkowego położenia. W naszej analizie pomijamy tę możliwość.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W celu oszacowania bilansu skutków ekspansji fiskalnej danego kraju istotna jest skala ostatecznego spadku stóp procentowych. Zauważmy, że spadek stóp procentowych w kraju B (z r_0 do r_2) jest mniejszy, niż gdyby kraj ten funkcjonował poza unią walutową (wówczas stopy spadłyby z r_0 do r_F). Wynika to stąd, że krzywa *IS* w kraju B przesuwa się w prawo w mniejszym stopniu (do położenia *IS*") niż krzywa *IS* odnośnie do całej unii, ponieważ przesunięcie krzywej *IS* w przypadku całej unii stanowi średnią przesunięcia krzywych *IS* w kraju A i w kraju B. W rezultacie nowym punktem równowagi w kraju B staje się punkt *G*, a nie punkt *F*, wobec czego stopy procentowe będą również odpowiednio wyższe. Jako członek unii walutowej kraj B ponosi zatem konsekwencje oddziaływania dwóch efektów zewnętrznych – z jednej strony korzysta na wzroście eksportu do kraju A, lecz z drugiej strony traci na skutek niższego spadku stóp procentowych. Pierwszy z efektów rozprzestrzenia się za pośrednictwem kanału handlu zagranicznego, drugi zaś za pośrednictwem kanału stóp procentowych.

Rysunek 8.6. Szok symetryczny i efekty zewnętrzne polityki fiskalnej w ramach unii walutowej w modelu IS-LM



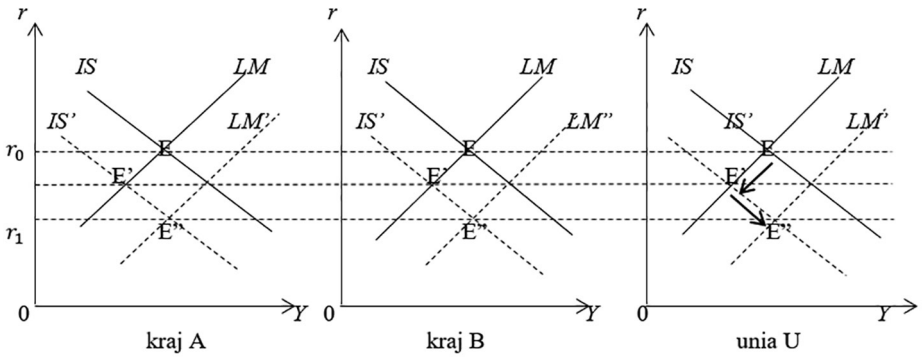
Źródło: opracowanie własne.

Łączny efekt dla kraju B zależy od tego, który z cząstkowych efektów przeważy. Gdy silniejszy staje się efekt zewnętrzny stóp procentowych, kraj B *per saldo* traci i może być zmuszony do zwiększenia wydatków rządowych kosztem wzrostu deficytu budżetowego Kiedy dominuje jednak efekt handlu, kraj B *per saldo* zyskuje dzięki ekspansji fiskalnej kraju A. Lecz w takiej sytuacji kraj A może wstrzymać się z ekspansją fiskalną, licząc na to, że kraj B zdecyduje się na zwiększenie wydatków. Wynika stąd, że poszczególne kraje członkowskie unii walutowej mogą niechętnie

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

angażować się samodzielnie w kosztowną ekspansję fiskalną, ponieważ z wypracowanych przez nie efektów w znacznej mierze korzystają też pozostałe państwa zrzeszone w unii (poprzez wzrost eksportu). Opłaca się więc czekać, aż inne kraje zdecydują się na zwiększenie wydatków rządowych. W sytuacji szoku symetrycznego indywidualna polityka fiskalna poszczególnych państw może więc okazać się mało skuteczna. Dopiero koordynacja polityki fiskalnej (skoordynowana ekspansja fiskalna w obu krajach) umożliwia skuteczne ograniczenie negatywnych efektów szoku popytowego.

Rysunek 8.7. Symetryczny szok popytowy i efekt wspólnej polityki pieniężnej w unii walutowej



Źródło: opracowanie własne.

Zauważmy jednocześnie, że w warunkach szoku symetrycznego w unii walutowej najlepszym instrumentem reagowania jest wspólna polityka pieniężna. Rysunek 8.7 pokazuje, że symetryczny szok popytowy przesuwa krzywe *IS* w lewo w dół o tę samą wielkość (zakładamy nadal, że kraje A i B są podobne). Wówczas wzrost podaży pieniądza (krzywa *LM* przemieszcza się do położenia *LM'*) przekłada się na spadek stóp procentowych w obu krajach z r_0 do r_1, co umożliwia każdej z gospodarek osiągnięcie nowego punktu równowagi *E''* i zapewnia powrót do początkowego poziomu produkcji.

W rozdziale 6 omawialiśmy skuteczność polityki fiskalnej w ramach modelu Mundella-Fleminga i zauważyliśmy, że zależy ona od przyjętego w danym kraju systemu kursowego. W szczególności zaś w warunkach stałych kursów i dużej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

transgranicznej mobilności kapitału polityka fiskalna w skali jednego kraju okazuje się w pełni skuteczna. Takie właśnie warunki panują w unii walutowej. I rzeczywiście, polityka fiskalna w kraju A jest w pełni skuteczna, co potwierdza wniosek płynący z analizy modelu Mundella-Fleminga. Niezależna polityka fiskalna kraju A pozostaje jednak nie bez wpływu na sytuację panującą w kraju B (i ewentualnie także w pozostałych państwach zrzeszonych w unii walutowej). Analiza przeprowadzona w ramce 8.5 pokazuje, że efekty zewnętrzne wynikające z zastosowania indywidualnej polityki w danym kraju mogą być różne w zależności od tego, czy w pozostałych państwach przeważa efekt handlowy, czy też efekt stóp procentowych.

Na tej podstawie można więc sformułować następujące wnioski dotyczące zastosowania poszczególnych rodzajów polityki w ramach unii walutowej:

- wspólna polityka pieniężna lub skoordynowana polityka fiskalna stanowią skuteczne instrumenty reagowania na wstrząsy *symetryczne* w unii walutowej;
- narodowa polityka fiskalna jest preferowanym instrumentem reagowania na wstrząsy *asymetryczne* w pojedynczym kraju unii walutowej, lecz pozostaje ona nie bez wpływu na pozostałe państwa członkowskie, w przypadku których bilans zmian może okazać się ujemny;
- występowanie efektów zewnętrznych uzasadnia koordynację narodowej polityki fiskalnej wśród poszczególnych krajów unii walutowej.

8.6.2. Potrzeba stosowania dyscypliny fiskalnej w obrębie unii walutowej: uzasadnienie teoretyczne

Jak pamiętamy z rozdziału 3, podstawowym uzasadnieniem w przypadku zachowania szczególnej troski o dyscyplinę fiskalną jest występowanie zjawiska „skłonności do deficytu" (*deficit bias*). Skłonność do deficytu to charakterystyczną cechą demokratycznych systemów parlamentarnych, w których władzę sprawują politycy wybierani w wolnych i powtarzalnych wyborach. W warunkach funkcjonowania takich ustrojów państwowych politycy mają skłonność do zwiększania wydatków publicznych w celu spełnienia oczekiwań swoich wyborców; z tego samego powodu unikają też podnoszenia podatków. Takie postępowanie ułatwia im reelekcję i utrzymanie władzy. Zjawisko to nasila się w okresach przedwyborczych, kiedy w kampaniach politycznych padają obietnice zaspokojenia potrzeb i żądań różnych grup wyborców za sprawą zwiększonych wydatków z budżetu.

Związany z wyborami cykl polityczny wpływa na okresowe zwiększanie wydatków publicznych, które mają najczęściej charakter trwały, wywołują wzrost deficytu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i są z reguły finansowane z długu[20]. Skłonność do deficytu nasila się w warunkach występowania tzw. problemu „wspólnego zasobu" (*common pool problem*), który również jest zjawiskiem typowym w przypadku współczesnych demokracji. Polityk reprezentujący określony okręg wyborczy ma zawsze szczególnie silną motywację do „załatwienia" dla swojego rejonu dodatkowych środków finansowych z budżetu ogólnokrajowego. Kwoty te najczęściej niewiele ważą w skali całego budżetu, lecz w warunkach danego okręgu wyborczego będą to kwoty znaczące, mogące przesądzić o dalszej karierze tego polityka. Występuje więc asymetria między kosztami zwiększenia budżetu ogólnokrajowego o pewną kwotę, która obciąża całe społeczeństwo, a korzyściami, jakie z tego tytułu przypadają wyborcom tylko jednego okręgu. Ponieważ większość polityków postępuje w ten właśnie sposób, wspominana asymetria bodźców sprawia, że w systemach demokratycznych pojawia się systematyczna skłonność do deficytu budżetowego, która jest znacznie silniejsza niż tendencja do nadwyżki budżetowej.

W ramach unii walutowej funkcjonują jednak dodatkowe mechanizmy wpływające na osłabienie dyscypliny fiskalnej w poszczególnych krajach, wykraczające poza zjawisko cyklu politycznego i problem „wspólnego zasobu". W celu zapewnienia wysokiej wiarygodności wspólnej waluty w europejskiej Unii Gospodarczo-Walutowej (UGW) należało zapobiec powstawaniu nadmiernego zadłużenia publicznego, ponieważ kraje nadmiernie zadłużone są bardziej skłonne do zmniejszania ciężaru zaciągniętych pożyczek poprzez dopuszczanie do wysokiej inflacji, która pozwala obniżyć realną wartość długu (McKinnon 1996, De Grauwe 2000). Utrzymanie wysokiej wiarygodności podjętego w ramach UGW zobowiązania do utrzymania niskiej inflacji wymagało zatem ścisłej kontroli poziomu długu publicznego w państwach członkowskich i niedopuszczenia do nadmiernego wzrostu zadłużenia. Ponadto chodziło o potrzebę wyeliminowania zjawiska „pasażera na gapę" (*free rider problem*) oraz o uniknięcie zwiększonej skłonności do osiągania deficytu wśród krajów unii walutowej, co wynika z tzw. pokusy nadużycia (*moral hazard*).

Problem „pasażera na gapę" polega na tym, że część kosztów związanych z nadmiernie ekspansywną polityką fiskalną jednego kraju – takich jak koszt wzrostu krótkookresowych i długookresowych stóp procentowych, wzrostu inflacji i spadku kursu waluty – jest ponoszona przez inne kraje należące do strefy wspólnej waluty. To rozproszenie kosztów przy koncentracji korzyści stwarza dla poszczególnych

[20] Wszystkie trzy wymienione cechy spełnia program wsparcia dla rodzin (tzw. program 500+), wprowadzony w 2016 roku w Polsce przez rząd Prawa i Sprawiedliwości.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

państw silne bodźce do stosowania luźnej polityki fiskalnej. Pojedynczy kraj, zwłaszcza kraj niewielki w stosunku do całej strefy wspólnej waluty, mógłby prowadzić taką politykę, nie ponosząc z tego tytułu praktycznie żadnych dodatkowych kosztów, ponieważ koszty te (na przykład w postaci wzrostu stóp procentowych) rozkładałyby się na wszystkie państwa członkowskie. Gdyby w podobny sposób zaczęły jednak postępować także pozostałe kraje, wspólna waluta straciłaby na wartości i wiarygodności, a kraje utrzymujące dyscyplinę fiskalną ponosiłyby z tego tytułu dodatkowe koszty, nie odnosząc przy tym żadnych korzyści.

Co więcej, ryzyko wystąpienia *deficit bias* rośnie dodatkowo w warunkach unii walutowej, ponieważ kraj, który prowadzi nadmiernie ekspansywną politykę fiskalną, może w sytuacji kryzysowej z większym prawdopodobieństwem liczyć na to, że otrzyma pomoc (*bail-out*) od swoich partnerów, niż gdyby znajdował się poza unią walutową. Dzieje się tak, ponieważ pozostałe państwa unii walutowej będą starały się nie dopuścić do bankructwa kraju członkowskiego nadmiernie zadłużonego, gdyż same poniosłyby wówczas duże straty kapitałowe (utrata wartości posiadanych przez nie obligacji rządowych kraju zadłużonego), a zarazem ucierpiałaby wiarygodność i wartość wspólnej waluty (co spowodowałoby spadek kursu i wzrost stóp procentowych). Świadomość tego, że inne kraje staną się zobligowane do udzielenia pomocy, osłabia jednak bodźce skłaniające do utrzymania dyscypliny fiskalnej w poszczególnych państwach, zwłaszcza w tych o tradycyjnie niskiej dyscyplinie fiskalnej. Stwarza to „pokusę nadużycia". Może ona prowadzić do nadmiernej ekspansji fiskalnej w przekonaniu, że inne kraje członkowskie będą zawsze zmuszone przyjść z pomocą[21]. Dobrą ilustracją przebiegu obu opisanych zjawisk była polityka fiskalna Grecji w latach 2000., która ostatecznie doprowadziła kraj do głębokiego kryzysu w 2010 roku.

Wprowadzenie wspólnej waluty i jednolitej polityki pieniężnej przy pozostawieniu poszczególnym krajom pełnej suwerenności w zakresie polityki fiskalnej zwiększało więc ryzyko stosowania nadmiernie ekspansywnej polityki fiskalnej, co mogło doprowadzić do osłabienia podstaw unii walutowej. Ze względu na to państwa członkowskie UE uznały, że niezbędnym warunkiem utrzymania wysokiej wiarygodności unijnej waluty jest nałożenie na politykę fiskalną poszczególnych krajów określonych ograniczeń prawnych i instytucjonalnych, wspólnych dla całej Unii.

[21] Zob. m.in. McKinnon (1996). Odmienne stanowisko reprezentował De Grauwe (2000), który twierdził, że „istnieją bardzo słabe dowody, że unia walutowa rozluźni dyscyplinę fiskalną, zwiększy ryzyko niewypłacalności i wsparcia finansowego".

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

8.7. Znaczenie konwergencji nominalnej. Kryteria konwergencji określone w Traktacie z Maastricht

Konwergencja nominalna oznacza zmniejszanie rozbieżności między wielkościami podstawowych nominalnych zmiennych makroekonomicznych w poszczególnych krajach unii walutowej. W Traktacie z Maastricht (1992) państwa należące do UE ustaliły, że członkami unii walutowej mogą zostać tylko te kraje członkowskie UE, które spełniają określone kryteria konwergencji nominalnej. Dotyczyły one czterech nominalnych zmiennych makroekonomicznych: inflacji, deficytu budżetowego, stóp procentowych i kursu walutowego.

Uzasadnieniem w przypadku wprowadzenia kryteriów członkostwa w postaci wymogu konwergencji nominalnej było dążenie do zapewnienia wysokiego stopnia homogeniczności gospodarek, które miały wchodzić w skład unii walutowej. Chodziło przede wszystkim o uniknięcie sytuacji, w której cele odnoszące się do stabilności makroekonomicznej byłyby rozbieżne w różnych krajach, co uniemożliwiałoby prowadzenie jednej, skutecznej polityki pieniężnej w obrębie całej unii walutowej i podważałoby stabilność wspólnej waluty. Chodziło także o wyeliminowanie ryzyka *free riding* i negatywnych efektów *spillovers* w sytuacji, gdy za sprawą funkcjonowania wspólnej waluty skutki polityki stosowanej przez poszczególne państwa członkowskie miały silny wpływ na sytuację gospodarczą w pozostałych krajach Unii.

8.7.1. Zapisy traktatowe

Przyjęte w Traktacie z Maastricht kryteria konwergencji odnosiły się do zmiennych nominalnych, nie dotyczyły natomiast stopnia konwergencji realnej, czyli zbieżności takich zmiennych realnych jak na przykład poziom rozwoju gospodarczego czy konkurencyjność międzynarodowa. Uznano, że państwa członkowskie UE znajdują się na zbliżonym poziomie rozwoju gospodarczego, a ewentualne różnice w zakresie zamożności czy konkurencyjności były na tyle niewielkie, że nie stanowiły przeszkody w prowadzeniu wspólnej polityki pieniężnej.

Ponadto, przyjęte kryteria konwergencji nominalnej odnosiły się do zmiennych *makroekonomicznych*, a więc nie dotyczyły kryteriów *mikroekonomicznych* sformułowanych przez Mundella, takich jak mobilność siły roboczej czy elastyczność rynków pracy. Uznano, że kryteria te zostały już spełnione na skutek wprowadzenia

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w UE jednolitego rynku zapewniającego, między innymi, swobodę przepływu pracowników między krajami.

Formalne kryteria konwergencji nominalnej (*nominal convergence criteria*), zgodnie z zapisami art. 140 ust. 1 Traktatu o Funkcjonowaniu Unii Europejskiej (TFUE) obejmują jedno kryterium fiskalne, dwa kryteria pieniężne oraz jedno kryterium stabilności kursu walutowego – łącznie cztery kryteria:

- kryterium fiskalne (deficytu budżetowego) wymaga, aby kraj kandydujący do unii walutowej nie był objęty procedurą nadmiernego deficytu (*excessive deficyt procedure* – EDP) w rozumieniu art. 126.6 TFUE, co oznacza, że deficyt budżetowy w tym kraju nie może być wyższy niż 3% PKB;
- kryterium inflacji wymaga, aby roczny kroczący (12-miesięczny) zharmonizowany wskaźnik cen konsumpcyjnych (*Harmonized Index of Consumer Prices* – HICP) w kraju kandydującym nie był wyższy niż średni roczny wskaźnik HICP w trzech krajach członkowskich UE, „które osiągają najlepsze rezultaty pod względem stabilności cen", plus 1,5 punktu procentowego;
- kryterium stóp procentowych wymaga, aby rynkowa stopa oprocentowania długoterminowych papierów skarbowych nie była wyższa niż średnia odpowiednich wartości stóp procentowych w trzech krajach członkowskich UE, „które osiągają najlepsze rezultaty pod względem stabilności cen", plus 2 punkty procentowe;
- kryterium stabilności kursu walutowego wymaga, aby wahania kursu waluty kraju kandydującego mieściły się w „normalnych marginesach przewidzianych mechanizmem wymiany walut europejskiego systemu walutowego przez co najmniej dwa lata, bez dewaluacji w stosunku do waluty innego Państwa Członkowskiego" (system ERM 2), przy czym marginesy te wynoszą maksymalnie +/–15%, wokół parytetu centralnego.

Niekiedy wśród kryteriów konwergencji wymienia się także górny pułap długu publicznego (60% PKB), lecz nie stanowi on formalnego kryterium konwergencji, ponieważ nie został wymieniony *expressis verbis* w art. 140 TFUE ani w Protokole nr 13. Limit długu na poziomie 60% PKB pozostaje natomiast jedną z reguł fiskalnych obowiązujących wszystkie państwa członkowskie UE, zgodnie z zapisami art. 126 oraz Protokołu nr 12.

W art. 140 zawarto jeszcze jedno kryterium akcesji do unii walutowej, odnoszące się do przepisów regulujących funkcjonowanie banku centralnego, które stanowi, że ustawodawstwo dotyczące banku centralnego musi być zgodne z wymogami

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

niezależności banku centralnego, zapisanymi w art. 130 i 131 TFUE[22]. Jest to więc kryterium konwergencji instytucjonalnej. Pełna treść art. 140 ust. 1 odnośnie do kryteriów konwergencji nominalnej i konwergencji instytucjonalnej zawarta została w ramce 8.6. Z kolei w ramce 8.7 przytoczono treść Protokołu nr 13 w sprawie kryteriów konwergencji.

Ramka 8.6. Kryteria konwergencji zapisane w art. 140 ust. 1 TFUE

Art. 140 ust. 1 TFUE przyjmuje następującą treść (podkreślenia – D.R.):

Co najmniej co dwa lata lub na wniosek Państwa Członkowskiego objętego derogacją, Komisja i Europejski Bank Centralny składają Radzie sprawozdania w sprawie postępów dokonanych przez Państwa Członkowskie objęte derogacją w wypełnianiu ich zobowiązań w zakresie urzeczywistnienia unii gospodarczej i walutowej. Sprawozdania te określają, czy **ustawodawstwo krajowe** każdego z tych Państw Członkowskich, w tym statuty jego krajowego banku centralnego, jest zgodne z art. 130 i 131 i ze Statutem ESBC i EBC. Sprawozdania określają również, czy osiągnięty został **wysoki poziom trwałej konwergencji**, analizując, w jakim stopniu każde Państwo Członkowskie spełniło następujące kryteria:

- osiągnięcie wysokiego stopnia stabilności cen – będzie to wynikało ze stopy inflacji zbliżonej do istniejącej w co najwyżej trzech Państwach Członkowskich, które mają najlepsze rezultaty w dziedzinie stabilności cen;
- stabilna sytuacja finansów publicznych – będzie to wynikało z sytuacji budżetowej, która nie wykazuje nadmiernego deficytu budżetowego w rozumieniu art. 126 ust. 6;
- poszanowanie zwykłych marginesów wahań kursów przewidzianych mechanizmem kursów wymiany walut europejskiego systemu walutowego przez co najmniej dwa lata, bez dewaluacji w stosunku do euro;
- trwały charakter konwergencji osiągniętej przez Państwo Członkowskie objęte derogacją i jego udziału w mechanizmie kursów wymiany walut, co odzwierciedla się w poziomach długoterminowych stóp procentowych.

[22] Art. 130 TFUE ma następującą treść: „W wykonywaniu uprawnień oraz zadań i obowiązków, które zostały im powierzone Traktatami i Statutem ESBC i EBC, ani Europejski Bank Centralny, ani krajowy bank centralny, ani członek któregokolwiek z ich organów decyzyjnych nie zwracają się o instrukcje ani ich nie przyjmują od instytucji, organów ani jednostek organizacyjnych Unii, rządów Państw Członkowskich, ani jakiegokolwiek innego organu. Instytucje, organy i jednostki organizacyjne Unii, jak również rządy Państw Członkowskich zobowiązują się szanować tę zasadę i nie dążyć do wywierania wpływu na członków organów decyzyjnych Europejskiego Banku Centralnego lub krajowych banków centralnych przy wykonywaniu ich zadań". Art. 131 TFUE przyjmuje zaś treść następującą: „Każde Państwo Członkowskie zapewnia zgodność swojego ustawodawstwa krajowego, w tym statutu krajowego banku centralnego, z Traktatami i Statutem ESBC i EBC".

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Cztery kryteria określone w niniejszym ustępie i odpowiednie okresy, podczas których każde ma być przestrzegane, są sprecyzowane w Protokole załączonym do Traktatów (Protokół nr 13). Sprawozdania Komisji i Europejskiego Banku Centralnego biorą również pod uwagę wyniki integracji rynków, sytuację i rozwój równowagi płatności bieżących oraz ocenę rozwoju jednostkowych kosztów pracy i innych wskaźników cen.

Jak można zauważyć, w art. 140 podkreśla się, że spełnienie kryteriów konwergencji musi mieć charakter trwały („wysoki poziom trwałej konwergencji"). Oznacza to, że nie wystarczy wykazać się spełnieniem kryteriów w danym momencie. Jeśli okaże się, że realizacja kryteriów jest skutkiem jednorazowych lub tymczasowych dostosowań bądź też została osiągnięta w wyniku „kreatywnej księgowości" lub manipulacji statystycznych, Komisja Europejska i Europejski Bank Centralny mogą stwierdzić, że kraj kandydujący do unii walutowej nie spełnia wymaganych kryteriów członkostwa.

Ramka 8.7. Protokół nr 13 w sprawie kryteriów konwergencji

WYSOKIE UMAWIAJĄCE SIĘ STRONY, PRAGNĄC ustalić szczegóły dotyczące kryteriów konwergencji, którymi Unia będzie się kierować przy podejmowaniu decyzji dotyczących uchylenia derogacji w stosunku do Państw Członkowskich objętych derogacją, o których mowa w art. 140 ust. 1 Traktatu o funkcjonowaniu Unii Europejskiej, UZGODNIŁY następujące postanowienia, które są dołączone do Traktatu o Unii Europejskiej i do Traktatu o funkcjonowaniu Unii Europejskiej:

Art. 1

Kryterium stabilności cen, określone w art. 140 ust. 1 tiret pierwsze Traktatu o funkcjonowaniu Unii Europejskiej, oznacza, że Państwo Członkowskie ma trwały poziom stabilności cen, a średnia stopa inflacji, odnotowana w tym państwie w ciągu jednego roku poprzedzającego badanie, nie przekracza o więcej niż 1,5 punktu procentowego inflacji trzech Państw Członkowskich o najbardziej stabilnych cenach. Inflacja mierzona jest za pomocą wskaźnika cen artykułów konsumpcyjnych na porównywalnych podstawach, z uwzględnieniem różnic w definicjach krajowych.

Art. 2

Kryterium sytuacji finansów publicznych, określone w art. 140 ust. 1 tiret drugie wspomnianego Traktatu, oznacza, że w czasie badania Państwo Członkowskie nie jest objęte

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

decyzją Rady zgodnie z art. 126 ust. 6 wspomnianego Traktatu, stwierdzającą istnienie nadmiernego deficytu.

Art. 3

Kryterium udziału w mechanizmie kursów walut Europejskiego Systemu Walutowego, określonego w art. 140 ust. 1 akapit trzeci wspomnianego Traktatu, oznacza, że Państwo Członkowskie stosowało normalne granice wahań, przewidziane w mechanizmie kursów walut Europejskiego Systemu Walutowego, bez poważnych napięć przynajmniej przez dwa lata przed badaniem. W szczególności Państwo Członkowskie nie zdewaluowało dwustronnego centralnego kursu swojej waluty wobec euro z własnej inicjatywy przez ten sam okres.

Art. 4

Kryterium konwergencji stóp procentowych, określone w art. 140 ust. 1 tiret trzecie wspomnianego Traktatu, oznacza, że w ciągu jednego roku przed badaniem Państwo Członkowskie posiadało średnią nominalną długoterminową stopę procentową nie przekraczającą więcej niż o dwa punkty procentowe stopy procentowej trzech Państw Członkowskich o najbardziej stabilnych cenach. Stopy procentowe oblicza się na podstawie długoterminowych obligacji państwowych lub porównywalnych papierów wartościowych, z uwzględnieniem różnic w definicjach krajowych.

Art. 5

Dane statystyczne, z których korzysta się przy stosowaniu niniejszego Protokołu, są dostarczane przez Komisję.

Art. 6

Rada, stanowiąc jednomyślnie, na wniosek Komisji i po konsultacji z Parlamentem Europejskim, EBC oraz z Komitetem Ekonomiczno-Finansowym, przyjmuje odpowiednie przepisy w celu ustanowienia szczegółów kryteriów konwergencji, wymienionych w art. 140 wspomnianego Traktatu, które zastąpią niniejszy Protokół.

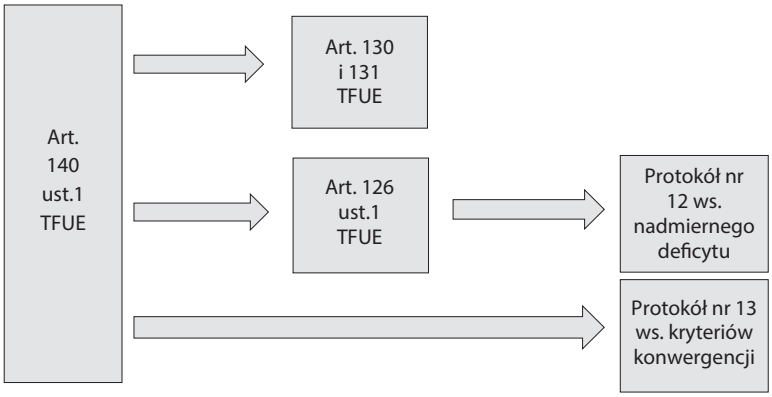
Należy podkreślić, że opisane w Protokole wartości referencyjne oraz sposób ich liczenia i interpretacji, mogą zostać zmienione na skutek jednomyślnej decyzji Rady UE, podjętej na wniosek Komisji Europejskiej. Modyfikacja ta nie wymaga wdrażania procedury ratyfikacyjnej, co zapewnia tym przepisom większą elastyczność niż w przypadku zapisów będących częścią samego Traktatu.

Przepisy dotyczące kryteriów konwergencji nominalnej można odnaleźć w kilku miejscach Traktatu. Podstawowym zapisem jest oczywiście art. 140 TFUE. W jego treści pojawiają się jednak trzy odwołania do innych zapisów traktatowych,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

tj. po pierwsze, odwołanie do Protokołu nr 13 w sprawie kryteriów konwergencji nominalnej, po drugie, odwołanie do art. 130 i 131 TFUE w sprawie niezależności banku centralnego, i, po trzecie, odwołanie do art. 126 ust. 6 TFUE w sprawie procedury nadmiernego deficytu. Z kolei w art. 126 znajduje się odniesienie do Protokołu nr 12 w sprawie procedury nadmiernego deficytu. Wzajemne powiazania tych zapisów zostały ukazane na rysunku 8.8.

Rysunek 8.8. Zapisy dotyczące kryteriów konwergencji w Traktacie o Funkcjonowaniu Unii Europejskiej



Źródło: opracowanie własne.

8.7.2. Ocena kryteriów konwergencji nominalnej

Jak wspomniano, konwergencja nominalna gospodarek wchodzących w skład unii walutowej jest potrzebna w celu uniknięcia rozbieżności celów w polityce makroekonomicznej i zapewnienia skuteczności wspólnej polityki pieniężnej. Na przykład, gdyby do unii walutowej wszedł kraj, w którym poziom inflacji jest wyraźnie wyższy niż w pozostałych państwach, bank centralny unii musiałby zaostrzyć politykę pieniężną, tak aby średnia inflacja w całej unii utrzymała się na dotychczasowym poziomie. To z kolei oznaczałoby dodatkowy koszt dla tych krajów unii walutowej, którym zależałoby na zachowaniu niskiej inflacji. Co więcej, państwa tolerujące wysoką inflację, na przykład ze względu na wysoki poziom długu publicznego, mogłyby sprzeciwiać się w ramach EBC podnoszeniu stóp procentowych, licząc na to, że wyższa inflacja pozwoli im wydostać się z pułapki zadłużenia. Wszystko to pozwala zrozumieć, dlaczego wśród kryteriów konwergencji nominalnej znalazło się kryterium inflacji.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Zapisane w Traktacie z Maastricht kryterium inflacyjne wykazuje jednak również pewne poważne mankamenty. Po pierwsze, nie uwzględnia wzrostu cen wynikającego z oddziaływania efektu Balassy-Samuelsona (efekt „doganiania" – *catching up*). Jak pamiętamy z rozdziału 6 (punkt 6.8.3), efekt ten wynika z szybkiego przyrostu wydajności pracy w produkcji towarów wymiennych w krajach „doganiających", co wywołuje szybszy wzrost płac w produkcji towarów niewymiennych i w rezultacie przekłada się na wyższą średnią inflację w całej gospodarce. Według różnych oszacowań, efekt ten w krajach Europy Środkowej w latach 90. wahał się między 1% a 6% (Rosati 2002). Ponieważ kryterium inflacyjne jest zdefiniowane w sposób dość restrykcyjny (jako średnia inflacja w trzech krajach UE „o najlepszych wynikach w zakresie stabilności cen" plus 1,5 punktu procentowego), efekt B-S może utrudniać niektórym państwom spełnienie tego warunku. Dodajmy, że oddziaływanie tego efektu staje się silniejsze w krajach, które rozwijają się szybciej, co oznacza, że bezwzględne spełnienie kryterium inflacyjnego wymagałoby od tych państw hamowania wzrostu gospodarczego w celu obniżenia inflacji do wymaganego poziomu. W świetle tych spostrzeżeń ekonomiczny sens restrykcyjnego stosowania kryterium inflacyjnego staje się wątpliwy.

Druga, równie ważna wątpliwość dotyczy interpretacji wartości referencyjnej. Aby ją wyliczyć, Europejski Bank Centralny oblicza po prostu nieważoną średnią arytmetyczną stopę inflacji w trzech krajach o *najniższej stopie inflacji*. Tymczasem art. 140.1 TFUE stanowi wyraźnie, że punktem wyjścia przy wyliczeniu wartości referencyjnej jest średnia inflacja w trzech krajach UE „o *najlepszych wynikach w zakresie stabilności cen*". Różnica jest wyraźna – „najniższe stopy inflacji" to nie to samo, co stopy inflacji zapewniające „stabilność cen". Sam EBC definiuje stabilność cen na potrzeby polityki pieniężnej jako stopę inflacji „poniżej 2%, ale blisko 2%" (zob. rozdział 5), tymczasem najniższa stopa inflacji może być bliska zeru (w celu przeprowadzenia obliczeń EBC nie bierze pod uwagę krajów, w których inflacja jest ujemna). Ponieważ w ramach tego samego Traktatu nie mogą funkcjonować dwie różne interpretacje tego samego terminu, przyjęta przez EBC definicja „stabilności cen" nakazywałaby wyliczać wartość referencyjną na podstawie średniej inflacji z trzech krajów, w których stopy inflacji kształtują się „poniżej, ale blisko 2%", a nie z trzech krajów o „najniższej stopie inflacji"[23].

[23] Sam EBC przyjął zresztą formalnie taką właśnie interpretację, wskazując, że chodzi o "najniższe stopy inflacji spójne z celem stabilności cen" ("with regard to «an average rate of inflation, observed over a period of one year before the examination», the inflation rate is calculated using the increase in the latest available 12-month average of the Harmonised Index of Consumer Prices (HICP) over the previous

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

W przypadku kryterium fiskalnego uzasadnieniem dla wymogu niskiego deficytu jest z jednej strony dążenie do wypracowania niskich stóp procentowych w unii walutowej (wysoki deficyt zwiększa potrzeby pożyczkowe państwa i wywołuje wzrost rynkowych stóp procentowych) oraz chęć utrzymania w ryzach długu publicznego. Lecz wartość referencyjną deficytu (3% PKB) przyjęto dość arbitralnie – kraje rozwijające się szybciej mogą wykazywać wyższy deficyt bez wzrostu zadłużenia (zob. rozdział 3, punkt 3.7.3 na temat ścisłego związku między deficytem, długiem i tempem wzrostu gospodarczego).

Często krytykowane jest także założenie dotyczące stabilności kursu walutowego, zgodnie z którym ma on się kształtować w normalnych granicach wahań w ramach parytetu centralnego bez poważnych napięć przez co najmniej dwa lata. W szczególności, kraj kandydujący nie może w tym okresie zdewaluować swojej waluty wobec euro z własnej inicjatywy. Przyjęte przez EBC dopuszczalne pasmo wahań wynosi +/–2,25% wokół ustalonego parytetu centralnego. Uzasadnieniem jest potrzeba osiągnięcia stabilnego kursu równowagi średniookresowej przed wejściem do strefy euro, aby kurs początkowy nie był ani za mocny, ani za słaby, bo mogłoby to prowadzić do systematycznej nierównowagi na rachunku bieżącym bilansu płatniczego i grozić kryzysem.

Kryterium to budzi jednak wątpliwości. Zostało ono przyjęte na początku lat 90., kiedy istniały jeszcze ograniczenia w transgranicznym przepływie kapitałów i skala tych przepływów była niewielka. Zapewniało to znacznie większą stabilność kursów. Obecnie przepływy te wyraźnie wzrosły, w związku z czym i fluktuacje kursów są teraz większe. W tym świetle przyjęty przez EBC przedział tzw. „normalnych wahań" wydaje się zdecydowanie za wąski, zwłaszcza że odnosi się do okresu co najmniej dwóch lat. Analiza zmienności kursu euro wobec innych głównych walut wskazuje, że na przykład w latach 2010–2015 odchylenie standardowe od średniego kursu USD/EUR wynosiło 7,1%, a odchylenie najwyższej i najniżej wartości kursu od średniej wyniosło odpowiednio +14,7% i –18,8%. Wyższa zmienność kursu w ostatnich latach jest czymś naturalnym i bynajmniej nie musi świadczyć o niestabilności waluty.

12-month average. The notion of «at most, the three best-performing Member States **in terms of price stability**», which is used for the definition of the reference value, is applied by using the unweighted arithmetic average of the rate of inflation in the three countries with **the lowest inflation rates, given that these rates are compatible with price stability** and there are no outliers". Zob. European Monetary Institute 1998, s. 32–33, podkreślenia – D.R.).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Stosunkowo najmniej zastrzeżeń budziło kryterium długookresowych stóp procentowych. Miało ono zapobiegać nadmiernym fluktuacjom na rynku długu suwerennego w przypadku krajów wchodzących do strefy euro. Przed utworzeniem UGW utrzymywały się duże różnice w długookresowych stopach rentowności obligacji rządowych poszczególnych państw, co wynikało głównie z ryzyka kursowego, a także z ryzyka płynności i ryzyka wypłacalności. Ponieważ wprowadzenie jednolitej waluty eliminuje ryzyko kursowe, spowodowałoby to gwałtowny wzrost cen obligacji wysoko oprocentowanych i spadek cen obligacji niżej oprocentowanych. To z kolei pociągnęłoby za sobą duże straty kapitałowe wśród posiadaczy obligacji niżej oprocentowanych. Drugim argumentem uzasadniającym przyjęcie tego kryterium była chęć obniżenia i ustabilizowania oczekiwań inflacyjnych. Ponieważ długookresowe stopy zawierają w sobie premię za ryzyko, ich wysokość zależy od oczekiwań inflacyjnych. Trwałość osiągniętej niskiej inflacji można więc oceniać po tym, że długie stopy procentowe również utrzymują się na niskim poziomie.

8.8. Koordynacja polityki gospodarczej w Unii Europejskiej

8.8.1. Ramy instytucjonalne koordynacji polityki gospodarczej w Unii Europejskiej

W pierwszym okresie istnienia unii monetarnej w ramach instytucjonalnej koordynacji polityki gospodarczej w UE funkcjonowały dwa główne mechanizmy. Jednym z nich były Ogólne Wytyczne Polityki Gospodarczej (*Broad Economic Policy Guidelines* – BEPG), które wraz z Ogólnymi Wytycznymi w sprawie Zatrudnienia (*Broad Employment Guidelines* – BEG) miały ułatwiać koordynację narodowej polityki gospodarczej w celu uniknięcia różnokierunkowych impulsów fiskalnych wśród państw strefy euro i zapewniać wewnętrzną spójność polityki gospodarczej w ramach Unii. BEPG i BEG utworzyły razem tak zwane Zintegrowane Wytyczne (*Integrated Guidelines*).

Stosowanie BEPG i BEG reguluje kolejno art. 121 oraz art. 148 Traktatu o Funkcjonowaniu Unii Europejskiej (TFUE). Stwierdza się w nich, że polityka gospodarcza realizowana przez poszczególne państwa członkowskie jest przedmiotem wspólnego zainteresowania i podlega koordynacji. Rada UE na zlecenie Komisji opracowuje projekt Ogólnych Wytycznych Polityki Gospodarczej dla państw członkowskich

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i całej Unii oraz składa sprawozdanie Radzie Europejskiej. Ta z kolei, na podstawie otrzymanego sprawozdania, debatuje nad konkluzją w sprawie ogólnych założeń polityki gospodarczej, realizowanej zarówno wśród poszczególnych państw członkowskich, jak i w ramach Unii. W nawiązaniu do tego orzeczenia Rada UE przyjmuje zalecenia określające BEPG (*Council Recommendation*) oraz podejmuje decyzję w sprawie BEG (*Council Decision*), informując Parlament Europejski o swoim zaleceniu. Rada UE na podstawie sprawozdań przedstawionych przez Komisję nadzoruje rozwój sytuacji gospodarczej w każdym z państw członkowskich oraz w całej Unii i dokonuje oceny zgodności bieżącej polityki gospodarczej poszczególnych państw z ogólnymi kryteriami. Jeżeli Komisja stwierdzi, że polityka gospodarcza wybranego państwa członkowskiego nie jest zgodna z BEPG lub zagraża prawidłowemu funkcjonowaniu unii gospodarczej i walutowej, ma prawo wystosować do danego kraju ostrzeżenie, a Rada – na podstawie wniosku Komisji – może kierować do rządu wskazanego państwa niezbędne zalecenia. Przepisy dotyczące BEPG i BEG, choć dają uprawnienia organom Unii do oceniania polityki gospodarczej państw członkowskich i formułowania zaleceń, mają charakter ogólny (ramowy), nie uwzględniają instrumentów zapewniających wdrożenie zaleceń i nie przewidują możliwości stosowania sankcji w przypadku ich naruszenia. By posłużyć się przykładem, Zintegrowane Wytyczne przyjęte przez Radę UE na 2015 rok zostały wymienione w ramce 8.8.

Ramka 8.8. Zintegrowane Wytyczne na 2015 rok

Zintegrowane Wytyczne na 2015 rok obejmują łącznie osiem punktów – cztery spośród nich dotyczą polityki gospodarczej i przyjęte zostały przez Radę UE w formie zalecenia (*recommendation*), podczas gdy cztery pozostałe odnoszą się do polityki zatrudnienia i uchwalone zostały przez Radę UE w formie decyzji (*decision*):

- Wytyczna 1. Pobudzanie inwestycji
- Wytyczna 2. Sprzyjanie wzrostowi gospodarczemu przez państwa członkowskie poprzez wdrożenie reform strukturalnych
- Wytyczna 3. Usuwanie największych barier dla wzrostu gospodarczego i zatrudnienia na poziomie Unii
- Wytyczna 4. Lepsze zrównoważenie finansów publicznych i zwiększenie ich prowzrostowego charakteru
- Wytyczna 5. Pobudzanie popytu na pracę

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

- Wytyczna 6. Wspieranie podaży pracy, kwalifikacji i kompetencji
- Wytyczna 7. Wzmacnianie funkcjonowania rynków pracy
- Wytyczna 8. Sprzyjanie włączeniu społecznemu, walce z ubóstwem I promowanie równych szans zawodowych.

W dokumentach Rady UE każda z Wytycznych jest rozwinięta i uzupełniona kilku- lub kilkunastozdaniowym tekstem. Niemniej, pozostają one nadal ogólnymi wskazówkami dla Państw Członkowskich. W szczególności zaś nie towarzyszą im żadne formalne procedury weryfikacji oraz przepisy wykonawcze, zapewniające ich realizację[24].

Źródło: Council Recommendation (EU) no. 2015/1184 of 14 July 2015 on broad guidelines for the economic policy of the Member States and of the European Union; Council Decision (EU) of 5 October on guidelines for the employment policies of the Member States for 2015.

Drugim instrumentem służącym koordynacji jest Pakt Stabilności i Wzrostu (*Stability and Growth Pact* – SGP), który ustanawia reguły oraz narzuca ograniczenia mające zwiększyć dyscyplinę w zakresie polityki fiskalnej wśród państw strefy euro. Postanowienia wynikające z Paktu Stabilności i Wzrostu to operacyjne rozwinięcie i uszczegółowienie odpowiednich zapisów wprowadzonych do prawa unijnego na podstawie Traktatu z Maastricht w 1992 roku. Kluczowe znaczenie ma tu art. 126 TFUE, który nakłada na państwa członkowskie obowiązek unikania „nadmiernego deficytu" oraz określa działania, które podejmują organy UE w sytuacji, gdy państwa członkowskie naruszają ten obowiązek. Nadmierny deficyt (*excessive deficit*) jest zdefiniowany jako deficyt, który przekracza tzw. wartość referencyjną (*reference value*)[25]. Uzupełnienie art. 126 stanowi Protokół nr 12 w sprawie procedury dotyczącej nadmiernego deficytu (dołączony do traktatu), gdzie określono m.in. konkretne wielkości liczbowe w przypadku tzw. wartości referencyjnych deficytu i długu (odpowiednio 3% PKB i 60% PKB).

[24] Jedyną sankcją za brak zgodności narodowej polityki gospodarczej z Wytycznymi jest przepis wynikający z art. 121 ust. 4, gdzie czytamy: „Jeżeli w ramach procedury, o której mowa w ustępie 3, zostanie stwierdzone, że polityki gospodarcze Państwa Członkowskiego nie są zgodne z ogólnymi kierunkami, o których mowa w ustępie 2, lub zagrażają prawidłowemu funkcjonowaniu unii gospodarczej i walutowej, Komisja może skierować do danego Państwa Członkowskiego ostrzeżenie. Rada, na zalecenie Komisji, może kierować do danego Państwa Członkowskiego niezbędne zalecenia. Rada, na wniosek Komisji, może zadecydować o podaniu swoich zaleceń do publicznej wiadomości".

[25] Rada UE w głosowaniu podejmuje decyzję o uznaniu deficytu przekraczającego 3% PKB za „nadmierny".

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Pakt Stabilności i Wzrostu pozostaje najdalej zaawansowaną formą koordynacji polityki fiskalno-budżetowej w UE[26].

8.8.2. Pakt Stabilności i Wzrostu

Pakt Stabilności i Wzrostu został formalnie przyjęty w 1997 roku. Składały się na niego trzy dokumenty: rezolucja Rady Europejskiej (przyjęta na szczycie w Amsterdamie 17 czerwca 1997 roku) oraz dwa Rozporządzenia Rady ECOFIN (nr 1466/97 i 1467/97). Mechanizm kontroli i zwiększania dyscypliny fiskalnej w ramach paktu obejmował trzy grupy instrumentów: system wielostronnego nadzoru i kontroli sytuacji fiskalnej w poszczególnych państwach (*multilateral surveillance*), reguły fiskalne w postaci określonych limitów ilościowych i nakazów dotyczących wielkości deficytu i długu publicznego oraz szczegółowe operacyjne rozwinięcie procedury nadmiernego deficytu (*excessive deficit procedure* – EDP), wprowadzonej na podstawie Traktatu z Maastricht (art. 126 TFUE).

System wielostronnego nadzoru i kontroli oraz działania dyscyplinujące przewidziane w ramach Paktu dzielą się na dwie części: część prewencyjną, która ma *ex ante* zapobiegać powstawaniu nadmiernego deficytu, oraz część korekcyjną, służącą eliminowaniu nadmiernego deficytu *ex post*. Część prewencyjna SGP (*preventive arm* – Rezolucja Rady nr 1466/97) obejmuje system wielostronnego nadzoru i kontroli, który nakłada na państwa członkowskie UGW należące do strefy euro obowiązek przedkładania co roku programów stabilności (*stability programs*), a na państwa nienależące do strefy euro (państwa członkowskie UGW objęte derogacją) obowiązek przedkładania programów konwergencji (*convergence programs*). Strategie te stają się następnie przedmiotem oceny dokonywanej przez Komisję Europejską z punktu widzenia przestrzegania reguł fiskalnych oraz – w odniesieniu do krajów objętych derogacją – pod względem stopnia wymaganej konwergencji nominalnej. Komisja sporządza sprawozdanie i przedkłada je Radzie, która może podjąć działania mające na celu zwiększenie dyscypliny fiskalnej w państwach członkowskich. Część korekcyjna SGP (*corrective arm* – Rezolucja Rady nr 1467/97)

[26] Oprócz BEPG, BEG i SGP, państwa członkowskie UE koordynują swoją politykę także w ramach tzw. otwartej metody koordynacji (*open method of coordination* – OMC). Została ona przyjęta w 2000 roku i jest stosowana w celu koordynowania polityki narodowej w dziedzinach, w których Unia nie posiada kompetencji decyzyjnych, takich jak polityka przemysłowa, polityka edukacyjna, polityka badań naukowych czy polityka ochrony zdrowia. Metoda OMC stanowi „miękką" formę koordynacji, wykorzystującą wzajemną wymianę informacji i dobrych praktyk oraz mechanizm wzajemnej oceny w ramach przeglądu „rówieśniczego" (*peer review*) i wzajemnej zachęty (*peer pressure*).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

obejmuje procedurę nadmiernego deficytu (EDP), mającą na celu obniżenie deficytu do dopuszczalnego poziomu, i przewiduje możliwość zastosowania sankcji finansowych wobec krajów, które nie wykonują zaleceń naprawczych i uporczywie uchylają się od obowiązku unikania nadmiernego deficytu.

Dyscyplina fiskalna w państwach członkowskich UE oceniana jest pod względem realizacji trzech głównych kryteriów fiskalnych. Obejmują one:

a) ograniczenie dopuszczalnego (planowanego i rzeczywistego) poziomu deficytu sektora finansów publicznych do maksimum 3% PKB,
b) ograniczenie poziomu długu publicznego brutto do maksimum 60% PKB,
c) obowiązek utrzymywania równowagi budżetowej w średnim okresie (w średnim okresie budżet powinien być zrównoważony lub wykazywać nadwyżkę)[27].

8.8.3. Procedura nadmiernego deficytu

Procedura nadmiernego deficytu (*excessive deficit procedure* – EDP) stanowi rozwinięcie zapisów art. 126 TFUE i pozostaje zasadniczym elementem części korekcyjnej Paktu Stabilności i Wzrostu. Punktem wyjścia jest przepis ust. 1, który nakłada na państwa członkowskie obowiązek unikania nadmiernego deficytu budżetowego. W przypadku, gdy deficyt przekroczy określoną w Traktacie wartość odniesienia (3% PKB), Komisja składa sprawozdanie do Rady UE, a ta z kolei może podjąć decyzję o wszczęciu wobec danego państwa procedury nadmiernego deficytu. Zgodnie z tym rozporządzeniem, państwo członkowskie ma obowiązek obniżenia deficytu do dozwolonego poziomu w określonym czasie i na określonych warunkach. Ramka 8.9 zawiera szczegółowy opis procedury nadmiernego deficytu.

Ramka 8.9. Art. 126 TFUE: procedura nadmiernego deficytu

Definicja „nadmiernego deficytu" (*excessive deficit*) oraz procedura postępowania w przypadku wystąpienia tego zjawiska w kraju członkowskim jest zapisana

[27] Reguły a) i b) są bezpośrednim przeniesieniem odpowiednich zapisów Traktatu z Maastricht i Protokołu nr 12, dołączonego do traktatu. Reguła c) nie znajduje bezpośredniego odpowiednika w Traktacie i została wprowadzona formalnie w postaci Rezolucji Rady Europejskiej. W Traktacie istnieje jeszcze jeden zapis regulujący politykę fiskalną. Jest nim zakaz wykupu długów zadłużonych państw członkowskich UE przez narodowe banki centralne, EBC, inne instytucje UE i rządy pozostałych państw członkowskich UE oraz zakaz uprzywilejowanego dostępu rządów państw UE do krajowych instytucji kredytowych (art. 123 TFUE).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

w art. 126 TFUE ust. 2–14 oraz w Protokole nr 12, dołączonym do Traktatu. Deficyt uważa się za „nadmierny", gdy Komisja Europejska uzna, że jego wielkość w danym państwie członkowskim przekracza wartość odniesienia (3% PKB), i kiedy na tej podstawie Rada UE podejmie formalną decyzję o wystąpieniu „nadmiernego deficytu". Wartość odniesienia zapisana jest w Protokole nr 12.

Procedura nadmiernego deficytu ma charakter wieloetapowy. Składa się na nią:

- **etap 1**: komisja bada na podstawie sprawozdań i planów budżetowych dostarczonych przez państwa członkowskie, czy nie występuje nadmierny deficyt (lub możliwość jego wystąpienia);
- **etap 2**: jeśli deficyt przekracza 3% PKB, Komisja „w zasadzie" stwierdza, że występuje nadmierny deficyt, chyba że uzna, że występują wyjątkowe okoliczności i deficyt ma charakter „tymczasowy";
- **etap 3**: komisja sporządza sprawozdanie i po zaopiniowaniu go przez Komitet Ekonomiczno-Finansowy, wysyła sprawozdanie do Rady UE;
- **etap 4**: Rada UE w trybie głosowania ustala, czy występuje „nadmierny deficyt"; jeśli tak, Rada formułuje „rekomendację" z zaleceniem ograniczenia deficytu do dopuszczalnych rozmiarów w określonym czasie; przyjęcie rekomendacji przez Radę oznacza początek procedury EDP;
- **etap 5**: jeśli państwo członkowskie nie podejmuje koniecznych działań naprawczych lub podjęte przez nie starania są niewystarczające, Rada może podać zalecenia do publicznej wiadomości oraz zażądać okresowych sprawozdań;
- **etap 6**: jeżeli w dalszym ciągu państwo członkowskie nie realizuje zaleceń Rady, organ ten może skierować wezwanie do podjęcia działań (*notice*) i zażądać dostarczenia precyzyjnego harmonogramu działań;
- **etap 7**: sankcje – funkcjonują tylko w odniesieniu do państw członkowskich strefy euro; jeśli dany kraj nie realizuje określonych zaleceń, Radzie przysługuje zastosowanie jednego lub więcej spośród następujących środków:
  - może zażądać opublikowania dodatkowych informacji przed kolejną emisją obligacji rządowych;
  - może zwrócić się do Europejskiego Banku Inwestycyjnego z sugestią zaostrzenia warunków udzielania pożyczek danemu państwu;
  - może zażądać złożenia nieoprocentowanego depozytu w określonej wysokości do czasu, aż nadmierny deficyt zostanie skorygowany (część stała 0,2% PKB plus część zmienna 0,1% PKB za każdy punkt procentowy deficytu utrzymującego powyżej limitu 3% PKB. Wielkość depozytu nie może jednak przekroczyć 0,5% PKB).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Jeśli nadmierny deficyt utrzymuje się przez dwa lata, depozyt zamienia się w karę finansową (nie podlega zwrotowi). Kwota sankcji dzielona jest między kraje członkowskie (nie mające nadmiernego deficytu) proporcjonalnie do wielkości ich dochodu narodowego brutto. Korekta musi być dokonana w ciągu roku po ogłoszeniu EDP o tyle, o ile nie występują „szczególne okoliczności".

Procedura EDP była wielokrotnie stosowana w UE. Nałożono ją m.in. na Grecję (kilkakrotnie w latach 1994–2009, choć deficyty z lat 1997–2004 zostały ujawnione dopiero w 2005 roku), Portugalię (lata 2002 i 2006), Niemcy (lata 2003–2005), Francję (lata 2003–2005) oraz Włochy (lata 2005–2007). W 2010 roku w związku z kryzysem finansowym procedurą EDP objęte zostały 23 państwa członkowskie UE. Ich liczba spadła pod koniec 2015 roku do dziewięciu.

Procedura EDP była krytykowana ze względu na przewlekłość i uznaniowość decyzji (zarówno na etapie oceny, jak i rekomendacji). W całym okresie funkcjonowania SGP Rada UE nigdy nie nałożyła sankcji finansowych na żadne państwo członkowskie. Powstrzymywanie się od nakładania kar (nawet w przypadku Grecji, która w latach 1999–2010 nigdy nie odnotowała deficytu niższego niż 3% PKB) oraz odrzucenie przez Radę UE wniosku Komisji o rozpoczęcie procedury EDP wobec Niemiec w 2002 roku wskazuje na silne upolitycznienie całego procesu.

8.8.4. Ocena systemu koordynacji polityki fiskalnej w Unii Europejskiej w latach 1999–2009

Reguły zapewniające dyscyplinę w zakresie finansów publicznych w ramach UE od początku były przedmiotem powszechnej krytyki. Po pierwsze, zarzucano im arbitralność i brak solidnych podstaw teoretycznych (Buiter, Corsetti i Roubini 1993, s. 57–100; Wyplosz 2002; Fatas i Mihov 2003). Podkreślano, że nie istnieje powszechnie akceptowana teoria optymalnych rozmiarów zadłużenia publicznego. W świetle tych spostrzeżeń wymóg równowagi średniookresowej może zostać zakwestionowany, ponieważ w długim okresie prowadzi do zerowego zadłużenia, co nie musi być optymalne z punktu widzenia użyteczności społecznej (przy założeniu, że istnieje dodatnia społeczna stopa dyskonta). Po drugie, część autorów wskazywała na nadmierną restrykcyjność przyjętych reguł. Podkreślano, że w ramach unii walutowej, gdy poszczególne kraje zostają pozbawione samodzielnej polityki pieniężnej i kursowej, polityka fiskalna staje się jedynym instrumentem makroekonomicznym pozwalającym łagodzić skutki szoków asymetrycznych

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(Eichengreen i Wyplosz 1998, s. 67–113; Melitz 2000, s. 3–22; Buti i Sapir, 1998). Tymczasem przyjęte reguły fiskalne mogą nadmiernie usztywniać politykę fiskalną poprzez nałożenie na nią ilościowych ograniczeń, co osłabia jej skuteczność jako instrumentu stabilizacji makroekonomicznej. Niektórzy ekonomiści byli zdania, że twórcy SGP zbyt tendencyjnie kierowali się obawą o niestabilność deficytu i długu, nie uwzględniając w dostatecznym stopniu konieczności zapewnienia niezbędnej elastyczności budżetu. Przestrzegali przed tym, że narzucony za sprawą reguł fiskalnych brak elastyczności w zakresie polityki budżetowej w krajach UGW stwarza większe zagrożenie niż ryzyko niewypłacalności i wymuszonego wsparcia finansowego dla krajów nadmiernie zadłużonych (zob. np. De Grauwe 2000)[28]. Ten sam autor w innym miejscu ubolewał nad tym, że reguły fiskalne stanowią „wotum nieufności wobec demokratycznych instytucji w państwach członkowskich" i narzucają suwerennym państwom niedopuszczalne warunki zewnętrzne (De Grauwe 2002, cyt. za: Buti 2006). Po trzecie, kwestionowano wybór kategorii długu brutto jako wskaźnika dyscypliny fiskalnej, wskazując, nie bez racji, że o długookresowej stabilności finansowej kraju decyduje raczej kategoria majątku netto (*net wealth*), uwzględniająca także należności i inne aktywa (Buiter i in. 1993; Calmfors 2005; Wyplosz 2005, s. 70–84).

Ostatecznie te i inne zastrzeżenia nie zostały uwzględnione, a reguły fiskalne przyjęto w zaproponowanej formie, głównie pod naciskiem Niemiec (zob. np. Stark 2001). Jak jednak pokazały późniejsze wydarzenia, całe to rozbudowane instrumentarium nie było w stanie zapobiec nadmiernemu zadłużaniu się państw członkowskich i w konsekwencji nie uchroniło strefy euro przed głębokim kryzysem zadłużenia. W praktyce głównym problemem okazała się jednak nie tyle nadmierna sztywność reguł, ile możliwość ich bezkarnego naruszania.

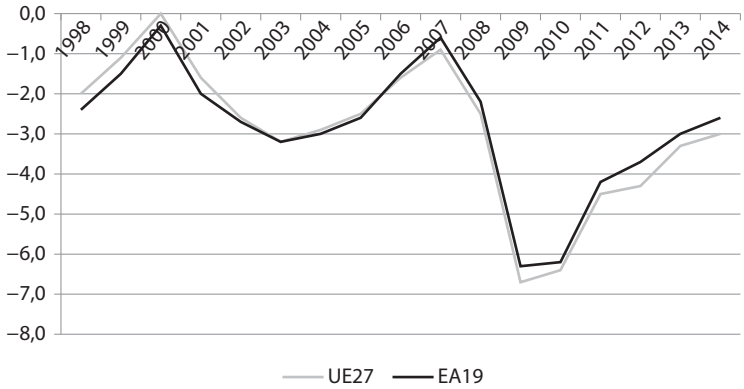
Oceniając skuteczność reguł fiskalnych w UGW w pierwszym okresie ich funkcjonowania (do czasów kryzysu finansowego i gospodarczego z lat 2008–2012), należy przede wszystkim spojrzeć na dane dotyczące deficytu i długu oraz ich zmiany w okresie funkcjonowania wspólnej waluty. Wykres 8.2 przedstawia zagregowane wielkości salda sektora finansów publicznych (definiowanego według systemu ESA2010 jako potrzeby pożyczkowe netto sektora publicznego – *net lending/borrowing of the general government sector*) odnośnie do 27 państw członkowskich UE (UE27) i 19 państw strefy euro (EA19). Jak widać, zagregowane

[28] Kryzys grecki dowodzi, że w tej kwestii De Grauwe nie uwzględnił w odpowiedniej mierze zagrożeń związanych z nadmiernym zadłużaniem się niektórych krajów strefy euro.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

saldo sektora finansów publicznych było ujemne w całym analizowanym okresie obowiązywania SGP (lata 1998–2014), z wyjątkiem 2000 roku, kiedy łączne saldo w UE27 osiągnęło wartość bliską zera. Średnia wielkość deficytu w krajach UE i państwach strefy euro wyniosła w całym siedemnastoletnim okresie około 2,9% PKB. Nawet jeśli w sposób szczególny potraktuje się lata kryzysu 2008–2012, kiedy deficyt budżetowy wielu krajów wzrósł silnie za sprawą spadku dochodów i dyskrecjonalnych działań rządów, mających na celu pobudzanie koniunktury i ratowanie banków, trudno uznać, aby państwa członkowskie UE i EA wypełniały zobowiązanie utrzymywania równowagi budżetowej w średnim okresie[29].

Wykres 8.2. Saldo sektora finansów publicznych w Unii Europejskiej (27 państw) i w strefie euro (19 państw) w latach 1998–2014 (dane w % PKB)



Źródło: opracowanie własne na podstawie danych Eurostatu.

W tabeli 3.4 w rozdziale 3 przedstawiono szczegółowe wielkości deficytu budżetowego (jako % PKB) utrzymującego się w poszczególnych krajach UE27 w okresie obowiązywania SGP. Wspomniana tabela pokazuje, że na łączną liczbę 457 obserwacji rocznych w 192 przypadkach odnotowano wyższy deficyt niż referencyjna wartość 3% PKB (zaznaczone na szaro) co stanowiło pełne 39% wszystkich rocznych obserwacji. Oczywiście, nie we wszystkich przypadkach oznaczało to naruszenie reguł unijnych – część naruszeń miała bowiem miejsce w krajach,

[29] W istocie, przytoczone dane wskazują, że kraje UE w praktyce realizowały długookresowy cel w postaci deficytu budżetowego na poziomie około 2,9% PKB, co wyraźnie odbiegało od przyjętego formalnie celu zrównoważonego budżetu w długim okresie.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

które wówczas nie należały jeszcze do UE lub nie były członkami strefy euro (EA). Jeśli z całego zbioru wyłączy się obserwacje dotyczące wspomnianych podmiotów, to okazuje się, że w grupie UE27 (w latach 1998–2014) doszło do 158 naruszeń obowiązującego limitu 3% na ogólną liczbę 379 rocznych obserwacji (41,7% całości), a w grupie EA19 miało miejsce 105 naruszeń w porównaniu z ogólną liczbą 228 obserwacji (46,1% całości).

Z pewnością odosobnionym przypadkiem jest Grecja, która dopuściła się największej liczby naruszeń (15 przypadków na 15 obserwacji) i gdzie ani razu w analizowanym okresie deficyt finansów publicznych nie spadł poniżej 3% PKB. Spośród pozostałych krajów strefy euro regułę deficytu łamały najczęściej Portugalia (13 razy), Francja (10 razy) oraz Włochy (9 razy). Do najbardziej zdyscyplinowanych państw należały zaś Estonia (która od czasu przystąpienia do strefy euro w 2011 roku nie zanotowała naruszeń) oraz Luksemburg (raz od 1998 roku), a także Finlandia (2 razy).

Spośród krajów nienależących do strefy euro najlepszą dyscypliną fiskalną wykazywały się Szwecja (brak naruszeń) i Dania (1 naruszenie), najczęściej łamały natomiast regułę deficytu Polska (10 razy w okresie 11 lat), Wielka Brytania (9 razy w okresie 17 lat)[30] i Węgry (8 razy w okresie 11 lat). Ogólnie można więc stwierdzić, że sytuacja fiskalna w państwach UE pozostawiała wiele do życzenia, co dowodzi, że system dyscypliny budżetowej w ramach Paktu Stabilności i Wzrostu był zasadniczo mało skuteczny.

8.9. Kryzys w strefie euro

8.9.1. Kryzys w strefie euro w latach 2009–2012

Kryzys zadłużenia, który w końcu 2009 roku dotknął Grecję, a rok później także Irlandię i Portugalię, wywołał falę alarmistycznych komentarzy wśród polityków, ekonomistów, a zwłaszcza dziennikarzy. Wielu z nich prześcigało się w snuciu

[30] Wielka Brytania na mocy wyłączenia ze strefy euro korzysta ze specjalnych i nieco łagodniejszych przepisów dotyczących deficytu. Przypomnijmy, że art. 126 ust. 1 stanowi, że „kraje członkowskie unikają nadmiernego deficytu". Lecz Wielka Brytania jest wyłączona spod działania tego zapisu, zgodnie z postanowieniem Protokołu nr 15 w sprawie niektórych postanowień dotyczących Zjednoczonego Królestwa Wielkiej Brytanii i Irlandii Północnej. W punkcie 5 tego Protokołu stwierdza się, że „Zjednoczone Królestwo *podejmuje wysiłki* w celu uniknięcia nadmiernego deficytu", co stanowi łagodniejszy wymóg niż zapis art. 126.1, który obowiązuje pozostałe państwa UE.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

katastroficznych wizji dotyczących dalszych losów wspólnej waluty. Dominującym przekazem medialnym w latach 2009–2012 było przepowiadanie rychłego rozpadu strefy euro i powrotu państw członkowskich do walut narodowych. Powoływano się przy tym na opinie znanych ekonomistów, zwłaszcza anglosaskich, którzy od początku krytykowali projekt integracji walutowej. Do największych sceptyków należeli m.in. Martin Feldstein, Paul Krugman i Nuriel Roubini, którzy twierdzili, że wprowadzenie wspólnej waluty w Europie było aktem czysto politycznym, pozbawionym dostatecznych podstaw ekonomicznych, i że strefa euro rozpadnie się w ciągu kilku lat (zob. Feldstein 1997, 2000; Roubini i Mihm 2011; Krugman 2012)[31]. W polskiej literaturze również krytycznie o unii walutowej wypowiadali się m.in. Kawalec i Pytlarczyk (2012)[32].

Dziś już wiadomo, że strefa euro przetrwała, wspólna waluta ma się nieźle, a państwa członkowskie podjęły działania na rzecz dalszej integracji. Lecz nadal warto zadawać pytanie o to, na czym właściwie polegał kryzys z lat 2009–2012 i czy u jego źródeł nie legł jakiś błąd popełniony w momencie utworzenia strefy euro?

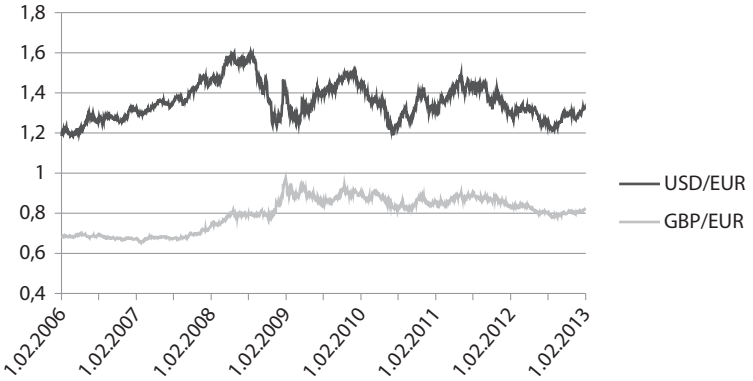
[31] Oto przykład typowych dla tego nurtu opinii: „…Prawda jest taka, że brak dyscypliny fiskalnej nie jest jedynym, ani nawet najważniejszym źródłem kłopotów Europy – i to nawet w Grecji, gdzie rząd istotnie prowadził nieodpowiedzialna politykę… Nie, prawdziwe przyczyny euro-bałaganu nie leżą w rozrzutności polityków, ale w arogancji elit – a zwłaszcza elit politycznych, które popchnęły Europę do przyjęcia wspólnej waluty, gdy kontynent zupełnie nie był przygotowany do takiego eksperymentu…" (Krugman 2010).

[32] Ich zdaniem główną wadą strefy euro jest brak możliwości dewaluacji waluty w krajach dotkniętych kryzysem zadłużenia. Jak piszą autorzy: „Wprowadzenie euro okazało się posunięciem chybionym, które zamiast scementować Unię Europejską może doprowadzić do poważnych konfliktów i zagrozić współpracy europejskiej. Głównym problemem związanym ze wspólną walutą jest to, że państwa członkowskie pozbawione zostały możliwości posługiwania się bardzo skutecznym i trudnym do zastąpienia instrumentem dostosowawczym, jakim jest możliwość zmiany kursu walutowego. W przypadku utraty międzynarodowej konkurencyjności lub nagłej utraty zaufania rynku, osłabienie waluty (połączone z odpowiednią polityką fiskalną i monetarną) jest instrumentem skutecznie przywracającym konkurencyjność gospodarki i równowagę w obrotach międzynarodowych, pozwalającym na szybkie wejście na ścieżkę wzrostu gospodarczego. Próby usunięcia poważnej luki w konkurencyjności kraju, bez zmiany kursu walutowego, poprzez restrykcje fiskalne lub zacieśnienie monetarne, powodują duże koszty w postaci spadku realnego PKB i wzrostu bezrobocia, a w warunkach systemu demokratycznego kończą się bardzo często niepowodzeniem" (Kawalec i Pytlarczyk 2012). Autorzy nie dostrzegają jednak, że to nie euro stało się przyczyną kryzysu zadłużenia, lecz nieodpowiedzialna polityka budżetowa niektórych państw strefy euro (Grecja, Portugalia) lub nieskuteczny nadzór bankowy (Irlandia, Hiszpania). Euro mogło rzeczywiście ułatwiać zaciąganie pożyczek, lecz nie było obowiązku zadłużania się ponad miarę. Ponadto, jak wskazywaliśmy w rozdziale 6, rekomendowana przez nich dewaluacja kursu jest, po pierwsze, kosztowna i na dłuższą metę nieskuteczna, a po drugie, pozostaje jedynie reakcją na popełnione już wcześniej błędy w zakresie polityki fiskalnej. Optymalną strategią (*first-best*) byłoby oczywiście, w pierwszej kolejności, unikanie błędów i utrzymanie dyscypliny budżetowej – nie istniałaby wówczas potrzeba dewaluowania waluty krajowej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Nie był to na pewno kryzys wspólnej waluty. Typowymi objawami kryzysu walutowego są bowiem gwałtowny spadek kursu danej waluty i skokowy wzrost stóp procentowych. Żadne z tych zjawisk nie wystąpiło wówczas w strefie euro. Wykres 8.3 przedstawia zmiany kursu euro wobec dolara amerykańskiego i funta brytyjskiego bezpośrednio przed i w czasie kryzysu. Można zauważyć, że perturbacje związane z kryzysem greckim spowodowały przejściowe osłabienie euro wobec dolara w 2009 roku, lecz potem kurs unijnej waluty powrócił do poprzedniego poziomu. Z kolei wobec funta kurs euro umacniał się nawet w 2009 roku. Wahania kursu, naturalne w okresach perturbacji na rynkach finansowych, nie były jednak bynajmniej objawem kryzysu wspólnej waluty. Euro pozostawało nadal drugą co do ważności walutą rezerwową na świecie, notując udział w rezerwach międzynarodowych na poziomie około 25%. Podobny wniosek płynie z analizy zmian stóp procentowych (wykres 8.4). W latach 2009–2012 zarówno oficjalne stopy procentowe EBC, jak i rentowności niemieckich obligacji rządowych spadały, osiągając historyczne minima, co świadczyło o utrzymującym się wysokim zaufaniu do waluty europejskiej.

Wykres 8.3. Kurs euro wobec dolara amerykańskiego i funta brytyjskiego – notowania dzienne w latach 2006–2012



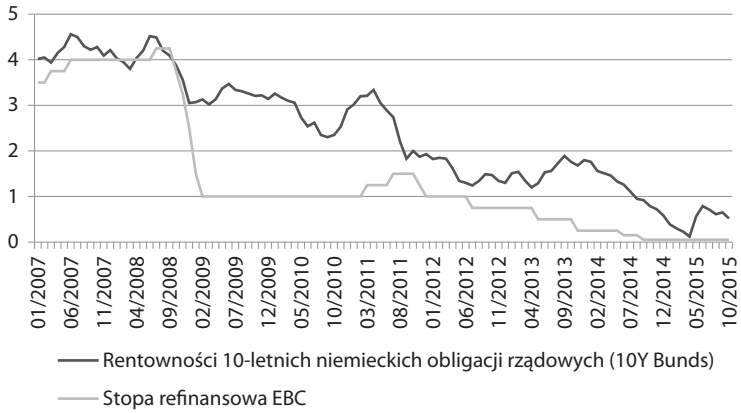
Źródło: opracowanie własne na podstawie danych EBC.

Nie był to również kryzys unii walutowej. W wielu krajach strefy euro, takich jak Niemcy, Holandia, Belgia, Austria, Finlandia czy Luksemburg, utrzymana została równowaga finansowa i makroekonomiczna mimo wystąpienia niekorzystnych zjawisk w gospodarce światowej i bezpośrednim ich otoczeniu. Jak wynika z danych zamieszczonych w tabeli 8.3, strefa euro jako całość nie była bardziej zadłużona niż

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

inne główne gospodarki świata, a poziom łącznego deficytu budżetowego okazał się w jej przypadku niższy niż w innych gospodarkach.

Wykres 8.4. Stopa refinansowa EBC (stan na koniec miesiąca) i stopa rentowności 10-letnich niemieckich obligacji skarbowych (średnia w miesiącu) w latach 2007–2015 (dane w %)



Źródło: opracowanie własne na podstawie danych EBC.

Tabela 8.3. Saldo sektora finansów publicznych i dług publiczny w strefie euro, USA, Wielkiej Brytanii i Japonii w latach 2009–2013 (dane w % PKB)

Kraj	Saldo sektora finansów publicznych					Dług publiczny				
	2009	2010	2011	2012	2013	2009	2010	2011	2012	2013
Strefa euro	–6,4	–6,2	–4,1	–3,6	–2,9	80,0	92,4	85,8	89,1	90,9
USA	–12,8	–12,2	–7,5	–6,8	–4,1	76,0	87,1	95,2	99,4	100,1
UK	–11,4	–10,1	–7,6	–8,3	–5,7	67,1	78,4	81,8	85,8	87,3
Japonia	–3,5	–8,3	–8,7	–9,0	–9,2	188,7	193,3	210,6	218,8	226,1

Źródło: opracowanie własne na podstawie danych Eurostatu i OECD.

Kryzys lat 2009–2012 nie był więc ani kryzysem wspólnej waluty, ani kryzysem strefy euro. W rzeczywistości był to kryzys zadłużenia suwerennego kilku państw członkowskich – przede wszystkim Grecji, Irlandii i Portugalii, choć zasadnicze przyczyny wzrostu zadłużenia publicznego były różne w przypadku każdego z państw. Nadmierne zadłużanie się sektora rządowego stało się główną przyczyną kryzysu w Grecji oraz w Portugalii, a także spowodowało przejściowe kłopoty

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

finansowe Włoch. Z kolei nadmierna ekspansja kredytowa oraz inwestycyjna banków i związane z tym nadmierne zadłużanie się sektora prywatnego w sektorze bankowym okazały się główną przyczyną kryzysu finansowego w Irlandii oraz w Hiszpanii. Każdy z tych czynników wpływał ostatecznie na wzrost zadłużenia suwerennego. Gdy w 2009 roku udział zagrożonych kredytów w bilansach banków Irlandii i Hiszpanii zaczął szybko rosnąć, rządy tych państw zmuszone były do ratowania banków, udzielając im wsparcia w postaci gwarancji finansowych na dużą skalę. W rezultacie nadmierny dług prywatny przekształcił się w dług publiczny za pośrednictwem strat banków gwarantowanych lub bezpośrednio przejmowanych przez rząd (Rosati 2013a).

Oczywiście, w ramach unii walutowej łatwiej było się zadłużać, ponieważ stopy procentowe pozostawały zasadniczo niskie, a rynki finansowe przez długi czas były przekonane, że strefa euro nie dopuści do bankructwa żadnego z krajów członkowskich. Gdyby kraje, takie jak Grecja czy Portugalia, dysponowały własną walutą, nie byłyby zapewne w stanie zadłużyć się na tak dużą skalę, ponieważ rynki finansowe prawdopodobnie nigdy nie zgodziłyby się na finansowanie rosnącego zadłużenia tych krajów powyżej pewnego poziomu uznanego za bezpieczny. Trzeba jednak również zauważyć, że żadna z tych odmian kryzysu nie jest powiązana wyłącznie ze wspólną walutą lub ze strefą euro. Kryzysy zadłużenia suwerennego stały się powszechnym zjawiskiem w krajach *emerging markets* w latach 80. i 90., a kryzysy zadłużenia prywatnego (bankowe) miały też miejsce w krajach rozwiniętych funkcjonujących poza strefą euro (USA, Islandia, Łotwa).

Kryzys zadłużenia jest przewlekły, trudny do przezwyciężenia i dewastujący w skutkach w przypadku sektora realnego gospodarki. Nadmierne zadłużenie rządów wymusza oszczędności budżetowe w postaci ograniczeń wydatków i podnoszenia podatków (*austerity measures*), co działa prorecesyjnie. Z kolei nadmierne zadłużenie sektora prywatnego (przedsiębiorstw i gospodarstw domowych) zmusza podmioty gospodarcze do delewarowania, co ogranicza bieżące wydatki i również działa prorecesyjnie. Wreszcie, straty bilansowe banków pociągają za sobą tworzenie rezerw na ich pokrycie, a to ogranicza skalę udzielanych kredytów (*credit-crunch*) i przyczynia się do dalszego osłabiania aktywności gospodarczej. Te jednokierunkowe zachowania głównych sektorów gospodarki wywołują przewlekłą recesję, która – w odróżnieniu od cyklicznej recesji – bywa nazywana recesją bilansową (zob. ramka 8.10).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 8.10. Recesja bilansowa według Richarda Koo

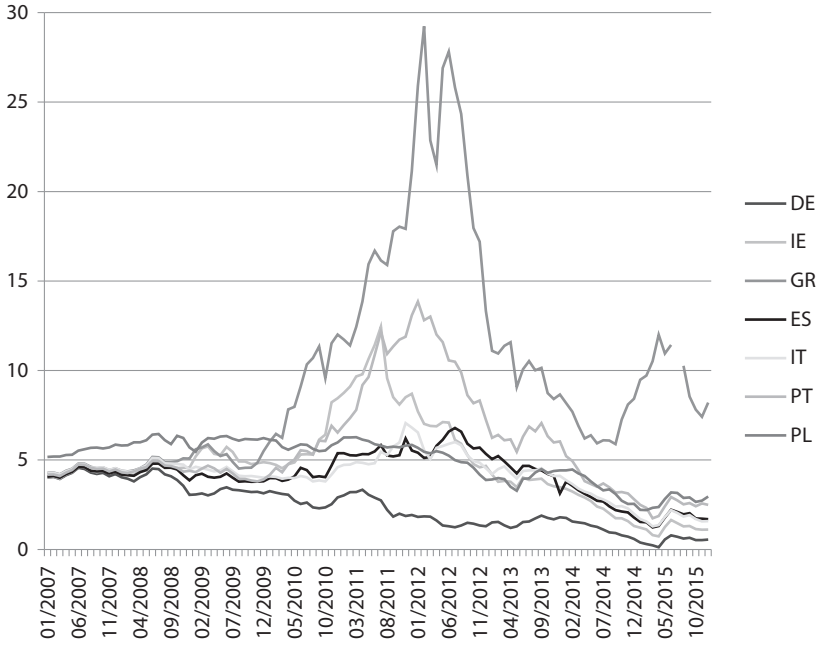
Recesja bilansowa jest szczególną odmianą recesji gospodarczej, spowodowaną wysokim poziomem zadłużenia sektora prywatnego, a nie normalnym cyklicznym spowolnieniem zagregowanego popytu. W latach 2008–2010 w krajach rozwiniętych mieliśmy do czynienia z ograniczaniem poziomu wydatków bieżących oraz inwestycyjnych jednocześnie we wszystkich trzech głównych sektorach gospodarki (sektor publiczny, sektor prywatny, sektor bankowy) w związku z koniecznością obniżania poziomu nadmiernego zadłużenia. Jak zauważył amerykański ekonomista Richard Koo (2009, 2011), w warunkach makroekonomicznej równowagi wewnętrznej i zewnętrznej zadłużenie netto sektora prywatnego powinno znajdować pokrycie w aktywach netto sektora bankowego, a budżet sektora rządowego powinien pozostawać względnie zrównoważony. Gdy sektor publiczny staje się jednak nadmiernie zadłużony i ogranicza swoje wydatki, a na dodatek sektor bankowy ogranicza kredyty ze względu na konieczność pokrycia strat bilansowych, ta ogólna równowaga ulega naruszeniu i gospodarka wchodzi w recesję. Spowolnienie gospodarcze jeszcze bardziej ogranicza dochody, a w ten sposób zwiększa się relatywny poziom długu, co zmusza podmioty gospodarcze do dalszego zaciskania pasa. Takie negatywne sprzężenie zwrotne może trwać kilka lat. Dlatego też recesja bilansowa jest z reguły dłuższa i trudniejsza do przezwyciężenia niż normalna recesja cykliczna.

Kryzys zadłużenia w strefie euro rozwijał się stopniowo, najpierw uderzając w Grecję, a następnie przenosząc się także na inne wysoko zadłużone kraje strefy euro. W maju 2010 roku Grecja otrzymała od Unii Europejskiej pomoc finansową w kwocie 110 mld EUR. W listopadzie wsparcia takiego udzielono Irlandii (na kwotę 85 mld EUR), w kwietniu 2011 roku Portugalii (na kwotę 75 mld EUR), a w maju 2013 roku otrzymał je Cypr (na kwotę 10 mld EUR). Na wykresie 8.5 ukazano zmiany w zakresie rentowności obligacji rządowych najbardziej zadłużonych państw strefy euro i Niemiec. Wyraźnie widać, że kłopoty Grecji uruchomiły proces spadku zaufania na rynku długu suwerennego wśród pozostałych państw zadłużonych. Okres od maja 2009 roku do końca 2012 roku wyznacza ostrą fazę kryzysu zadłużenia w strefie euro – rentowności greckie skoczyły do niemal 30% (co oznacza spadek cen greckich obligacji o 80–85%), a rentowności

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

irlandzkie i portugalskie do 15%. W lecie 2011 roku strefa euro podjęła w uzgodnieniu z inwestorami prywatnymi decyzję o częściowej utracie wartości greckiego długu znajdującego się w ich rękach. Porozumienie to nazwano dość eufemistycznie „zaangażowaniem" inwestorów prywatnych (*private sector involvement* – PSI). Była to bardzo nieprzyjemna niespodzianka dla rynków finansowych, ponieważ decyzja o PSI oznaczała *de facto* uznanie częściowego bankructwa suwerennego Grecji. Związane z tym nasilenie efektu „zarażenia" spowodowało, że pod presją znalazły się również obligacje włoskie i hiszpańskie, a przejściowo także dług francuski. W tym samym czasie silnie spadały rentowności obligacji niemieckich, które były masowo nabywane przez inwestorów jako bezpieczne aktywa.

Wykres 8.5. Rentowności 10-letnich obligacji rządowych wybranych krajów strefy euro oraz Polski w latach 2007–2015 (dane w %)



Uwaga: rentowności obligacji greckich w lipcu 2015 roku nie były kwotowane z powodu zawieszenia działalności banków greckich w okresie ostrego kryzysu związanego z negocjowaniem trzeciego pakietu pomocowego dla Grecji.

Źródło: opracowanie własne na podstawie danych EBC.

Kryzys został zażegnany dopiero w drugiej połowie 2012 roku dzięki determinacji Europejskiego Banku Centralnego, który zapowiedział, że uczyni wszystko, aby uratować wspólną walutę (pisaliśmy o tym w rozdziale 5), ogłaszając gotowość

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

do uruchomienia programu nieograniczonego skupu obligacji państw zadłużonych (program OMT). Deklaracja EBC zadziałała jako „zapora ogniowa" (*firewall*), oddzielająca kraje ewidentnie niewypłacalne (Grecja) od krajów fundamentalnie wypłacalnych, choć przeżywających kłopoty z utrzymaniem płynności (Portugalia, Włochy, Hiszpania). W ten sposób stworzono barierę, która chroniła rynki długu pozostałych państw strefy euro przed „zarażeniem" ze strony Grecji po ogłoszeniu decyzji o obciążeniu inwestorów prywatnych kosztami częściowej redukcji długu greckiego. W rezultacie od trzeciego kwartału 2012 roku zaufanie stopniowo powraca na rynek długu suwerennego państw strefy euro i sytuacja ulega stabilizacji.

8.9.2. Czy brak konwergencji mógł być przyczyną kryzysu w strefie euro?

Konwergencja jest pojęciem mającym wiele różnych znaczeń. W kontekście dyskusji o unii walutowej wymienia się trzy rodzaje konwergencji. Po pierwsze, może ona oznaczać stopniowe zbliżanie się do siebie wartości pewnych nominalnych zmiennych ekonomicznych, takich jak na przykład poziom cen, stopa inflacji czy rynkowe stopy procentowe wśród poszczególnych krajów członkowskich unii walutowej. Tak definiowana jest konwergencja **nominalna** w Traktacie o Funkcjonowaniu Unii Europejskiej, a spełnienie określonych kryteriów konwergencji nominalnej jest warunkiem wejścia kraju członkowskiego UE do strefy euro. Konwergencja **realna** jest zaś najczęściej rozumiana jako trwałe zbliżanie się wartości PKB na mieszkańca (według parytetu siły nabywczej) w uboższych krajach członkowskich unii walutowej do poziomu PKB na mieszkańca w krajach najbardziej rozwiniętych. Polega ona więc na „doganianiu" krajów zamożnych przez kraje uboższe pod względem poziomu dochodu przypadającego na mieszkańca, co wiąże się ze zmniejszaniem rozpiętości dochodowych między krajami członkowskimi unii walutowej[33]. Temu zjawisku „doganiania" krajów bogatszych przez kraje uboższe (*catching up*) towarzyszy najczęściej proces szybszego wzrostu cen w krajach „doganiających", zwany efektem Balassy-Samuelsona (zob. rozdział 6,

[33] W literaturze ekonomicznej wyróżnia się dwa typy konwergencji realnej. Konwergencja typu beta (*β-convergence*) ma miejsce, gdy kraje znajdujące sią na niższym poziomie rozwoju rozwijają się szybciej niż kraje zamożne, czyli następuje proces „doganiania" pod względem poziomu dochodu na mieszkańca (*catching up*). Konwergencja typu sigma (*σ-convergence*) zachodzi natomiast wówczas, gdy maleją rozpiętości w poziomie dochodu na mieszkańca między poszczególnymi krajami. Konwergencja typu beta jest warunkiem wystąpienia konwergencji typu sigma (zob. Matkowski i Próchniak 2009; Próchniak i Rapacki 2009).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

punkt 6.8.3). Stąd też ścisły związek między konwergencją realną a konwergencją nominalną (Rogoff 1996). Wreszcie, konwergencja **strukturalna** oznacza zbliżenie struktur gospodarek państw unii walutowej, zarówno pod względem produkcji, jak i handlu zagranicznego. Konwergencja strukturalna oznacza z reguły silniejszą synchronizację cyklu koniunkturalnego, a to z kolei zmniejsza prawdopodobieństwo wystąpienia wstrząsów asymetrycznych w ramach unii walutowej (Buti i Turrini 2015).

Krytycy idei integracji walutowej w Europie twierdzili, że o ile zapisy traktatowe wymagały od krajów członkowskich pewnego zakresu konwergencji nominalnej *ex-ante*, o tyle pod względem stopnia konwergencji realnej i konwergencji strukturalnej kraje UE pozostawały bardzo zróżnicowane, a zarazem brakowało skutecznych mechanizmów umożliwiających przeprowadzenie konwergencji (Buti i Sapir 1998). W rezultacie kraje członkowskie strefy euro – zdaniem krytyków – nie stanowiły optymalnego obszaru walutowego. W takiej sytuacji mogły pojawić się dwa zagrożenia: jedno dotyczyło braku konwergencji realnej, a drugie braku konwergencji strukturalnej.

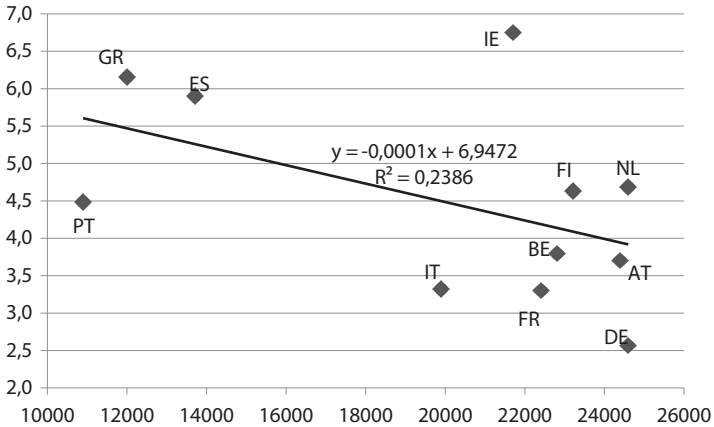
W warunkach braku konwergencji **realnej** i przy dużych różnicach w poziomie dochodu przypadającego na mieszkańca jednolita polityka pieniężna nie jest w stanie zapewnić zbliżonych stóp inflacji we wszystkich krajach członkowskich unii walutowej. Przy danych jednakowych nominalnych stopach procentowych w całej unii walutowej kraje znajdujące się na niższym poziomie rozwoju będą notowały wyższe stopy inflacji niż kraje najbardziej rozwinięte. Oznacza to, że w warunkach jednakowych **nominalnych** stóp procentowych w unii walutowej **realne** stopy procentowe stają się niższe w krajach „doganiających" niż w krajach rozwiniętych. W rezultacie wspólna polityka pieniężna będzie bardziej ekspansywna w państwach „doganiających" i bardziej restrykcyjna w krajach rozwiniętych. W pierwszej grupie krajów wzrost gospodarczy oraz inflacja zaczną przyspieszać, podczas gdy w drugiej grupie państw wskaźniki te ulegną spowolnieniu. W efekcie może pojawić się rosnąca nierównowaga, wobec czego niemożliwe stanie się zapewnienie realnej oraz nominalnej konwergencji w unii walutowej.

Jednym z pierwszych ekspertów, który już w 1986 roku dostrzegł to niebezpieczeństwo, był brytyjski ekonomista Alan Walters, wcześniej doradca ekonomiczny premier Margaret Thatcher. Tzw. krytyka Waltersa (*Walters critique*), w wersji oryginalnej odnosząca się do skuteczności sztywnych kursów w ramach Europejskiego Systemu Walutowego, stała się następnie w latach 90. popularnym argumentem przemawiającym na rzecz pozostania Wielkiej Brytanii poza strefą euro oraz

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

przeciwko utworzeniu unii walutowej przez znajdujące się na różnym poziomie rozwoju gospodarczego kraje Europy kontynentalnej (Miller i Sutherland 1991).

Wykres 8.6. Konwergencja realna w strefie euro: poziom PKB/mieszkańca w 1998 roku (euro) i średnie tempo wzrostu PKB/mieszkańca według PPS w latach 1998–2008 (dane w %)



Źródło: opracowanie własne na podstawie danych Eurostatu.

Badania empiryczne nie potwierdziły jednak hipotezy Waltersa. Rzeczywiście zaobserwowano, że w krajach o początkowo wyższej inflacji i niższych realnych stopach procentowych tempo wzrostu PKB było szybsze niż w pozostałych państwach, lecz wbrew przepowiedniom Waltersa różnice w stopach inflacji między poszczególnymi krajami nie rosły, a spadały. W ciągu kilku lat po utworzeniu unii walutowej w Europie doszło do wyraźnej nominalnej konwergencji w zakresie inflacji (Mongelli i Wyplosz 2008). Równolegle postępowała również konwergencja realna. Wykres 8.6 przedstawia zależność występującą między wielkością PKB na mieszkańca w 1998 roku (oś pozioma) a średnim tempem wzrostu PKB (%) w 11 krajach członkowskich strefy euro w latach 1998–2008. Jak można zauważyć, mniej zamożne państwa strefy euro rozwijały się zasadniczo szybciej niż kraje bogatsze. W rezultacie rozpiętości dochodowe między państwami członkowskimi stopniowo zmniejszały się[34]. Dane te rzeczywiście mogą sugerować, że szybsze tempo wzrostu gospodarczego w Grecji, Hiszpanii czy Irlandii mogło wiązać się

[34] Stosunek odchylenia standardowego od średniego dochodu wśród 11 krajów strefy euro zmniejszył się w latach 1998–2008 z 0,25 do 0,23.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

z szybszym wzrostem zadłużenia, lecz w żaden sposób nie wskazują na to, że brak konwergencji realnej czy też duże rozpiętości dochodowe w momencie powstania unii walutowej mogły stać się przyczyną kryzysu[35]. Jeśli istniał jakiś związek między tymi zjawiskami, to raczej polegał on na tym, że to szybkie tempo konwergencji realnej mogło sprzyjać wzrostowi zadłużenia.

Drugie z wymienianych zagrożeń wiązało się z brakiem konwergencji strukturalnej. Jak pamiętamy z wcześniejszych rozważań, unia walutowa ma ekonomiczny sens wówczas, gdy prawdopodobieństwo wystąpienia szoków asymetrycznych w jej ramach jest niewielkie. Tymczasem strukturalne zróżnicowanie poszczególnych państw członkowskich zwiększało potencjał dla wystąpienia wstrząsów asymetrycznych. W szczególności zaś, jak wskazywano od samego początku, tzw. kraje „peryferyjne" unii, wykazujące odmienne charakterystyki strukturalne w stosunku do krajów „centrum", mogą być w większym stopniu narażone na szoki asymetryczne. Badania przeprowadzone w latach 90. wskazywały na znaczny potencjał występowania szoków asymetrycznych w krajach „peryferyjnych" (zob. m.in. Bayoumi i Eichengreen 1992, Sørensen i Yosha 1998, Arreaza i in. 1998). Liczono jednak na to, że wprowadzenie wspólnej waluty i związany z tym dalszy rozwój w zakresie integracji oraz wzajemnego handlu spowoduje silniejszą konwergencję strukturalną między państwami członkowskimi, a także przełoży się na spadek prawdopodobieństwa wystąpienia wstrząsów asymetrycznych w strefie euro (Frankel i Rose 1996). Ze wspomnianych badań wynikało także, że w obliczu zagrożenia szokami asymetrycznymi państwa strefy euro powinny stworzyć mechanizmy zabezpieczające przed skutkami tych szoków, w tym zwłaszcza systemy podziału ryzyka (*risk sharing*) – zarówno prywatne, jak i publiczne – które pozwalałyby niwelować te skutki poprzez wygładzanie dochodów i konsumpcji w krajach dotkniętych wstrząsami.

Jedną z możliwych metod oszacowania skali wstrząsów asymetrycznych jest analiza odchyleń zaobserwowanych stóp wzrostu PKB od trendu, czyli od średnich stóp wzrostu w wybranej grupie krajów w danym okresie (zob. IMF 2013). Odchylenia te, reprezentujące wstrząsy wpływające na wzrost PKB, są następnie

[35] Pogląd mówiący o tym, że osiągnięcie pełnej konwergencji realnej (czyli uzyskanie przez kraj kandydujący do strefy euro poziomu PKB na mieszkańca porównywalnego z poziomem najbardziej rozwiniętych krajów unii walutowej) jest niezbędnym warunkiem wejścia do strefy euro pozostaje również popularny wśród niektórych polskich ekonomistów (zob. np. Kaźmierczak 2013). Trudno jednak bronić tego stanowiska zarówno ze względu na doświadczenia uboższych krajów unii walutowej, jak i w świetle teorii integracji walutowej. Wejście do unii walutowej kraju „doganiającego" przyspiesza jego wzrost gospodarczy i tym samym konwergencję realną o tyle, o ile spełnione są warunki wysokiego stopnia integracji w ramach handlu oraz wysokiego stopnia synchronizacji cyklu koniunkturalnego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

dekomponowane na składnik wspólny dla całej grupy krajów oraz na składnik indywidualny w przypadku konkretnego państwa (*country-specific*). Ten ostatni element reprezentuje wstrząsy asymetryczne. W ramce 8.11 przedstawiono szczegółową metodologię i wyniki badania mającego na celu zidentyfikowanie wstrząsów asymetrycznych w strefie euro w latach 1996–2015.

Ramka 8.11. Wstrząsy asymetryczne w strefie euro: konwergencja czy dywergencja strukturalna?

Badaniem objęto zmiany PKB w 19 krajach członkowskich strefy euro w okresie 1996–2015, wykorzystując dane Eurostatu. Punktem wyjścia jest wyliczenie średnich stóp wzrostu PKB na podstawie równań trendu w przypadku poszczególnych krajów oraz odnośnie do strefy euro jako całości w badanym okresie, zgodnie z równaniem:

$$y_{i,t} = y_{i,0}\left(1+\hat{r}_i\right)^t, \quad \text{dla } i = 1, 2,..., 17 \tag{8.18}$$

gdzie $\hat{r}_i$ jest średnią stopą wzrostu PKB w kraju *i*. Na tej podstawie obliczono odchylenia rzeczywistych stóp wzrostu PKB od wartości średnich w przypadku poszczególnych krajów. Odchylenia te – reprezentujące wstrząsy „brutto" (symetryczne i asymetryczne) wpływające na zmienność stóp wzrostu PKB w poszczególnych państwach – stanowią sumę dwóch składników. Pierwszy z nich jest efektem wstrząsów wspólnych dla całej strefy euro i pozostaje równy odchyleniom stopy wzrostu PKB dla całej strefy euro w danym roku od średniej stopy wzrostu PKB w przypadku całej strefy euro. Składnik ten mierzy wstrząsy symetryczne. Drugi z nich – składnik rezydualny – powstaje za sprawą wstrząsów indywidualnych występujących w poszczególnych krajach. Ten drugi składnik mierzy wstrząsy asymetryczne – zarówno dodatnie, jak i ujemne. Oznaczmy odchylenia rzeczywistej stopy wzrostu PKB od trendu w strefie euro oraz w kraju *i* odpowiednio jako $u_{EA,t}$ oraz $u_{i,t}$, a średnią stopę wzrostu PKB w przypadku całej strefy euro jako $\hat{r}_{EA}$. Otrzymujemy wówczas:

$$r_{i,t} - \hat{r}_i = u_{EA,t} + u_{i,t}, \tag{8.19}$$

gdzie:

$$u_{i,t} = r_{i,t} - \hat{r}_i - u_{EA,t}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

oraz

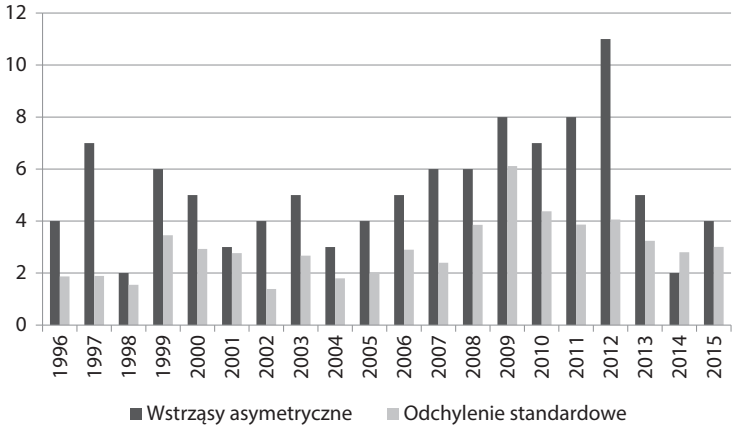
$$u_{EA,t} = r_{EA,t} - \widehat{r}_{EA}$$

i ostatecznie:

$$u_{i,t} = \left(r_{i,t} - \widehat{r}_i\right) - \left(r_{EA,t} - \widehat{r}_{EA}\right). \tag{8.20}$$

Wielkość $u_{i,t}$ reprezentuje indywidualny składnik odchylenia stopy wzrostu PKB w kraju *i* w roku *t* od średniej. Równanie (8.20) zostało wykorzystane do wyliczenia wielkości $u_{i,t}$, przy czym umownie przyjęto, że wstrząsem asymetrycznym jest wielkość $u_{i,t}$ przekraczająca 2 punkty procentowe[36]. Otrzymane wyniki zostały przedstawione na wykresach 8.7 i 8.8[37].

Wykres 8.7. Liczba wstrząsów asymetrycznych oraz odchylenia standardowe w przypadku krajów strefy euro w latach 1996–2015



Uwaga: wstrząsy asymetryczne zdefiniowane jako wielkości $u_{i,t}$ przekraczające 2 punkty procentowe; odchylenie standardowe przekrojowe w przypadku wielkości $u_{i,t}$ dotyczy wszystkich krajów w danym roku *t* (*cross-sectional standard deviation*).

Źródło: opracowanie własne.

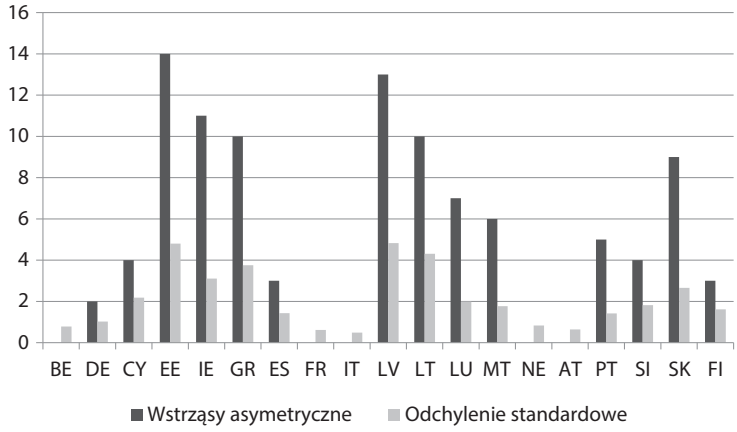
[36] Założenie to ma oczywiście charakter umowny.

[37] Przedstawiona metodologia została zastosowana m.in. przez Międzynarodowy Fundusz Walutowy (IMF, 2013). Trzeba jednak zauważyć, że miara zastosowana przez MFW w istocie mierzy stopień dyspersji stóp wzrostu gospodarczego, i nie w każdym przypadku pozwala identyfikować rzeczywisty wstrząs asymetryczny. Na przykład, gdy oba elementy po prawej stronie równania (8.20) będą dodatnie (lub ujemne), wyrażenie $u_{i,t}$ mierzy po prostu różnice w odchyleniach stóp wzrostu od trendu, ale nie pokazuje asymetrii *sensu stricto*. Alternatywne miary wstrząsów asymetrycznych, pozbawione tej wady, proponuje Rosati (2016).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Wykres 8.7 pokazuje, że liczba wstrząsów asymetrycznych – zarówno dodatnich, jak i ujemnych – w grupie krajów objętych badaniem bynajmniej nie zmalała po utworzeniu strefy euro w 1999 roku, a nawet nieco wzrosła w latach 2009–2012. Zmiany odchyleń standardowych potwierdzają tę obserwację – przekrojowe odchylenia standardowe (*cross-sectional*) odnośnie poszczególnych lat nie malały po utworzeniu strefy euro, a w latach 2008–2011 nawet wzrosły. Jednocześnie w latach 2013–2015 nastąpił wyraźny spadek wartości obu zmiennych.

Wykres 8.8. Liczba wstrząsów asymetrycznych oraz odchylenia standardowe w przypadku poszczególnych krajów strefy euro w latach 1996–2015



Uwaga: wstrząsy asymetryczne zdefiniowane jako wielkości $u_{i,t}$ przekraczające 2 punkty procentowe; odchylenia standardowe w przypadku wielkości $u_{i,t}$ dotyczą poszczególnych krajów w całym okresie 1990–2015.

Źródło: opracowanie własne.

Z kolei wykres 8.8 pokazuje, że w strefie euro rzeczywiście występuje wyraźny podział na dwie odrębne grupy krajów członkowskich – kraje „twardego rdzenia" („centrum") i kraje pozostałe („peryferie"). W pięciu państwach (Belgia, Francja, Włochy, Holandia i Austria) w badanym okresie nie wystąpiły wstrząsy asymetryczne, co świadczy o wysokim stopniu strukturalnej konwergencji tych gospodarek. Do grupy krajów „centrum" można też niewątpliwie zaliczyć Niemcy, gdzie pojawiły się tylko dwa wstrząsy, oraz – w dużym stopniu – także Hiszpanię i Finlandię (3 wstrząsy). Na drugim biegunie znajdują się państwa, w których wstrząsy asymetryczne były częste lub bardzo częste – Słowacja (9), Litwa (10), Grecja (10), Irlandia (11), Łotwa (13) i Estonia (14). Między tymi dwoma grupami znajduje się grupa krajów o średniej liczbie wstrząsów asymetrycznych, obejmująca Cypr,

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Luksemburg, Malta, Portugalia i Słowenia[38]. Dane te pokazują, że stopień konwergencji strukturalnej w strefie euro pozostaje zróżnicowany. Również wielkości odchyleń standardowych zmiennej $u_{i,t}$ odnośnie do poszczególnych krajów potwierdzają wysoki stopień konwergencji strukturalnej w pierwszej grupie państw (obejmującej w tym przypadku Belgię, Niemcy, Hiszpanię, Francję, Włochy, Luksemburg, Maltę, Holandię, Austrię, Portugalię i Słowenię) i niski stopień konwergencji strukturalnej w pozostałych krajach[39].

Trzeba jednak pamiętać, że na występowanie wstrząsów asymetrycznych ma także wpływ polityka fiskalna. Kraje mniej zadłużone, posiadające odpowiedni stopień swobody fiskalnej (*fiscal space*), mogły reagować na wstrząs asymetryczny przy pomocy wzrostu wydatków publicznych lub obniżenia podatków. W przypadku tych krajów wstrząsy asymetryczne byłyby więc mniejsze, lub zostałyby całkowicie zneutralizowane. Natomiast w przypadku krajów zadłużonych, pozbawionych przestrzeni fiskalnej, wstrząsy asymetryczne ujawniały się w pełni. W rezultacie przedstawione dane dotyczące liczby wstrząsów asymetrycznych w poszczególnych krajach strefy euro mogą zaniżać rzeczywistą liczbę tych wstrząsów w krajach posiadających dużą przestrzeń fiskalną, w stosunku do pozostałych krajów. W świetle tej obserwacji stopień konwergencji strukturalnej w ramach strefy euro może być w rzeczywistości wyższy niż wynikałoby to z danych dotyczących różnic w stopach wzrostu PKB.

8.9.3. Co ujawnił kryzys? Niedostosowanie instytucjonalne w strefie euro

Kryzys lat 2009–2012 ujawnił cztery główne słabości systemu zarządzania ekonomicznego w strefie euro:

– istniejący mechanizm dyscypliny budżetowej w postaci Paktu Stabilności i Wzrostu okazał się nieskuteczny w zapobieganiu nadmiernemu zadłużeniu publicznemu w państwach członkowskich;

[38] Wysoka liczba wstrząsów asymetrycznych w przypadku Luksemburga (12) może wynikać z niewielkich rozmiarów tej gospodarki i dużego udziału sektora usług – zwłaszcza usług finansowych – w PKB.

[39] Należy pamiętać, że otrzymane wyniki zależą też od wielkości poszczególnych gospodarek. Im większa jest dana gospodarka, tym silniejszy wywiera wpływ na średnią stopę wzrostu w całej strefie euro i tym niższe – *ceteris paribus* – stają się odchylenia obserwacji krajowych od średniej w strefie euro.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

- w ramach systemu zarządzania gospodarczego w Unii nie istniał system wspólnotowej kontroli i nadzoru nad sektorem finansowym, w tym bankowym;
- Unia Europejska nie dysponowała skutecznym mechanizmem reagowania na kryzys zadłużenia;
- nieskuteczny okazał się mechanizm koordynacji polityki gospodarczej.

Na brak dyscypliny budżetowej w niektórych państwach strefy euro, mimo formalnego obowiązywania reguł fiskalnych zapisanych w SGP, złożyło się wiele przyczyn. Po pierwsze, w SGP brakowało silnych bodźców skłaniających państwa członkowskie do przestrzegania reguł fiskalnych. Zasady nadzoru i kontroli w ramach części prewencyjnej Paktu nie zapobiegały naruszaniu reguł, a procedura EDP w części korekcyjnej Paktu była w praktyce rozciągana na bardzo długie okresy czasu (teoretycznie nawet do 9 lat – zob. Calmfors 2005). Przejściowe naruszenie reguł nie powodowało żadnych istotnych negatywnych konsekwencji w przypadku danego państwa członkowskiego, a kary finansowe nigdy nie zostały zastosowane. Po drugie, procedura EDP była w dużym stopniu dyskrecjonalna (uznaniowa) i w efekcie nadmiernie upolityczniona, ponieważ Komisja Europejska wnioskowała jedynie do Rady UE o rozpoczęcie procedury, lecz ostateczną decyzję o tym, czy istnieje „nadmierny deficyt" w rozumieniu Traktatu podejmowała Rada kwalifikowaną większością głosów[40]. Po trzecie, sankcje odnosiły się tylko do naruszenia reguły fiskalnej, dotyczącej dopuszczalnego deficytu, natomiast w przypadku naruszenia pozostałych dwóch reguł – poziomu długu i średniookresowej równowagi budżetowej – żadne sankcje nie zostały przewidziane ani w Traktacie, ani w SGP. Wreszcie, system wielostronnego nadzoru i kontroli nie zapobiegał stosowaniu kreatywnej księgowości budżetowej, a niekiedy wręcz fałszowaniu danych statystycznych przez państwa członkowskie (więcej na ten temat zob. Rosati 2010, 2011)[41].

Lecz nie tylko system kontroli i nadzoru nad polityką fiskalną w ramach SGP okazał się nieskuteczny w zapewnieniu dyscypliny budżetowej. Co ciekawe, również rynki finansowe zawiodły przy wycenie długu suwerennego poszczególnych państw. Wykres 8.5 pokazuje, że praktycznie do połowy 2008 roku inwestorzy traktowali dług rządu greckiego za niemal równie bezpieczną inwestycję jak dług rządu

[40] Co wymagało w obowiązującym wówczas systemie głosowania uzyskania większości ponad 74% głosów ważonych.

[41] Najbardziej spektakularnym przykładem było ukrywanie przez wiele lat faktycznego stanu finansów publicznych przez władze greckie, lecz także w innych krajach dochodziło do przypadków „twórczej księgowości" (np. we Włoszech, a wśród krajów nowo przyjętych na Węgrzech).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

niemieckiego – rentowności obligacji greckich były jedynie o 30–50 punktów bazowych wyższe niż w przypadku obligacji niemieckich. Aż do początku 2009 roku trzy największe międzynarodowe agencje ratingowe (Standard and Poor's, Moody's, Fitch) nadal przyznawały Grecji wysoki rating inwestycyjny na poziomie A+ lub A (zob. Dziawgo 2012). U podstaw tego niefrasobliwego zachowania leżało błędne założenie o solidarnym pokrywaniu zobowiązań w strefie euro i – prawdopodobnie – nadmierna wiara w skuteczność unijnych reguł fiskalnych. Nie bez znaczenia był także fakt ukrywania rzeczywistych rozmiarów deficytu przez kolejne greckie rządy. Dopiero od stycznia 2009 roku nastąpił wyraźny wzrost rentowności obligacji greckich, a od marca 2010 roku mamy już do czynienia z paniczną wyprzedażą długu greckiego.

Kolejną wadą systemu zarządzania gospodarczego w Unii Europejskiej był brak jakiegokolwiek systemu wspólnotowej kontroli i nadzoru nad sektorem finansowym, w tym zwłaszcza bankowym, który zapobiegałby nadmiernej ekspansji kredytu bankowego w poszczególnych państwach członkowskich. Na obszarze unii walutowej, gdzie funkcjonuje jedna waluty i jednolita polityka pieniężna, istniało kilkanaście odrębnych narodowych systemów regulacji oraz nadzoru bankowego i finansowego. Zachęcało to do arbitrażu regulacyjnego i lokowania kapitałów w krajach o bardziej liberalnym nadzorze bankowym. W rezultacie na przykład w bankach irlandzkich depozyty złożone w 2010 roku przez nierezydentów były większe niż depozyty rezydentów (odpowiednio 203 mld EUR i 166 mld EUR), a ich łączna suma niemal trzykrotnie przekraczała irlandzki poziom PKB[42].

Zarazem międzynarodowe grupy bankowe działały w wielu państwach członkowskich jednocześnie, otwierając swoje oddziały lub powołując odrębne spółki na podstawie miejscowego prawa. W tych warunkach trudno było ocenić skalę ryzyka systemowego związanego z działalnością poszczególnych banków. W rezultacie w wielu przypadkach krajowe banki gromadziły w swoich portfelach nadmiernie ryzykowne aktywa, finansując spekulacyjne inwestycje na rynkach nieruchomości (głównie w Irlandii i Hiszpanii, a w mniejszym stopniu także w innych krajach), podczas gdy ryzyko systemowe tych inwestycji rozprzestrzeniało się na systemy bankowe innych państw. Groźba załamania się systemów bankowych zmusiła rządy wielu krajów strefy euro do udzielenia bankom pomocy finansowej na dużą skalę, co pogłębiło kryzys w zakresie finansów publicznych. Niepewna sytuacja wielu banków stanowi nadal potencjalne zagrożenie dla stabilności budżetów narodowych

[42] "The Economist", 27 November, 2010.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wielu państw (Włochy, Hiszpania, Cypr, Austria). Również odwrotnie, ryzyko niewypłacalności niektórych krajów grozi upadłością banków, które zainwestowały w papiery dłużne tych państw (Grecja, Portugalia). Wzajemne powiązania rządów i banków utworzyły błędne koło, które niełatwo było rozerwać.

Trzecią niedoskonałością systemu zarządzania gospodarczego w Unii Europejskiej był brak skutecznego mechanizmu reagowania na kryzys zadłużenia – zapisy traktatowe nie przewidywały w ogóle sytuacji, w której kraj członkowski strefy euro mógłby mieć kłopoty z wypłacalnością, a w skrajnym przypadku nawet zbankrutować. Z dzisiejszej perspektywy widać, że twórcom unii walutowej po prostu zabrakło wyobraźni. W rezultacie reakcje UE na kryzys miały nie tylko charakter doraźny, lecz były także z reguły spóźnione i niewystarczające, ponieważ wobec braku wcześniej uzgodnionych procedur działania i przy sprzecznych interesach poszczególnych państw antykryzysowe decyzje o udzieleniu pomocy finansowej wymagały każdorazowo żmudnych politycznych negocjacji i z reguły kończyły się ułomnymi kompromisami.

Czwartą wreszcie wadą systemu zarządzania okazała się nieskuteczność mechanizmu koordynacji polityki gospodarczej w postaci BEPG. Nie uwzględniał on bowiem żadnych instrumentów skłaniających państwa członkowskie do korygowania narastającej nierównowagi i podejmowania niezbędnych reform. Konkretne rekomendacje w ramach systemu ogólnych wytycznych były traktowane przez poszczególne państwa jako zbiór niewiążących sugestii.

Wszystkie cztery wymienione mankamenty świadczyły o poważnych lukach zarówno w instytucjonalnej strukturze UE, jak i w systemie przestrzegania unijnych przepisów. UGW została stworzona na „dobrą pogodę", a założycielom unii walutowej zabrakło wyobraźni i niezbędnej przezorności. Wady systemu koordynacji polityki w Unii Europejskiej wyszły na jaw wówczas, gdy w Europę uderzył w 2008 roku kryzys finansowy (Shambaugh 2012).

W okresie od stycznia 2010 roku do czerwca 2012 roku przywódcy UE odbyli 18 spotkań Rady Europejskiej, a ponadto szefowie państw lub rządów krajów strefy euro zebrali się na 10 spotkaniach – wszystkie w całości lub w przeważającej mierze zostały poświęcone walce z kryzysem. Komunikaty ogłaszane po każdym szczycie pełne były krzepiących zapewnień o determinacji w walce z kryzysem i zapowiedzi podjęcia skutecznych działań. W praktyce okazywało się jednak, że reakcje zarówno samej UE, jak i krajów strefy euro były z reguły spóźnione, połowiczne i doraźne – brakowało całościowej strategii walki z kryzysem, a osiągane w toku wielogodzinnych debat kompromisy nie zdołały przywrócić zaufania na rynkach

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

finansowych. Dobrym przykładem nieskuteczności podejmowanych działań jest przebieg zjawisk kryzysowych w Grecji (ramka 8.12), gdzie przez wiele miesięcy nie udało się opanować sytuacji. W rezultacie kryzys systematycznie się pogłębiał i rozprzestrzeniał na coraz to nowe kraje.

Ramka 8.12. Przebieg kryzysu zadłużenia w Grecji

W październiku 2009 roku, po wyborach parlamentarnych, nowy rząd grecki ujawnił fakt fałszowania danych dotyczących wielkości deficytu budżetowego przez poprzedni rząd i ogłosił, że deficyt w 2009 roku wyniesie 12,7% PKB zamiast zaplanowanych 6,0% PKB (ostatecznie wyniósł 13,5%). Gwałtowny spadek zaufania inwestorów do papierów skarbowych Grecji i związany z tym odpływ kapitału spowodował wzrost spreadu na rentownościach ze 100–120 bps ponad poziom rentowności obligacji niemieckich do 250–300 bps w grudniu 2009 i do 400–500 bps w styczniu 2010 roku. W lutym 2010 roku Rada ECOFIN ogłosiła, że UE udzieli Grecji wsparcia. Ponieważ wśród państw członkowskich nie było jednak zgody co do warunków i zakresu pomocy, rentowność obligacji greckich nadal rosła. Równolegle, w reakcji na ogólną niepewność i niepokój na rynkach, zaczęła wzrastać rentowność obligacji portugalskich, hiszpańskich i irlandzkich.

Na początku maja 2010 roku rentowności wzrosły do 10–12%, co praktycznie uniemożliwiało zaciąganie pożyczek na rynkach finansowych. W obliczu tej sytuacji rząd grecki zwrócił się oficjalnie o pomoc do UE. W związku z tym 2 maja 2010 państwa strefy euro podjęły decyzję o przyznaniu Grecji dwustronnych pożyczek w ramach doraźnego programu pomocowego koordynowanego przez Komisję Europejską (tzw. *Greek Loan Facility* – GLF) na kwotę 80 mld EUR, do wykorzystania w ratach w okresie od maja 2010 roku do czerwca 2013 roku[43]. Pomoc finansowa strefy euro została uzupełniona przez pożyczkę z MFW na kwotę 30 mld EUR w ramach porozumienia *stand-by arrangement* (SBA). W zamian za otrzymaną pomoc finansową Grecja miała dokonać oszczędności budżetowych rzędu 12,5% PKB (30 mld EUR) w latach 2011–2013 i obniżyć deficyt do 3% w 2014 roku. Jednocześnie rząd grecki zobowiązał się do przeprowadzenia szeregu

[43] Kwota ta ostatecznie była niższa o 2,7 mld EUR, ponieważ Słowacja odmówiła uczestniczenia w programie, a Portugalia i Irlandia musiały zrezygnować w związku z objęciem ich odrębnymi programami pomocy.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

reform strukturalnych, w tym reformy emerytalnej, uelastycznienia rynku pracy, poprawy ściągalności podatków i usunięcia barier na rynkach produktów i usług.

Równolegle (9 maja 2010 roku) państwa strefy euro podjęły decyzję o utworzeniu europejskiego mechanizmu stabilizacyjnego (*European Stability Mechanism*), dysponującego łączną kwotą 750 mld EUR i obejmującego Europejski Instrument Stabilności Finansowej (*European Financial Stability Facility* – EFSF) na kwotę 440 mld EUR, Europejski Mechanizm Stabilności Finansowej (*European Financial Stabilisation Mechanism* – ESFM) na kwotę 60 mld EUR, oraz pożyczki z MFW na kwotę do 250 mld EUR).

Realizacja programu dostosowawczego w Grecji odbiegała jednak od przyjętych założeń zarówno ze względu na niekorzystne warunki zewnętrzne, jak i – zwłaszcza – ze względu na opóźnienia w wypełnianiu przez Grecję podjętych zobowiązań. W szczególności nie udało się skierować gospodarki na ścieżkę wzrostu gospodarczego. W rezultacie w maju 2012 roku zatwierdzono drugi program pomocowy na kwotę 130 mld EUR, finansowany tym razem z Europejskiego Instrumentu Stabilności Finansowej (*European Financial Stability Facility* – EFSF) i z Europejskiego Mechanizmu Stabilności Finansowej. Kluczowym elementem programu była częściowa redukcja zadłużenia, która obciążyła prywatnych inwestorów posiadających obligacje greckie (na kwotę około 50% zdyskontowanej wartości długu)[44]. Miało to ułatwić Grecji wyjście na ścieżkę wzrostu, lecz w istocie oznaczało wymuszone częściowe bankructwo suwerenne tego państwa.

Poprawa sytuacji gospodarczej w Grecji przebiegała tymczasem bardzo wolno. Dopiero w 2014 roku udało się zahamować spadek produkcji, lecz dług publiczny wzrósł w międzyczasie do ponad 170% PKB. Pojawiła się realna groźba chaotycznego bankructwa i wyjścia Grecji ze strefy euro[45]. Odpowiedzią był trzeci program pomocowy (na kwotę 86 mld EUR), ostatecznie przyjęty w lipcu 2015 roku po wielu miesiącach negocjacji.

[44] Zob. wspomniany wyżej program *private sector involvement*.

[45] Wyjście Grecji ze strefy euro musiałoby *de facto* nastąpić w momencie wstrzymania przez EBC finansowania greckiego systemu bankowego w ramach awaryjnego planu dostarczania płynności (*Emergency Liquidity Assistance* – ELA), i to niezależnie od stanowiska rządu greckiego. W takiej bowiem sytuacji rząd grecki musiałby w trybie natychmiastowym dostarczyć alternatywne instrumenty zapewniające płynność banków – najpierw w doraźnej postaci weksli (*promissory notes*) oraz instrumentów typu IOU (*I owe you*), a w końcu zmuszony byłby wprowadzić równoległą walutę.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

8.9.4. Brak mechanizmu podziału ryzyka

Jak już wspomniano, w ramach unii walutowej kraje członkowskie tracą możliwość wykorzystania instrumentów polityki pieniężnej i polityki kursowej w ramach niwelowania skutków wstrząsów asymetrycznych. Mogą wprawdzie nadal stosować w tym celu politykę fiskalną, lecz tylko w takim stopniu, w jakim pozwala im na to posiadana przestrzeń fiskalna (*fiscal space* – zob. rozdział 3, punkt 3.7.5). Jak już jednak wiemy, w krajach najbardziej zadłużonych przestrzeń fiskalna była bardzo ograniczona, co praktycznie pozbawiało te państwa możliwości zastosowania antycyklicznej polityki fiskalnej.

Lecz, jak wskazują niektórzy ekonomiści, nawet gdy w danym kraju istnieje dostatecznie duża przestrzeń fiskalna, narodowa polityka pieniężna może nie być w stanie w pełni niwelować skutków wstrząsów asymetrycznych (Baldacci i in. 2009). Ekonomiści ci wskazują, że w takiej sytuacji w ramach unii walutowej powinny istnieć inne, dodatkowe mechanizmy łagodzenia konsekwencji wstrząsów asymetrycznych, zwłaszcza gdy pozostałe „klasyczne" kryteria Mundella, takie jak mobilność siły roboczej i elastyczność płac, są spełnione tylko w ograniczonym zakresie (Atkeson i Bayoumi 1993, Bayoumi i Masson 1995, Sørensen i Yosha 1998). Okazało się *ex-post*, że brak takich mechanizmów rzeczywiście mógł stanowić istotną niedoskonałość unii walutowej w Europie, przyczyniając się do pogłębienia kryzysu zadłużenia (IMF 2013, Furceri i Zdzienicka 2013).

Literatura ekonomiczna zna kilka rodzajów mechanizmu podziału ryzyka między krajami. Można wśród nich wyróżnić prywatne mechanizmy podziału ryzyka (*private risk-sharing*) i publiczne mechanizmy podziału ryzyka (*public risk-sharing*). Do mechanizmów prywatnych należy transgraniczny rynek kapitałowy, umożliwiający dywersyfikację aktywów finansowych w celu stabilizowania wielkości majątku i wielkości dochodu z posiadanych aktywów. Zalicza się tu także krajowy i transgraniczny rynek kredytowy, umożliwiający zaciąganie pożyczek w celu wygładzania wahań konsumpcji w czasie. Z kolei do mechanizmów publicznych należy budżet państwa, pozwalający na przeprowadzanie zmian w zakresie poziomu wydatków w celu reagowania na wstrząsy asymetryczne, oraz system międzynarodowych podatków i transferów, umożliwiający wygładzanie wahań dochodów i konsumpcji wywołanych przez wstrząsy asymetryczne w poszczególnych państwach. Dzięki tym mechanizmom kraje mogą zabezpieczać posiadany poziom dochodu i konsumpcji w okresach spadku produkcji i popytu.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Skala wygładzania skutków wstrząsów asymetrycznych w ramach prywatnych i publicznych mechanizmów podziału ryzyka była przedmiotem wielu badań. W pionierskiej pracy na ten temat Asdrubali i in. (1996) analizowali skuteczność systemu podziału ryzyka w wygładzaniu produktu krajowego brutto w przeliczeniu na jednego mieszkańca Stanów Zjednoczonych. Wykorzystując metodę dekompozycji wariancji, autorzy wykazali, że w ramach unii walutowej, jaką stanowią USA, 62% wpływu wstrząsów na poziom PKB/mieszkańca jest wygładzane za pośrednictwem mechanizmów prywatnych oraz rynkowych w formie transakcji zakupu i sprzedaży aktywów na rynkach finansowych oraz w formie zaciągania pożyczek i kredytów. Pozostałe 13% udaje się zniwelować za pomocą federalnego systemu podatków i transferów międzystanowych, a 25% nie jest wygładzane w ogóle. Sørensen i Yosha (1998), stosując podobną metodologię, wykazali, że skala wygładzania wstrząsów za pośrednictwem mechanizmu fiskalnych transferów między krajami Wspólnot Europejskich w latach 1981–1990 była znacznie mniejsza niż w USA i wynosiła około 5%, z kolei podstawowym mechanizmem podziału ryzyka były pożyczki i kredyty krajowe, odpowiedzialne za niwelowanie około 46% wstrząsów. Bardzo podobne wyniki uzyskali ci sami autorzy w przypadku krajów OECD. Znacznie nowsze badania potwierdzają, że transgraniczne mechanizmy podziału ryzyka w Unii Europejskiej mają nadal marginalne znaczenie (Furceri i Zdzienicka 2013, Bąk i Maciejewski 2015). Metodologia tych badań opisana została krótko w ramce 8.13. Z kolei w tabeli 8.4 przedstawiono wyniki badań uzyskane przez różnych autorów w przypadku rozmaitych grup krajów i różnych okresów, odnoszące się do skali wygładzania wstrząsów asymetrycznych za pośrednictwem różnych kanałów podziału ryzyka.

Ramka 8.13. Szacowanie skuteczności kanałów podziału ryzyka metodą dekompozycji wariancji*

Mechanizm podziału ryzyka pozwala łagodzić wpływ zmian PKB na takie zmienne jak dochód narodowy i konsumpcja. Dzięki temu wstrząsy oddziałujące na poziom PKB nie wpływają lub wpływają tylko w ograniczonym stopniu na poziom dochodów i konsumpcji. Punktem wyjścia jest następująca tożsamość, obowiązująca w przypadku każdego okresu *t* (Sørensen i Yosha 1998):

$$GDP_i = \frac{GDP_i}{GNI_i}\frac{GNI_i}{NI_i}\frac{NI_i}{DNI_i}\frac{DNI_i}{\left(C_i + G_i\right)}\left(C_i + G_i\right), \tag{8.21}$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

gdzie: *GDP* – produkt krajowy brutto, *GNI* – dochód krajowy brutto, *NI* – dochód krajowy netto, *DNI* – dochód krajowy do dyspozycji, *C* – konsumpcja prywatna, *G* – konsumpcja publiczna, indeks *i* oznacza zaś kraj. Z definicji rachunków narodowych wynikają następujące tożsamości: *GNI* = *GDP* + saldo dochodów wypłaconych i otrzymanych z zagranicy, *NI* = *GNI* – amortyzacja środków trwałych, *DNI* = *NI* + transfery międzynarodowe (publiczne i prywatne) oraz (*C*+*G*) = *DNI* – oszczędności netto.

Podstawiamy następnie logarytmy oraz ich pierwsze różnice (Δ) w zakresie wszystkich zmiennych po obu stronach równania (8.21):

$$\Delta\ln(GDP_i)=\left[\Delta\ln(GDP_i)-\Delta\ln(GNI_i)\right]+\left[\Delta\ln(GNI_i)-\Delta\ln(NI_i)\right]+$$
$$+\left[\Delta\ln(NI_i)-\Delta\ln(DNI_i)\right]+\left[\Delta\ln(DNI_i)-\Delta\ln(C_i+G_i)\right]+\Delta\ln(C_i+G_i). \quad (8.22)$$

Równanie (8.22) pokazuje, że zmiany PKB nie przenoszą się z reguły na proporcjonalne zmiany w poziomie łącznej konsumpcji (*C*+*G*) – czynnikami amortyzującymi wpływ zmian PKB na zmiany konsumpcji są wyrażenia w nawiasach kwadratowych po prawej stronie równania. Pierwszy nawias kwadratowy ukazuje wpływ z tytułu dochodów zagranicznych, drugi nawias odzwierciedla wpływ amortyzacji środków trwałych, trzeci nawias przedstawia wpływ podatków i płatności transferowych, a czwarty nawias ukazuje wpływ w zakresie zmian oszczędności. Te cztery wyrażenia ujęte w nawiasach kwadratowych reprezentują więc cztery różne kanały wygładzania zmian PKB, czyli mechanizmy podziału ryzyka związanego ze zmianami PKB.

Pomnóżmy następnie obie strony (8.22) przez $\Delta\ln(GDP_i)$, przyjmijmy średnie zmiennych oraz wykorzystajmy wzór na wariancję i kowariancję. W efekcie uzyskujemy następującą dekompozycję wariancji zmiennej GDP:

$$var[\Delta\ln(GDP_i) = cov[\Delta\ln(GDP_i) - \Delta\ln(GNP_i), \Delta\ln(GDP_i)] +$$
$$+ cov[\Delta\ln(GNP_i - \Delta\ln(NI_i), \Delta\ln(GDP_i)] + cov[\Delta\ln(NI_i) - \Delta\ln(DNI_i),$$
$$\Delta\ln(GDP_i)] + cov[\Delta\ln(DNI_i) - \Delta\ln(C_i + G_i), \Delta\ln(GDP_i) +$$
$$+ cov[\Delta\ln(C_i + G_i), \Delta\ln(GDP_i) \quad (8.23)$$

przy czym:

$$var[X]=\frac{1}{N}\sum_{i}^{N}\left(X_i-\dot{X}\right)^2,$$

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

$$cov[X,Y]=\frac{1}{N}\sum_{i}^{N}\left(X_i-\dot{X}\right)\left(Y_i-\dot{Y}\right),$$

gdzie $var[X]$ jest przekrojową wariancją (*cross-sectional variance*), $\dot{X}$ oraz $\dot{Y}$ stanowią średnie arytmetyczne, a N oznacza liczbę krajów.

Dzieląc (8.20) przez $\Delta\ln\left(GDP_i\right)$, otrzymujemy:

$$1=\beta_f+\beta_d+\beta_T+\beta_s+\beta_u. \tag{8.24}$$

Zmienne β reprezentują udział poszczególnych kanałów wygładzania zmian PKB w całkowitym zasobie tych przekształceń i odnoszą się odpowiednio do dochodów zagranicznych (β_f), amortyzacji środków trwałych (β_d), podatków i płatności transferowych (β_T) oraz zmian oszczędności (β_s). Zmienna β_u reprezentuje niewygładzoną część zmian PKB.

Tabela 8.4. Skuteczność kanałów podziału ryzyka w wygładzaniu zmian dochodu i konsumpcji odnośnie różnych grup krajów i różnych okresów – wyniki rozmaitych badań (dane w %)

Kanał podziału ryzyka	Zmienna z równania (8.24)	EU8 (1981–1990) (Sørensen i Gosh 1998)	OECD (1981–1990) (Sørensen i Gosh 1998)	EU15 (1979–2010) (Furceri i Zdzienicka 2013)	OECD (1979–2010) (Furceri i Zdzienicka 2013)	EMU16 (1999–2014) (Bąk i Maciejewski 2015)
Dochody zagraniczne	β_f	2	–2	7,6	0,6	6
Amortyzacja środków trwałych	β_d	–8	–9	–8,4	–9,7	7
Międzynarodowe transfery	β_T	5	3	3,9	2,6	0
Oszczędności	β_s	24	44	31,0	32,9	20
W tym: publiczne prywatne				9,2 21,8	8,5 24,4	
Część niewygładzona	β_u	78	65	65,8	73,6	75

Źródło: opracowanie własne na podstawie Sørensen i Gosh (1998), Furceri i Zdzienicka (2013) oraz Bąk i Maciejewski (2015).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Jak widać, zarówno w Unii Europejskiej, jak i w grupie krajów OECD mechanizmy podziału ryzyka nie odgrywały dużej roli w wygładzaniu wstrząsów na poziomie PKB – niewygładzona część wstrząsów stanowiła między dwie trzecie a trzy czwarte całej zmienności PKB. Zwraca uwagę szczególnie niewielki udział transgranicznych mechanizmów podziału ryzyka w Unii Europejskiej – międzynarodowe transfery korygowały zaledwie między 0 a 5% wstrząsów i w miarę upływu czasu sytuacja ta nie ulegała większym zmianom. Podstawowym mechanizmem niwelowania skutków wstrząsu pozostawały oszczędności krajowe, w tym przede wszystkim prywatne.

8.9.5. „Trójkąt" współzależności

Kryzys finansowy pokazał, że bez daleko idących zmian unia walutowa w Europie nie będzie mogła sprawnie funkcjonować. Dalsze utrzymywanie *status quo* stawało się bardzo ryzykowne ze względu na duże prawdopodobieństwo „rozlewania się" negatywnych efektów zewnętrznych (*spillovers*) pod wpływem braku koordynacji polityki narodowej. Kraje strefy euro stanęły w ten sposób wobec strategicznego „trylematu":

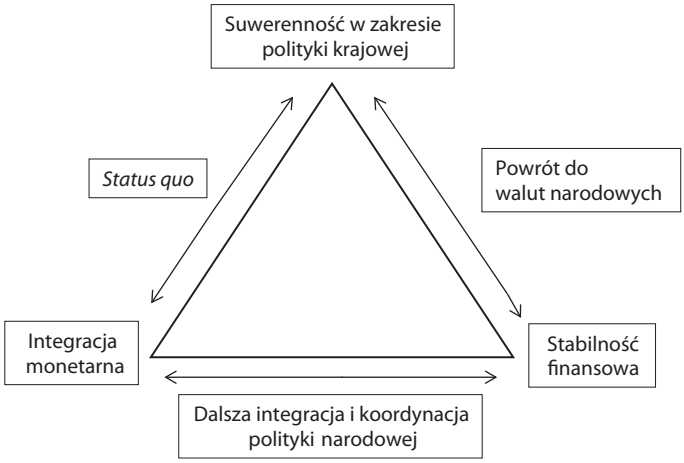
- albo pogłębią integrację i wzmocnią koordynację narodowej polityki, tak aby wyeliminować możliwość negatywnych efektów *spillovers* między krajami, co oznacza jednak dalsze ograniczanie tradycyjnie pojmowanej suwerenności;
- albo zrezygnują ze wspólnej waluty i powrócą do walut narodowych, co wiąże się jednak z utratą korzyści wynikających z posługiwania się wspólną walutą;
- albo utrzymają *status quo*, licząc się z ryzykiem powstania kolejnych kryzysów.

Ten strategiczny trylemat zilustrowany został na rysunku 8.9 w postaci trójkąta współzależności polityki. Wierzchołki trójkąta reprezentują ogólne cele, jakie chcą osiągnąć państwa członkowskie unii walutowej. Pierwszym z nich jest zachowanie suwerenności w zakresie polityki gospodarczej, czyli dostatecznie szerokiego zakresu swobody w podejmowaniu decyzji, w tym zwłaszcza w obszarze polityki fiskalno-budżetowej, finansowej i strukturalnej. Drugim z celów jest czerpanie korzyści z integracji gospodarczej i walutowej, w tym zwłaszcza z posługiwania się wspólną walutą. Trzeci wreszcie cel stanowi zachowanie stabilności finansowej w gospodarce. Jak łatwo zauważyć, równoczesne osiągnięcie wszystkich trzech celów jest niemożliwe. Posiadanie wspólnej waluty i zachowanie suwerenności w zakresie polityki krajowej pozwala wprawdzie odnosić korzyści z integracji walutowej, lecz nie zapewnia stabilności finansowej ze względu na wpływ indywidualnych, niezależnych decyzji

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

poszczególnych krajów na sytuację gospodarczą i finansową pozostałych państw członkowskich unii walutowej. Taki stan rzeczy ilustruje lewy bok trójkąta – jest to dotychczasowa sytuacja strefy euro (*status quo*). Istnieją dwie możliwości rozwiązania tego problemu. Pierwsza dotyczy utrzymania dotychczasowego zakresu suwerenności w zakresie polityki krajowej i rezygnacji ze wspólnej waluty na rzecz powrotu do walut narodowych – ten wariant opisuje prawy bok trójkąta. Druga możliwość to utrzymanie wspólnej waluty i ograniczenie suwerenności w zakresie polityki krajowej na rzecz większej koordynacji – ten wariant opisuje podstawa trójkąta.

Rysunek 8.9. Trójkąt współzależności w unii walutowej



Źródło: opracowanie własne.

Jaka była odpowiedź strefy euro na wyzwanie trylematu współzależności? Wydarzenia z lat 2009–2015 pokazały, że kraje strefy euro zdecydowały się dalej pogłębiać i poszerzać integrację, dokonując niezbędnych reform w wielu dziedzinach i zwiększając zakres koordynacji polityki krajowej. Nie ma wśród nich chętnych do rezygnacji ze wspólnej waluty. Mimo wielu krytycznych spostrzeżeń na temat dotychczasowego funkcjonowania unii walutowej, żadne państwo członkowskie nie zdecydowało się jeszcze z niej wystąpić. Wręcz przeciwnie, w czasie kryzysu i w latach następnych do strefy euro przystępowały kolejne kraje UE – Estonia (2011), Łotwa (2013) i Litwa (2015).

Nie oznacza to oczywiście, że opisana sytuacja nie może ulec zmianie w przyszłości. W większości krajów strefy euro obecne są partie polityczne, forsujące w swoich programach postulat wystąpienia ze strefy euro i powrotu do walut

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

narodowych. Brak postępów w niezbędnych reformach i niezdolność do powrotu na ścieżkę szybszego rozwoju będą umacniały nastroje populistyczne i eurosceptyczne, co może w pewnym momencie utorować wspomnianym partiom drogę do władzy. Może to ostatecznie doprowadzi do rozpadu strefy euro i poważnego kryzysu egzystencjalnego w Unii Europejskiej[46].

8.10. Nowy system zarządzania ekonomicznego w Unii Europejskiej

Jak wykazaliśmy, kryzys finansowy i gospodarczy z lat 2008–2012 ujawnił liczne niedoskonałości istniejącego systemu zarządzania ekonomicznego w Unii Europejskiej. Ponieważ państwa członkowskie zdały sobie sprawę z tego, że ich gospodarki są powiązane ze sobą silniej, niż wcześniej sądzono, a zarazem uznały, że chcą pozostać w unii walutowej, zdecydowano o dalszym umacnianiu koordynacji polityki narodowej wśród poszczególnych krajów i pogłębianiu integracji. Działania usprawniające system zarządzania ekonomicznego w UE po 2010 roku poszły w kierunku:

- wzmocnienia mechanizmu koordynacji i wspólnego nadzoru nad krajową polityką gospodarczą w ramach Semestru Europejskiego,
- reformy Paktu Stabilności i Wzrostu,
- utworzenia Unii Bankowej,
- stworzenia stałego mechanizmu stabilizacyjnego.

Reformy przeprowadzone w wymienionych czterech obszarach są szczegółowo omówione poniżej.

8.10.1. Semestr Europejski i Procedura Nierównowagi Makroekonomicznej

Semestr Europejski oznacza sześciomiesięczny okres każdego roku, kiedy kraje członkowskie UE przedstawiają i wspólnie omawiają swoją politykę fiskalno-budżetową, gospodarczą i strukturalną w celu zapewnienia koordynacji polityki, stabilizacji finansowej i gospodarczej oraz uniknięcia potencjalnych negatywnych

[46] Niektórzy autorzy twierdzą, że strefa euro ani się nie rozpadnie, ani w pełni nie zreformuje, lecz będzie nadal istnieć w niedoskonałym kształcie, przeżywając kolejne kryzysy (Cohen 2012).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

efektów *spillovers*. Roczny cykl koordynacji rozpoczyna się w końcu listopada każdego roku wraz z opublikowaniem przez Komisję Europejską Rocznego Przeglądu Wzrostu (*Annual Growth Survey* – AGS), w którym przedstawiona jest ocena aktualnej sytuacji gospodarczej w Unii oraz sformułowane są główne priorytety gospodarcze na następny rok. Równolegle Komisja publikuje także Raport Mechanizmu Alarmowego (*Alert Mechanism Report* – AMR), gdzie wymienione zostają główne czynniki zagrażające stabilności ekonomicznej w poszczególnych państwach członkowskich. Oba dokumenty są następnie dyskutowane w styczniu z udziałem przedstawicieli parlamentów narodowych i Parlamentu Europejskiego, a następnie omawiane i przyjmowane na posiedzeniu Rady Europejskiej w marcu. Na tej podstawie poszczególne państwa członkowskie sporządzają Narodowe Plany Reform (*National Reform Programs*), w których przedstawione są planowane reformy strukturalne i zmiany systemowe oraz krajowe Programy Stabilności (*Stability Programs* – dla państw strefy euro) i Programy Konwergencji (*Convergence Programs* – dla pozostałych państw UE), zawierające plany budżetowe i średniookresową strategię fiskalną poszczególnych państw. Następnie w maju każdego roku programy te przekazuje się do Komisji Europejskiej, która dokonuje oceny tych dokumentów z punktu widzenia ich zgodności z przyjętymi priorytetami (i z założeniami wieloletniego programu Europa 2020) oraz bada stan równowagi makroekonomicznej w poszczególnych państwach, wykorzystując w tym celu zestaw określonych wskaźników ekonomicznych. Na podstawie dokonanej analizy, Komisja formułuje zalecenia i rekomendacje pod adresem poszczególnych państw w postaci tzw. *Country Specific Recommendations* (CSR). Dyskusja nad CSR w Radzie UE oraz ich przyjęcie w czerwcu jako wiążących wytycznych dla poszczególnych państw kończy proces koordynacji polityki w ramach Semestru Europejskiego[47].

Podstawę prawną Semestru Europejskiego stanowi art. 121 TFUE (dotyczący koordynacji polityki gospodarczej), a także art. 146 TFUE (w sprawie polityki zatrudnienia) i art. 131 TFUE (odnośnie do specjalnych uregulowań wśród państw strefy euro). Szczegółowa procedura opisana jest w rozporządzeniu nr 1175/2011, stanowiącym część tzw. „sześciopaku" (*six-pack*), czyli pakietu sześciu aktów prawnych, który wszedł w życie w grudniu 2011 roku w ramach reformy system zarządzania ekonomicznego w UE. Rozporządzenie nr 1175/2011 zmienia i uzupełnia wcześniejsze rozporządzenie nr 1466/97, stanowiące część Paktu Stabilności i Wzrostu.

[47] Opisana procedura koordynacji w przypadku państw strefy euro obejmuje dodatkowe etapy, w tym opiniowanie projektów budżetów narodowych przez inne państwa członkowskie (w procesie „przeglądu rówieśniczego" – *peer review*).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Istotnym nowym elementem nadzoru i koordynacji polityki jest procedura makroekonomicznej nierównowagi (*Macroeconomic Imbalance Procedure* – MIP), wprowadzona w życie w grudniu 2011 roku na podstawie rozporządzenia nr 1174/2011 oraz 1176/2011, także w ramach „sześciopaku". Na podstawie tych przepisów makroekonomiczna równowaga w państwach członkowskich jest monitorowana za pomocą zestawu ustalonych wskaźników (*scoreboard*), takich jak saldo rachunku bieżącego, stopa bezrobocia, zadłużenie czy międzynarodowa pozycja inwestycyjna netto. Gdy wskaźniki przekraczają określone wartości progowe i powstaje ryzyko wystąpienia głębszej nierównowagi w danym państwie członkowskim, Komisja sporządza tzw. Pogłębiony Przegląd (*In-Depth Review*) w ramach Raportu Mechanizmu Alarmowego i może wnioskować do Rady UE o wszczęcie procedury nadmiernej nierównowagi (*Excessive Imbalance Procedure*) wobec danego państwa. Celem wdrożenia tej procedury jest podjęcie odpowiednich działań naprawczych (*corrective action*) przez państwo członkowskie, tak aby w określonym terminie przywrócić zachwianą równowagę makroekonomiczną. Wobec braku skutecznych działań, Rada UE może – na wniosek Komisji – nałożyć sankcje finansowe na państwo członkowskie (zob. ramka 8.14).

Ramka 8.14. Procedura Nierównowagi Makroekonomicznej

Narastanie nierównowagi makroekonomicznej w niektórych krajach stało się istotną przyczyną kryzysu w strefie euro w latach 2008–2010. W ramach „sześciopaku" państwa członkowskie UE przyjęły rozporządzenie nr 1174/2011 i 1176/2011, określające procedurę nadmiernej nierównowagi (*Macroeconomic Imbalance Procedure* – MIP) w zakresie wielostronnego nadzoru i kontroli nad polityką gospodarczą państw członkowskich[48]. Nierównowagę makroekonomiczną definiuje się tu jako wszelką tendencję, która zakłóca – lub może zakłócić – prawidłowe i zrównoważone funkcjonowanie gospodarek poszczególnych krajów członkowskich, UGW i całej UE. Na mocy tych rozporządzeń Komisja Europejska monitoruje stan równowagi makroekonomicznej w poszczególnych państwach, korzystając z tablicy wskaźników (*scoreboard*) oraz ich wartości progowych, a także przygotowuje roczne

[48] Regulation (EU) no. 1174/2011 of the European Parliament and the Council of 16 November 2011 on enforcement measures to correct excessive macroeconomic imbalances in the euro area oraz Regulation (EU) no. 1176/2011 of the European Parliament and the Council of 16 November 2011 on the prevention and correction of macroeconomic imbalances.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

raporty na ten temat (*Alert Mechanism Report*)[49]. Gdy Komisja stwierdzi istnienie nierównowagi, przeprowadza gruntowną analizę sytuacji w postaci tzw. Pogłębionego Przeglądu (*In-Depth Review*). Jeśli uda jej się w ten sposób potwierdzić występowanie nierównowagi, informuje o tym Radę UE, Parlament Europejski i Eurogrupę, a także Europejską Radę Ryzyka Systemowego. Gdy Komisja stwierdzi zaś istnienie *znacznej* nierównowagi (*severe imbalance*), może wnioskować do Rady o wszczęcie procedury nadmiernej nierównowagi (*Excessive Imbalance Procedure* – EIP) na podstawie art. 121(2) TFUE.

Na wniosek Komisji Rada UE wydaje rekomendację dla państwa członkowskiego z sugestiami dotyczącymi niezbędnych działań i terminem ich wykonania. Państwo członkowskie w odpowiedzi przedstawia w ciągu dwóch miesięcy plan działań naprawczych (*corrective action plan*), który podlega ocenie Komisji i Rady. Jeśli Rada uzna, że zgłoszony projekt jest niewystarczający, wzywa państwo członkowskie do przedstawienia uzupełnionego planu w terminie dwóch miesięcy.

Komisja monitoruje wdrażanie planu oraz informuje Radę o poczynionych postępach; kiedy realizacja programu naprawczego odbiega od ustaleń, Komisja informuje Radę o niewykonaniu przyjętych wcześniej założeń (*non-compliance*). Jeśli Rada nie odrzuci sprawozdania Komisji kwalifikowaną większością głosów (*qualified majority voting* – QMV), kieruje do państwa członkowskiego wezwanie do wykonania planu naprawczego. Na tym etapie Rada może nałożyć sankcje finansowe za niewykonywanie przyjętych ustaleń, zgodnie z rozporządzeniem nr 1174/2011.

Gdy Komisja stwierdza brak realizacji odpowiednich założeń (*non-compliance*), Rada ma prawo zażądać od państwa członkowskiego złożenia oprocentowanego depozytu. Z kolei kara finansowa może zostać nałożona decyzją Rady na wniosek Komisji, gdy:

a) w ramach tej samej procedury EIP Rada skierowała dwie kolejne rekomendacje, a mimo to państwo członkowskie przyjęło niewystarczający *plan* działań (*insufficient corrective action plan*);
b) w ramach tej samej procedury EIP Rada skierowała dwie kolejne rekomendacje, a mimo to państwo członkowskie podjęło niewystarczające *działania* (*insufficient corrective action*).

Decyzja o nałożeniu sankcji podejmowana jest w trybie odwróconej większości kwalifikowanej (*reverse qualified majority voting* – RQMV) w terminie 10 dni

[49] Analiza obejmuje 14 wskaźników makroekonomicznych i makrofinansowych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

na wniosek Komisji złożony nie później niż 20 dni od stwierdzenia *non-compliance*. Kara roczna nakładana jest w wysokości 0,1% PKB. Komisja może wnioskować o obniżenie lub anulowanie kary na podstawie wystąpienia „nadzwyczajnych okoliczności" (*exceptional circumstances*) lub za sprawą umotywowanego wniosku państwa członkowskiego (*a reasoned request by the Member State*). Jeśli nadmierna nierównowaga została wyeliminowana w ciągu roku, kara z odsetkami podlega zwrotowi (według zasady *pro rata temporis*).

8.10.2. Zreformowany Pakt Stabilności i Wzrostu

Zmiany w SGP, zatwierdzone na podstawie rozporządzenia nr 1175/2011 i 1177/2011[50], zmierzają zasadniczo w kierunku zwiększenia dyscypliny fiskalno-budżetowej poprzez rozbudowanie i udoskonalenie części prewencyjnej oraz skrócenia procedury nadmiernego deficytu w ramach części korekcyjnej Paktu.

Część prewencyjna Paktu odnosi się do tych państw członkowskich UE, które nie są objęte procedurą nadmiernego deficytu (EDP). Każdy z krajów posiada w tym przypadku wyznaczony średniookresowy cel (*Medium-Term Objective* – MTO) dostosowany do własnej sytuacji budżetowej. Cel MTO, wyrażony jako strukturalne saldo budżetu (czyli saldo skorygowane o wpływ cyklu koniunkturalnego), może przyjmować wartości: od –1% PKB do zera lub nadwyżki. Państwa członkowskie, które nie osiągnęły jeszcze swojego MTO, są zobowiązane do poprawy salda strukturalnego budżetu o 0,5% PKB w skali roku, przy czym faktyczna poprawa może być wyższa w fazie wzrostowej cyklu koniunkturalnego (w tzw. „dobrych czasach") i niższa w czasie spowolnienia lub recesji (w tzw. „złych czasach"). Gdy saldo strukturalne odchyla się od MTO – lub od wyznaczonej ścieżki dojścia do niego – o 0,5% PKB lub więcej, Komisja Europejska może skierować do państwa członkowskiego wstępne ostrzeżenie (*early warning*). Przejściowe odchylenia od MTO – lub od wyznaczonej ścieżki dojścia do niego – są dopuszczalne w trzech wyjątkowych sytuacjach:

[50] Regulation (EU) no. 1175/2011 of the European Parliament and of the Council of 16 November 2011 amending Council Regulation (EC) no. 1466/97 on the strengthening of the surveillance of budgetary positions and the surveillance and coordination of economic policies oraz Regulation of the Council (EU) no. 1177/2011 of 8 November 2011 amending Regulation (EC) no. 1467/97 on speeding up and clarifying the implementation of the excessive deficit procedure.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

– w przypadku podjęcia reform strukturalnych (np. reformy emerytalne lub reformy podatkowe) lub pewnych inwestycji publicznych[51], które mają dodatni wpływ na saldo budżetu w dłuższym okresie, lecz wywołują pogorszenie sytuacji budżetowej w krótkim okresie, jednakże pod warunkiem, że odchylenie to nie przekracza 0,5% PKB, saldo faktyczne budżetu utrzymuje się poniżej 3% PKB, a saldo strukturalne budżetu powraca do wyznaczonej wielkości w okresie nie dłuższym niż 4 lata;
– w przypadku wystąpienia nadzwyczajnego wydarzenia pozostającego poza kontrolą państwa członkowskiego i mającego znaczny wpływ na jego sytuację budżetową;
– w okresach poważnego ekonomicznego załamania (*severe economic downturn*) w strefie euro lub w całej UE o tyle, o ile dopuszczalne odchylenie nie zagraża jednak stabilności finansów publicznych w średnim okresie, a faktyczny deficyt nie przekracza 3% PKB.

Zalecenia odnoszące się do reguł części prewencyjnej Paktu wchodzą w skład corocznych CSR, czyli indywidualnie adresowanych do państw członkowskich rekomendacji (*country specific recommendations*), przyjmowanych w ramach Semestru Europejskiego. Ponadto, część prewencyjna Paktu została uzupełniona o regułę wydatkową (*expenditure benchmark*), nakładającą na te państwa członkowskie, które nie osiągnęły jeszcze swojego MTO, obowiązek utrzymania tempa wzrostu wydatków publicznych poniżej średniookresowego tempa wzrostu potencjalnego PKB[52]. Gdy saldo strukturalne budżetu odchyla się od MTO lub od ustalonej ścieżki dojścia do niego, Rada, działając na wniosek Komisji, może nałożyć na dane państwo członkowskie karę w postaci konieczności złożenia oprocentowanego depozytu w wysokości 0,2% PKB. Decyzja w tej sprawie podejmowana jest w trybie odwróconej większości kwalifikowanej (RQMV).

Część korekcyjna Paktu reguluje stosowanie procedury nadmiernego deficytu (EDP), która jest uruchamiana w następujących przypadkach:

a) gdy faktyczny (nominalny) deficyt budżetowy przekracza 3% PKB,
b) gdy państwo członkowskie, w którym dług publiczny przekracza 60% PKB, nie obniża relacji długu do PKB w dostatecznie szybkim tempie, czyli średnio co najmniej o jedną dwudziestą nadwyżki relacji długu do PKB ponad

[51] Inwestycje publiczne, które kwalifikują się do tej kategorii, to inwestycje współfinansowane z funduszy strukturalnych UE oraz z Europejskiego Funduszu Inwestycji Strategicznych (tzw. Plan Junckera).
[52] W Polsce funkcję tę pełni od 2013 roku stabilizująca reguła wydatkowa, opisana w rozdziale 3.10.5.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

wartość referencyjną 60% w okresie 3 lat, uwzględniając przy tym fazę cyklu koniunkturalnego.

Gdy Rada uzna (na wniosek Komisji), że deficyt przekracza 3% PKB, zleca państwu członkowskiemu skorygowanie nadmiernego deficytu w określonym czasie poprzez podjęcie „skutecznych działań" (*effective action*). W ramach tej rekomendacji Rada wyznacza roczne cele budżetowe w postaci tzw. wymaganego wysiłku fiskalnego (*fiscal effort*) w kategoriach strukturalnych lub nominalnych. Wartość tego wskaźnika musi być spójna z dążeniem do poprawy salda strukturalnego budżetu o co najmniej 0,5% PKB w ciągu roku, a także uwzględniać wiarygodne plany reform strukturalnych danego państwa członkowskiego (z wyłączeniem inwestycji publicznych). Uznaje się, że państwo członkowskie podjęło „skuteczne działania", jeśli uzyskana poprawa salda strukturalnego (*fiscal effort*) jest zgodna z zaleceniem Rady.

Rada może przedłużyć termin korygowania nadmiernego deficytu w dwóch sytuacjach:

a) gdy państwo członkowskie podjęło „skuteczne działania" (*effective action*) zmierzające do obniżenia deficytu, lecz cel nie został osiągnięty z powodu wystąpienia nieoczekiwanych wydarzeń, mających niekorzystny wpływ na finanse publiczne;
b) w okresach poważnego ekonomicznego załamania (*severe economic downturn*) w strefie euro lub w całej UE o tyle, o ile dopuszczalne odchylenie nie zagraża jednak stabilności finansów publicznych w średnim okresie.

Brak podporządkowania się tym zaleceniom może skutkować zaostrzeniem procedury EDP, włącznie z nałożeniem sankcji finansowych (wyłącznie w przypadku państw strefy euro). Gdy uruchamiana jest procedura EDP wobec państwa członkowskiego, które jest już objęte sankcjami w ramach części prewencyjnej Paktu, lub gdy naruszenie wartości referencyjnej staje się szczególnie duże, państwo członkowskie zobowiązane jest złożyć nieoprocentowany depozyt w wysokości 0,2% PKB. Brak podjęcia „skutecznych działań" w zakresie korygowania nadmiernego deficytu wiąże się z nałożeniem kary 0,2% PKB, z kolei gdy sytuacja taka powtarza się, następuje zwiększenie kary o zmienny składnik (dodatkowo 0,1% PKB za każdy punkt procentowy wynikający z przekroczenia wartości referencyjnej). Decyzje o nałożeniu sankcji podejmowane są w trybie odwróconej większości kwalifikowanej (*reverse* QMV), z wyjątkiem ostatniej sankcji, która wymaga zwykłej QMV.

Trzeba też podkreślić, że uporczywe łamanie przez państwo członkowskie reguł fiskalnych ujętych w zreformowanym SGP może spowodować także wstrzymanie wypłacania temu państwu funduszy strukturalnych i spójnościowych z budżetu

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

UE. W ramach przyjętego w końcu 2013 roku pakietu legislacyjnego regulującego europejską politykę spójności w latach 2014–2020 państwa członkowskie zgodziły się uwzględnić tzw. warunkowość makroekonomiczną (*macro-economic conditionality*). Ma ona polegać na tym, że wypłata funduszy unijnych w przypadku danego państwa członkowskiego może dokonać się tylko pod warunkiem prowadzenia przez to państwo odpowiedzialnej polityki makroekonomicznej (*sound macro-economic policies*), tak aby środki unijne mogły być wykorzystane w efektywny sposób. Dotyczy to zarówno krajów członkowskich strefy euro, jak i pozostałych państw członkowskich UE[53].

Istotnym elementem „sześciopaku" w zakresie dyscypliny fiskalnej jest dyrektywa Rady 2011/85/EU z listopada 2011 roku, nakładająca na państwa członkowskie obowiązek inkorporowania do ustawodawstwa krajowego reguł fiskalnych przyjętych na szczeblu unijnym. Zgodnie z Dyrektywą, każde państwo członkowskie wprowadza do prawa krajowego numeryczne reguły fiskalne (*numerical fiscal rules*), które zapewnią skuteczne wypełnienie zobowiązań TFUE w zakresie polityki fiskalnej sektora rządowego w horyzoncie wieloletnim, w tym w szczególności realizację wartości referencyjnych deficytu i długu publicznego zgodnie z wymogami TFUE oraz przyjęcie wieloletnich ram planowania fiskalnego zapewniających realizację MTO. Przepisy dotyczące reguł fiskalnych określają, między innymi, definicje reguł i zakres ich działania, zasady monitorowania przestrzegania reguł przez niezależne instytucje (*independent fiscal councils*) oraz konsekwencje braku przestrzegania reguł (*non-compliance*).

Uznając, że szczególna współzależność wynikająca z istnienia wspólnej waluty w strefie euro wymaga ściślejszej koordynacji polityki fiskalnej niż rozwiązania przyjęte w ramach „sześciopaku", oraz uwzględniając większy potencjał do generowania efektów *spillovers* na obszarze unii walutowej niż poza nią, państwa strefy euro zdecydowały o przyjęciu dodatkowych rozwiązań zapewniających zwiększoną

[53] "Under the preventive arm of the macro-economic conditionality the Commission may request amendments to a member state's partnership agreement in order to address for instance macroeconomic imbalances. Where a member state fails to respond satisfactorily to such a request, the Commission may propose to the Council to suspend a part or all of the payments for the programmes concerned. The proposed suggestion requires the support of a qualified majority of the Council to take effect. Under the corrective arm, the Commission must propose to suspend part or all commitments or payments when a member state fails to take corrective action within the economic governance procedures (e.g. the excessive deficit procedure and the macroeconomic imbalances procedure). A proposal to suspend commitments is deemed adopted unless it is rejected by the Council with qualified majority. A proposal to suspend payments requires the support of a qualified majority of the Council". Zob. *Council adopts cohesion policy package for 2014–2020*, Council of the European Union, Brussels, 17826/13, Presse 579, 16 December, 2013.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

dyscyplinę fiskalną i stabilność finansową. Regulacje te zostały ujęte w tak zwanym „dwupaku", czyli w pakiecie dwóch rozporządzeń, przyjętych w maju 2013 roku. Podstawę prawną w przypadku obu regulacji stanowi art. 136 TFUE, który pozwala na wzmocnioną współpracę w strefie euro. Pierwsze z rozporządzeń wprowadza jednolity kalendarz prac budżetowych w państwach strefy euro i przyznaje Komisji Europejskiej prawo do opiniowania planów budżetowych poszczególnych państw przed ich zatwierdzeniem w parlamentach narodowych[54]. W szczególności, w sytuacji zagrożenia stabilności finansów publicznych, Komisja może zażądać rewizji planów budżetowych. Drugie z rozporządzeń odnosi się do państw objętych procedurą EDP i nakłada na nie obowiązek przygotowywania strategii „korekcyjnych" w postaci programów ekonomicznego partnerstwa oraz daje Komisji prawo do dokładniejszego monitorowania sytuacji budżetowej w tych państwach[55].

8.10.3. Traktat o Stabilizacji, Koordynacji i Zarządzaniu

Traktat o Stabilizacji, Koordynacji i Zarządzaniu przyjęty w ramach Unii Gospodarczo-Walutowej (*Treaty on Stability, Coordination and Governance in EMU* – TSKZ) miał na celu dalsze zwiększanie dyscypliny fiskalnej, a także ściślejszą koordynację i poprawę zarządzania ekonomicznego w UE. TSKZ został podpisany 2 marca 2012 roku przez 25 państw członkowskich UE (Wielkiej Brytania i Czechy nie dołączyły do Traktatu, a Chorwacja nie była jeszcze członkiem UE) i ratyfikowany przez wszystkich sygnatariuszy do końca 2014 roku. Zapisy Tytułu III Traktatu, nazywanego Paktem Fiskalnym (*fiscal compact*), stanowią powtórzenie lub dalsze rozwinięcie reguł fiskalnych zapisanych w zreformowanym SGP. Główne postanowienia Paktu Fiskalnego są następujące:

– w art. 3 ust. 1 TSKZ potwierdzona została średniookresowa reguła równowagi budżetowej: budżet (sektora *general government*) uznany jest za zrównoważony lub z nadwyżką, gdy roczny deficyt strukturalny nie przekracza 0,5% PKB. Reguła ta powinna być wprowadzona do krajowego porządku prawnego na poziomie konstytucyjnym (lub równorzędnym) wraz z automatycznym

[54] Regulation (EU) no. 473/2013 of 21 May 2013 on common provisions for monitoring and assessing draft budgetary plans and ensuring the correction of excessive deficit of the Member States in the euro area.

[55] Regulation (EU) no. 472/2013 of 21 May 2013 on the strengthening of economic and budgetary surveillance of Member States experiencing or threatened with serious difficulties with respect to their financial stability in the euro area.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

mechanizmem korekcyjnym, zgodnie z kalendarzem zaproponowanym przez Komisję Europejską (art. 3 ust. 2);

- Europejski Trybunał Sprawiedliwości otrzymuje prawo do weryfikowania wypełnienia tych postanowień przez państwa członkowskie i do nakładania kar finansowych (do wysokości 0,1% PKB) na państwa, które nie wypełniają swoich zobowiązań w tym zakresie;
- także w art. 3 Traktat wprowadza zmienny górny limit strukturalnego deficytu budżetowego: niższy w przypadku krajów o wyższym poziomie zadłużenia (0,5% PKB dla krajów o długu > 60% PKB) i wyższy w przypadku krajów o niższym poziomie zadłużenia (1,0% PKB dla krajów o długu < 60% PKB);
- Traktat potwierdza wprowadzoną w ramach „sześciopaku" regułę „hamulca długu" (*debt brake*), zgodnie z którą kraje o zadłużeniu przekraczającym 60% PKB mają obowiązek obniżania co rok tej relacji o jedną dwudziestą nadwyżki długu w stosunku do PKB powyżej 60% (art. 4); ze względu na trwający wówczas kryzys gospodarczy przepis ten wszedł w życie trzy lata później, tj. w 2015 roku[56];
- państwa członkowskie objęte procedurą EDP przedkładają Komisji do zatwierdzenia (*endorsement*) programy ekonomicznego partnerstwa (*economic partnership programmes*) zawierające szczegółowe reformy strukturalne, które mają zapewnić skutecznie trwałą eliminację nadmiernego deficytu, realizacja tych reform jest zaś monitorowana przez Komisję i Radę (art. 5); obowiązek ten został wpisany już do prawa unijnego w ramach tzw. „dwupaku" (zob. wyżej).

W sytuacji, gdy państwo członkowskie nie przestrzega reguły średniookresowej równowagi budżetowej lub zasady „hamulca długu", uruchamia się zapisany w prawie krajowym automatyczny mechanizm korekcyjny, który ma niezwłocznie przywrócić deficyt strukturalny do realizacji celu średniookresowego lub przyczynić się do powrotu na wyznaczoną ścieżkę redukcji długu. Mechanizm ten może mieć postać automatycznej reguły wydatkowej (tak jak w Polsce) lub jakąkolwiek inną postać spełniającą ogólne wymogi sformułowane przez Komisję. W ramach wdrażania tego procesu państwa członkowskie powinny powołać niezawisłe Rady Polityki Fiskalnej (zob. rozdział 3, punkt 3.10.7) jako ciała odpowiedzialne za niezależną i obiektywną ocenę przestrzegania ustalonych reguł fiskalnych w zakresie krajowej polityki fiskalno-budżetowej.

[56] Omówienie różnic między paktem fiskalnym (TSKZ) a zreformowanym SGP można znaleźć w publikacji EBC, *A fiscal compact for a stronger economic and monetary union*, "ECB Monthly Bulletin", May 2012, s. 79–94.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Trzeba zaznaczyć, że w odniesieniu do państw-sygnatariuszy TSKZ nie będących członkami strefy euro wymogi paktu fiskalnego obowiązują dopiero od czasu uchylenia wobec nich derogacji – czyli od momentu przystąpienia do strefy euro – chyba że państwa te zadeklarują gotowość poddania się regułom fiskalnym TSKZ wcześniej (art. 14).

Tytuł IV TSKZ poświęcony jest koordynacji polityki gospodarczej. Zapisy tego ustępu powtarzają i rozwijają wcześniejsze postanowienia tzw. Paktu „Euro Plus", podpisanego w marcu 2011 roku przez państwa strefy euro i kilka pozostałych państw członkowskich UE[57]. W tej części państwa członkowskie zobowiązują się podejmować niezbędne działania w celu zapewnienia niezakłóconego funkcjonowania strefy euro i zwiększania konkurencyjności, a także wspierania zatrudnienia i wzmacniania stabilności finansów publicznych. Programy reform strukturalnych stają się więc przedmiotem wspólnej dyskusji i koordynacji między państwami członkowskimi strefy euro.

Tytuł V TSKZ odnosi się do zarządzania ekonomicznego w strefie euro i przewiduje zwoływanie formalnych spotkań szefów państw i rządów krajów strefy euro (*Euro Summits*), mających odbywać się regularnie, co najmniej dwa razy do roku. W spotkaniach tych mogą uczestniczyć także szefowie państw i rządów pozostałych krajów członkowskich UE, które dołączyły do Traktatu, choć tylko w sprawach związanych z „konkurencyjnością, globalną architekturą strefy euro i jej przyszłymi zasadami funkcjonowania" oraz bez prawa głosu.

Do czasu wpisania postanowień TSKZ do prawa unijnego Traktat funkcjonuje jako porozumienie o charakterze międzyrządowym, pozostające formalnie poza porządkiem prawnym Unii Europejskiej.

8.10.4. Unia Bankowa

Jedną z najważniejszych lekcji wyniesionych z kryzysu finansowego z lat 2008–2010 było uświadomienie sobie konieczności przeprowadzenia w UE harmonizacji przepisów regulujących funkcjonowanie sektora finansowego, w tym zwłaszcza

[57] Pakt Euro Plus został przyjęty jako zobowiązanie polityczne szefów państw i rządów strefy euro na posiedzeniu 24–25 marca 2011 roku. Do Paktu dołączyły również Bułgaria, Dania, Łotwa, Litwa, Polska i Rumunia na zasadzie dobrowolności. Porozumienie to dotyczy kwestii znajdujących się w kompetencji państw członkowskich, mających kluczowe znaczenie w zakresie poprawy konkurencyjności i unikania nierównowagi. Celem przyjęcia Paktu było wzmocnienie ekonomicznego wymiaru unii walutowej, poprawa koordynacji polityki gospodarczej i zwiększenie konkurencyjności. Postanowienia Paktu zostały formalnie wprowadzone w życie w ramach „sześciopaku" i paktu fiskalnego.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

banków. Istnienie kilkunastu odrębnych, zróżnicowanych narodowych systemów nadzoru nad bankami w warunkach funkcjonowania wspólnego pieniądza i wspólnej, jednolitej polityki pieniężnej w ramach unii walutowej okazało się kosztownym anachronizmem. Banki stosowały arbitraż regulacyjny, rozwijając bardziej ryzykowne rodzaje działalności w krajach o mniej wymagającym systemie nadzoru, co w warunkach istnienia międzynarodowych powiązań finansowych i właścicielskich między bankami powodowało rozprzestrzenianie się ryzyka systemowego w całej Unii. Prowadziło to także do segmentacji rynku usług bankowych i finansowych. Dążąc do wyeliminowania tych niedoskonałości, państwa strefy euro zdecydowały się na utworzenie unii bankowej. W jej skład weszły głównie następujące rozporządzenia:

- jednolite regulacje ostrożnościowe we wszystkich państwach członkowskich (tzw. *single rulebook*); regulacje te opierają się na rekomendacjach Komitetu Bazylejskiego (Basel II i Basel III – zob. rozdział 7). Jednolity zestaw reguł ma zapobiegać arbitrażowi regulacyjnemu między państwami członkowskimi i zapewniać jednakowe standardy mikroostrożnościowe, w tym wzmocnione wymogi kapitałowe dla banków;
- jednolity mechanizm nadzoru nad bankami (*single supervisory mechanism* – SSM); system ten wszedł w życie w listopadzie 2014 roku, a odpowiedzialność za jego funkcjonowanie spoczywa na Europejskim Banku Centralnym, który w wypełnianiu swoich funkcji współpracuje z narodowymi instytucjami nadzoru bankowego. Mechanizm SSM ma zapewniać ujednolicone zasady oceny sytuacji banków;
- jednolity mechanizm upadłości, likwidacji i restrukturyzacji banków (*single resolution mechanism* – SRM), który wszedł w życie w 2015 roku. Odpowiedzialność za jego funkcjonowanie spoczywa na nowo powołanej instytucji (*Single Resolution Board* – SRB), współpracującej z krajowymi instytucjami nadzorczymi. Mechanizm ten ma sprawić, że banki, które przechodzą trudności z płynnością lub niewypłacalnością, nie będą „zarażać" swoimi kłopotami innych banków, a w przypadku stwierdzenia ich upadłości koszty poniosą w pierwszej kolejności właściciele banków oraz inwestorzy, a nie podatnicy czy depozytariusze;
- jednolity mechanizm ubezpieczenia depozytów (*single deposit insurance mechanism* – SDIM), który ma wejść w życie do 2020 roku[58].

[58] Więcej na temat tworzącej się unii bankowej w Europie zob. *Europejska Unia Bankowa*, red. M. Zaleska, Difin, Warszawa 2015.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Podobne działania harmonizujące i koordynujące podejmowane są na szczeblu unijnym także w odniesieniu do pozostałych sektorów systemu finansowego, w których pojawia się ryzyko systemowe. Zabiegi te dotyczą między innymi sektora ubezpieczeniowego, giełd i pozostałych rynków papierów wartościowych oraz funduszy emerytalnych. Od 2011 roku na Europejski System Nadzoru Finansowego (*European Supervisory Authorities* – ESAs) składają się trzy instytucje odpowiedzialne za nadzór mikroostrożnościowy na szczeblu UE:

- Europejski Urząd Nadzoru Bankowego (*European Banking Authority* – EBA) z siedzibą w Londynie,
- Europejski Urząd Nadzoru nad papierami wartościowymi i rynkami kapitałowymi (*European Securities and Markets Authority* – ESMA) z siedzibą w Paryżu,
- Europejski Urząd Nadzoru nad Ubezpieczeniami i Pracowniczymi Systemami Emerytalnymi (*European Insurance and Occupational Pensions Authority* – EIOPA) z siedzibą we Frankfurcie nad Menem.

8.10.5. Europejski Mechanizm Stabilizacyjny

Jak wcześniej wspomniano, strefa euro nie dysponowała mechanizmem antykryzysowym, który umożliwiałby szybkie udzielenie pomocy finansowej krajom znajdującym się w trudnej sytuacji ekonomicznej. Reagując na kryzys zadłużenia w Grecji, państwa strefy euro utworzyły w maju 2010 roku dwa doraźne instrumenty wsparcia finansowego dla państw członkowskich dotkniętych kryzysem finansowym. W ramach tych rozwiązań wprowadzono Europejski Instrument Stabilności Finansowej (*European Financial Stability Facility* – EFSF) i Europejski Mechanizm Stabilności Finansowej (*European Financial Stability Mechanism* – EFSM).

Europejski Instrument Stabilności Finansowej został powołany przez państwa strefy euro 9 maja 2010 roku. Jego celem było zapewnienie stabilności ekonomicznej w UE poprzez udzielanie pomocy finansowej tym państwom strefy euro, które przechodziły trudności z płynnością lub wypłacalnością[59]. EFSF mógł emitować obligacje oraz inne rynkowe instrumenty dłużne w celu gromadzenia środków niezbędnych do udzielania pożyczek państwom strefy euro, dokapitalizowywania ich banków lub skupowania obligacji rządów państw strefy euro. Obligacje EFSF

[59] EFSF został utworzony jako „podmiot specjalnego przeznaczenia" (*special purpose vehicle* – SPV) z siedzibą w Luksemburgu. SPV jest instytucją finansową mającą na celu pozyskiwanie funduszy na rynku kapitałowym. Niezbędne wsparcie organizacyjno-techniczne zapewnia Europejski Bank Inwestycyjny na podstawie umowy.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

były gwarantowane przez państwa strefy euro w proporcji równej ich udziałom w kapitale własnym EBC do kwoty 440 mld EUR. EFSF zaczął formalnie funkcjonować w sierpniu 2010 roku i zakończył działalność w grudniu 2013 roku, gdy jego zadania i środki przejął Europejski Mechanizm Stabilizacyjny (ESM).

Uzupełnieniem finansowania z EFSF były pożyczki z Europejskiego Mechanizmu Stabilności Finansowej udzielane do kwoty 60 mld EUR. Finansowała je Komisja Europejska, wykorzystując w tym celu emisję obligacji na rynkach kapitałowych pod zastaw dochodów budżetu Unii Europejskiej. EFSF i EFSM mogły liczyć na dodatkowe wsparcie w postaci pożyczek z Międzynarodowego Funduszu Walutowego w sumie nieprzekraczającej kwoty 250 mld EUR, co pozwoliło stworzyć w pierwszych miesiącach kryzysu sieć bezpieczeństwa finansowego o łącznej wartości 750 mld EUR.

Europejski Mechanizm Stabilizacyjny (ESM) zastąpił EFSF i EFSM w 2013 roku. Funkcjonuje jako organizacja międzyrządowa, utworzona na podstawie porozumienia zawartego przez państwa członkowskie strefy euro. Kapitał subskrybowany ESM został ustalony na poziomie 700 mld EUR, z czego 80 mld stanowią wpłaty gotówkowe wniesione przez państwa strefy euro w okresie pięciu lat (od kwietnia 2014 roku ESM jest już w pełni dokapitalizowany) plus gwarancje rządowe na pozostałą kwotę 620 mld EUR. Tabela 8.5 przedstawia udział poszczególnych państw członkowskich strefy euro w kapitale ESM. Przy limicie pożyczkowym nie przekraczającym 500 mld EUR i ustalonym współczynniku wypłacalności CAR na poziomie powyżej 15% ESM otrzymał od agencji ratingowych najwyższy rating inwestycyjny (potrójne A). Pożyczki z ESM cieszą się przywilejem pierwszeństwa w spłacie (mają status preferowany), ustępując tylko pożyczkom IMF.

Zasady działania ESM przewidują, że wsparcie będzie udzielane tylko na podstawie rygorystycznego programu dostosowawczego (przygotowanego przez Komisję i MFW we współpracy z EBC). Pomoc przyznawana jest w zasadzie jednomyślnie przez członków Eurogrupy, przy czym w sytuacji kryzysowej, gdy Komisja i EBC stwierdzą, że obecny stan wymaga podjęcia pilnych działań, do zatwierdzenia decyzji wystarczy kwalifikowana większość 85% głosów. ESM został formalnie powołany do życia 27 września 2012 roku w ramach ratyfikacji dokonanej przez kraje reprezentujące 90% kapitału jako organizacja międzyrządowa zarejestrowana w Luksemburgu.

ESM gromadzi środki poprzez bezpośrednie zaciąganie pożyczek lub emisję obligacji gwarantowanych przez państwa członkowskie. Obligacje te zawierają tzw. klauzulę wspólnego działania (*common action clauses* – CAC), umożliwiającą zbiorowe

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

negocjowanie i zmianę warunków zadłużenia kwalifikowaną większością głosów wierzycieli. Udzielenie pomocy finansowej odbywa się na podstawie skierowania wniosku państwa członkowskiego do Eurogrupy, która z kolei zwraca się do Komisji o ocenę proponowanego programu dostosowawczego i rekomendacje; Komisja monitoruje także wykonanie programu pomocowego. Koszt pożyczki obejmuje wydatki związane z pozyskaniem środków plus standardową marżę w wysokości 200 bps (300 bps w przypadku pożyczek udzielanych na okres powyżej trzech lat).

Tabela 8.5. Udziały w ESM – stan na koniec 2015 roku

Kraj	Udział (dane w %)	Wpłacony kapitał (*paid-in capital*, dane w mln EUR)	Łączny kapitał subskrybowany (*total subscribed capital*, dane w mln EUR)	PKB (stan na 2011 rok, dane w mln EUR)
Germany	27,1464	21 717,1	190 024,8	3 315 643
France	20,3859	16 308,7	142 701,3	2 582 527
Italy	17,9137	14 331,0	125 395,9	2 055 114
Spain	11,9037	9 523,0	83 325,9	1 409 946
Netherlands	5,7170	4 573,6	40 019,0	783 293
Belgium	3,4771	2 781,7	24 339,7	465 676
Greece	2,8167	2 253,4	19 716,9	305 415
Austria	2,7834	2 226,7	19 483,8	376 841
Portugal	2,5092	2 007,4	17 564,4	229 336
Finland	1,7974	1 437,9	12 581,8	239 232
Ireland	1,5922	1 273,8	11 145,4	204 261
Slovakia	0,8240	659,2	5 768,0	86 262
Slovenia	0,4276	342,1	2 993,2	46 442
Luxemburg	0,2504	200,3	1 752,8	52 433
Cyprus	0,1962	157,0	1 373,4	22 752
Estonia	0,1860	148,8	1 302,0	19 220
Malta	0,0731	58,5	511,7	7 801
Total	100,00	80 000,0	700 000,0	12 202 194

Źródło: opracowanie własne na podstawie *ESM Treaty* (*Annex 1*) oraz innych dokumentów ESM.

Tabela 8.6 przedstawia porównanie struktury instytucjonalnej i kapitałowej, podstaw prawnych oraz sposobu funkcjonowania różnych unijnych instrumentów pomocowych mających na celu udzielanie wsparcia finansowego państwom członkowskim znajdującym się w trudnej sytuacji ekonomicznej. Jak można zauważyć, kraje strefy euro korzystają obecnie ze znacznie bardziej rozbudowanej

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

i zasobniejszej infrastruktury wsparcia finansowego niż pozostałe państwa członkowskie UE. Te ostatnie mogą bowiem liczyć tylko na pomoc finansową udzielaną w ramach art. 122 (sytuacje nadzwyczajne i klęski żywiołowe) oraz art. 143 (wsparcie w zakresie bilansu płatniczego) na łączną kwotę 110 mld EUR. Kraje strefy euro mają zaś obecnie także dostęp do ESM o zdolności pożyczkowej 500 mld EUR (wcześniej EFSF), a ponadto mogą korzystać z pomocy EBC.

Tabela 8.6. Charakterystyka unijnych instrumentów pomocy finansowej dla państw członkowskich

Wyszczególnienie	Wsparcie w zakresie bilansu płatniczego	EFSM	EFSF	ESM
Forma prawna	mechanizm traktatowy UE	mechanizm traktatowy UE	spółka prawa handlowego państw członkowskich strefy euro	organizacja międzyrządowa państw członkowskich strefy euro
Struktura kapitałowa	gwarantowana przez budżet UE	gwarantowana przez budżet UE	wkłady państw strefy euro plus gwarancje państw strefy euro	wkłady państw strefy euro plus gwarancje państw strefy euro
Zdolność pożyczkowa	50 mld EUR	60 mld EUR	440 mld EUR	500 mld EUR
Instrumenty	pożyczki, linie kredytowe	pożyczki, linie kredytowe	pożyczki, zakup obligacji na rynku pierwotnym i wtórnym	pożyczki, zakup obligacji na rynku pierwotnym i wtórnym, linie kredytowe, dokapitalizowanie banków
Rok utworzenia	1992	2010	2010	2013
Czas trwania	mechanizm trwały	mechanizm tymczasowy	mechanizm tymczasowy	mechanizm trwały
Organ decyzyjny	Rada UE, podejmująca decyzje QMV na wniosek Komisji	Rada UE, podejmująca decyzje QMV na wniosek Komisji	Rada Dyrektorów EFSF składająca się z ministrów finansów państw strefy euro (Eurogrupa)	Rada Dyrektorów ESM składająca się z ministrów finansów państw strefy euro (Eurogrupa)
Podstawa prawna finansowania	art. 143 TFUE	art. 122 TFUE	umowa międzyrządowa	traktat międzyrządowy i art. 132 TFUE
Kraje objęte pomocą (do 2016 roku)	Łotwa, Węgry, Rumunia	Grecja, Irlandia, Portugalia	Grecja (do 2013 roku), Irlandia, Portugalia	Cypr, Grecja (od 2013 roku)

Źródło: opracowanie własne na podstawie *Financial assistance provided to countries in difficulties*, Special Report, no. 18/2015, European Court of Auditors, Luksemburg.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

8.11. Przyszłość strefy euro

8.11.1. Czy strefa euro jest optymalnym obszarem walutowym?

W czasie kryzysu powróciły wątpliwości, czy strefa euro rzeczywiście stanowi optymalny obszar walutowy w rozumieniu teorii Roberta Mundella. Wielu autorów, zwłaszcza anglosaskich, od początku twierdziło, że Unia Europejska nie spełnia w tym przypadku odpowiednich kryteriów i że utworzenie unii walutowej w Europie było kosztownym błędem (przeglądu literatury amerykańskiej na ten temat dokonują m.in. Jonung i Drea 2009). Kryzys zadłużenia suwerennego ma zaś potwierdzać te sceptyczne przewidywania.

Niestety, nie dysponujemy żadnym precyzyjnym miernikiem, który pozwoliłby na określenie, czy dany region gospodarczy stanowi optymalny obszar walutowy. W tej sytuacji można zastosować analizę porównawczą. Wówczas jednym z możliwych sposobów uzyskania przybliżonej odpowiedzi na pytanie o to, czy strefa euro jest optymalnym obszarem walutowym, staje się porównanie struktury gospodarczej oraz instytucjonalnej w strefie euro z gospodarką Stanów Zjednoczonych, która w powszechnej opinii uchodzi za modelowy przykład optymalnego obszaru walutowego (Bayoumi i Eichengreen 1992). Tabela 8.7 przedstawia zasadniczą charakterystykę obu tych sfer gospodarczych.

Tabela 8.7 potwierdza, że pod wieloma względami UGW jest wewnętrznie słabiej zintegrowana niż Stany Zjednoczone. Jej główne niedoskonałości instytucjonalne wiążą się z jednej strony z brakiem unii fiskalnej zapewniającej przeprowadzanie transferów wygładzających skutki szoków asymetrycznych między uczestnikami unii walutowej oraz z drugiej strony z brakiem formalnych kompetencji umożliwiających EBC pełnienie funkcji pożyczkodawcy ostatniej szansy dla rządów. Ograniczona unia fiskalna może jednak powstać w przyszłości w Europie w wyniku utworzenia wspólnego budżetu strefy euro, o czym mówi się już od pewnego czasu (zob. następny punkt). Z kolei ten drugi mankament może przestać odgrywać istotną rolę na skutek zadeklarowanej gotowości EBC do zaangażowania się w operacje OMT, oznaczające *de facto* wsparcie adresowane do rządów państw znajdujących się w kryzysie zadłużenia (zob. rozdział 5, punkt 5.14.5).

Podobnie jest z różnicami na poziomie struktury gospodarczej. UGW niewątpliwie cechuje się mniejszą wewnętrzną mobilnością pracowników (zob. MFW 2012) i zapewne także mniejszą elastycznością wynagrodzeń niż Stany Zjednoczone

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

(zob. Duarte i Robalo Marques 2013). Z kolei pod względem stopnia integracji poprzez handel oraz poziomu synchronizacji cyklu koniunkturalnego oba obszary są do siebie zbliżone. Podsumowując, o ile można niewątpliwie stwierdzić, że USA w większym stopniu niż strefa euro spełnia kryteria optymalnego obszaru walutowego, o tyle nie daje się jednoznacznie przesądzić o tym, że strefa euro nie spełnia nawet minimalnych kryteriów, zwłaszcza jeśli weźmie się pod uwagę planowane reformy.

Tabela 8.7. Porównanie unii walutowej w UE i w USA: instytucje oraz struktura gospodarcza

Charakterystyka	USA	UGW	Obecne reformy w UGW
Instytucje			
Wspólna waluta	TAK	TAK	
Wspólny bank centralny	TAK	TAK	
Pożyczkodawca ostatniej szansy dla rządu	TAK	NIE	NIE
Możliwe wyjście z unii walutowej	NIE	NIE	NIE
Unia polityczna	TAK	NIE	NIE
Unia fiskalna (dochody własne, transfery wewnętrzne)	TAK (28%[a])	NIE (0,5%[a])	NIE***[b])
Unia bankowa	TAK	NIE	TAK
Struktura gospodarcza			
Integracja gospodarcza (handel wewnętrzny jako % PKB)	66%[c]	17%	???
Synchronizacja cyklu koniunkturalnego (korelacja części z całością)[d]	0,78/0,46[e]	0,51	
Mobilność siły roboczej	TAK	częściowa	???
Elastyczność płac	częściowa	NIE	???

a) Udział dochodów lokalnych kompensowany przez transfer centralny.
b) Reforma SGP i pakt fiskalny mają być substytutem unii fiskalnej, ale planowany jest budżet strefy euro.
c) Handel międzystanowy (uwaga: zawyża rzeczywisty wskaźnik integracji).
d) Roczne zmiany PKB w latach 1997–2007.
e) 8 makroregionów / 50 stanów dla USA.

Źródło: O'Rourke i Taylor (2013) oraz opracowanie własne.

8.11.2. Kierunki dalszych reform

Reakcje Unii Europejskiej na kryzys finansowy i gospodarczy zmierzały równolegle w dwóch zasadniczych kierunkach – doraźna walka z kryzysem zadłużenia i gruntowna przebudowa systemu zarządzania ekonomicznego. O ile w tym drugim przypadku Unii udało się w miarę szybko zbudować nowe, sprawniejsze

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

ramy koordynacji polityki krajowej, zreformować Pakt Stabilności i Wzrostu, zbudować podstawy unii bankowej i stworzyć trwały mechanizm antykryzysowy, o tyle w pierwszym przypadku państwa członkowskie długo nie były w stanie zażegnać niebezpieczeństwa na rynkach długu suwerennego (Rosati 2011). Dopiero zdecydowana interwencja Europejskiego Banku Centralnego w 2012 roku przywróciła zaufanie i stabilność na rynkach finansowych oraz przyczyniła się do zakończenia kryzysu.

Jak wcześniej wspomniano, kryzys zadłużenia suwerennego uświadomił politykom i opinii publicznej, że wraz z wprowadzeniem wspólnej waluty silnie wzrosła współzależność poszczególnych gospodarek wchodzących w skład strefy euro. Okazało się więc, że utrzymanie unii walutowej wymaga sprawniejszej koordynacji polityki wśród poszczególnych państw, aby ograniczyć skalę potencjalnych szkodliwych efektów *spillovers*. Koordynacja ta może nadal opierać się na uznanych powszechnie regułach prawnych (*rules-based system*) bądź też wymagać powołania nowych, wspólnych instytucji wyposażonych w kompetencje decyzyjne (*institution-based system*). Ze względu na utrzymujący się duży zakres suwerenności rządów narodowych w polityce gospodarczej, zwłaszcza fiskalnej, dotychczasowe reformy systemu zarządzania ekonomicznego zmierzały głównie w kierunku udoskonalania istniejących reguł, takich jak reguły fiskalne, oraz tworzenia nowych, takich jak reguły makroekonomicznej równowagi.

Równolegle powstają jednak także nowe instytucje – m.in. poszczególne składowe unii bankowej i permanentny mechanizm antykryzysowy. Tendencja do stopniowego przenoszenia kompetencji na nowe ponadnarodowe instytucje jest widoczna zwłaszcza w strefie euro – planuje się tu podjęcie dalszych działań na rzecz silniejszej koordynacji oraz integracji, w tym utworzenie odrębnego budżetu strefy euro wraz z możliwością dokonywania transferów fiskalnych między państwami (zob. ramka 8.15), utworzenie stanowiska „ministra finansów" strefy euro wyposażonego w prawo do emisji wspólnych obligacji czy też powołanie w ramach Parlamentu Europejskiego odrębnej izby złożonej z przedstawicieli państw strefy euro[60].

[60] Zob. np. *Towards a Genuine Economic and Monetary Union*, by Herman Van Rompuy, President of the European Council, in close collaboration with José Manuel Barroso, President of the European Commission, Jean-Claude Juncker, President of the Eurogroup, Mario Draghi, President of the European Central Bank, Brussels, December 5, 2012 oraz *Completing Europe's Economic and Monetary Union*, report by Jean-Claude Juncker in close cooperation with Donald Tusk, Jeroen Dijsselbloem, Mario Draghi, and Martin Schulz, European Commission, Brussels, 22 June, 2015.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Ramka 8.15. Wspólny budżet strefy euro

Omówione wcześniej niedoskonałości związane z funkcjonowaniem SGP i systemu dyscypliny fiskalnej pokazały, że krajowa polityka fiskalno-budżetowa nie była w stanie pełnić funkcji instrumentu reagowania na szoki asymetryczne. Nadmierne zadłużenie uniemożliwiło krajom, takim jak Grecja, Portugalia czy Włochy, wykorzystanie polityki fiskalnej w celu przezwyciężenia recesji, ponieważ nie dysponowały one praktycznie żadną swobodą fiskalną (*fiscal space* – zob. rozdział 3, punkt 3.7.5). Oznacza to, że te państwa, które najbardziej potrzebowały stymulacji fiskalnej, były jednocześnie tymi, które nie miały takiej możliwości, głównie ze względu na wcześniejsze naruszenia dyscypliny fiskalnej. Z kolei rozległa przestrzeń fiskalna istniała w krajach o niskim zadłużeniu, które nie potrzebowały fiskalnej stymulacji (np. Niemcy).

Jednym z rozpatrywanych rozwiązań tego paradoksu jest właśnie utworzenie w strefie euro wspólnego budżetu, zwanego też roboczo „zdolnością fiskalną" (*fiscal capacity*), który mógłby reagować na wstrząsy asymetryczne w krajach dysponujących niską lub zerową przestrzenią fiskalną. Taka propozycja pojawiła się m.in. w tzw. Raporcie Pięciu Przewodniczących z 2015 roku[61]. Wspólny budżet miałby częściowo niwelować – poprzez transfery międzyregionalne – skutki wstrząsów asymetrycznych. Zgodnie z tą propozycją, wspólny budżet pełniłby funkcję mechanizmu publicznego dzielenia się ryzykiem (*public rusk sharing*), umożliwiając jednocześnie rozkładanie skutków wstrząsów popytowych i podażowych między regionami unii walutowej. Taki mechanizm funkcjonuje na przykład w USA, gdzie federalne transfery dokonywane między poszczególnymi stanami pozwalają łagodzić od 15% do 30% skutków wstrząsów asymetrycznych (Arreaza i in. 1998, IMF 2013). W Unii Europejskiej taki mechanizm praktycznie nie istnieje, ponieważ budżet UE stanowi zaledwie 1% jej PKB, a ponadto nie jest ani podporządkowany celom stabilizacyjnym, ani nie stanowi mechanizmu dzielenia się ryzykiem.

Zgodnie z niektórymi pozostałymi propozycjami wspólny budżet mógłby też pełnić inne funkcje, na przykład odgrywać rolę swoistej rezerwy na czas gorszej koniunktury (tzw. *rainy fund*), uzupełniać krajowe systemy świadczeń dla

[61] *Completing Europe's Economic and Monetary Union*, report by Report by: Jean-Claude Juncker in close cooperation with Donald Tusk, Jeroen Dijsselbloem, Mario Draghi, and Martin Schulz, European Commission, Brussels, 22 June, 2015.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

bezrobotnych czy też wspierać inwestycje publiczne. Finansowanie wspólnego budżetu miałoby pochodzić głównie z emisji obligacji i ewentualnie z zysku EBC, a w mniejszym stopniu z egzekwowania dodatkowych podatków.

Propozycja utworzenia wspólnego budżetu w ramach strefy euro budzi jednak wiele wątpliwości. Przede wszystkim nie eliminuje to przyczyn niedostatecznej dyscypliny fiskalnej w niektórych krajach. Jak już wspomniano, głównym źródłem ich kłopotów było nierespektowanie reguł deficytu i długu, co prowadziło do nadmiernego zadłużania się. To z kolei ograniczało swobodę fiskalną i nie pozwalało na stymulowanie gospodarki w ramach polityki fiskalnej w okresach spowolnienia gospodarczego. Bez zasadniczego udoskonalenia systemu dyscypliny fiskalnej istnieje duże ryzyko „pokusy nadużycia", co oznacza, że lekkomyślne zachowania niektórych krajów będą mogły się powtarzać. W takich warunkach wspólny budżet przeobrazi się w mechanizm stałego, jednokierunkowego transferu finansowego od krajów fiskalnie odpowiedzialnych do krajów mniej odpowiedzialnych. Oczywiście, nikt się na to nie zgodzi.

Lecz zakładając nawet, że transferów tego typu można będzie w przyszłości uniknąć, pozostaje problem finansowania wspólnego budżetu. Zwiększenie zakresu dzielenia się ryzykiem za pośrednictwem przekazywania środków publicznych do wspólnej kasy może okazać się politycznie niewykonalne. Biorąc pod uwagę obecną trudną sytuację gospodarczą w wielu krajach strefy euro, w tym zwłaszcza wysoki poziom zadłużenia publicznego, wysokie bezrobocie i niskie tempo wzrostu, trudno sobie wyobrazić, aby państwa członkowskie były gotowe wyasygnować środki do wspólnego budżetu o wartości co najmniej 2–3% procent PKB (bo o takiej skali wspólnego budżetu trzeba mówić, jeśli ma on rzeczywiście pełnić funkcję stabilizacyjną). Powstaje także pytanie o to, jak zagwarantować, aby system zarządzania i kontroli nad całym mechanizmem redystrybucji środków był wiarygodny i wolny od nacisków politycznych.

Wreszcie, propozycja wspólnego budżetu nie uwzględnia tego, że system prywatnego podziału ryzyka jest znacznie skuteczniejszy niż system publicznego podziału ryzyka, ponieważ pozwala łagodzić znacznie większą część skutków wstrząsów asymetrycznych. Na przykład w USA za pośrednictwem budżetu federalnego niwelowane jest tylko 13% skutków wstrząsów asymetrycznych, podczas gdy w ramach system prywatnego podziału ryzyka – za pośrednictwem rynku kredytowego i rynków kapitałowych – wygładzane jest 62% skutków wstrząsów (Asdrubali, Sørensen i Yosha 1996; Sørensen i Yosha 1998). Oznacza to, że bardziej obiecującym kierunkiem budowy mechanizmu niwelowania skutków szoków

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

asymetrycznych w strefie euro byłaby pełna integracja rynków kapitałowych, czyli zbudowanie „unii kapitałowej", zamiast tworzenia wspólnego budżetu.

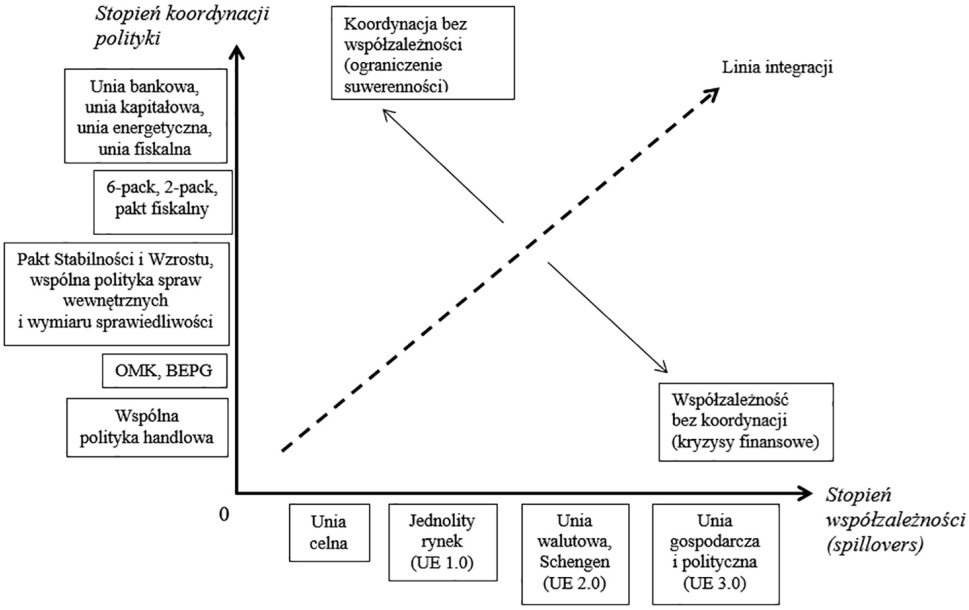
Współzależność między postępem integracji gospodarczej i politycznej a koniecznością coraz silniejszej koordynacji polityki gospodarczej wśród poszczególnych państw odzwierciedla linia integracji, przedstawiona na rysunku 8.10. Na osi poziomej odłożono stopień współzależności i skalę potencjalnych efektów *spillovers*, które wynikają z realizacji kolejnych wielkich projektów integracyjnych – utworzenia unii celnej, powołania Jednolitego Rynku Europejskiego i unii walutowej oraz planowanego utworzenia zaawansowanej unii gospodarczej i politycznej. Na osi pionowej odłożono stopień koordynacji polityki narodowej poszczególnych państw w ramach coraz bardziej rozbudowanych instrumentów polityki unijnej. Przerywana linia integracji przedstawia ogólny kierunek zależności – wyższym poziomom współzależności gospodarczej i finansowej powinny odpowiadać coraz bardziej rozwinięte instrumenty koordynacji polityki w skali unijnej. Odchylanie się od linii integracji pociąga za sobą dwa możliwe rodzaje negatywnych konsekwencji. Rosnąca współzależność bez dostatecznej koordynacji – obszar położony na prawo i poniżej linii integracji – rodzi ryzyko szoków i niestabilności ze względu na wpływ negatywnych efektów *spillovers*. Z kolei rosnąca koordynacja polityki bez rzeczywistej współzależności – obszar położony na lewo i powyżej linii integracji – stwarza ryzyko nieuzasadnionej utraty suwerenności w polityce.

W jakim kierunku będzie zmierzał dalszy rozwój Unii Europejskiej? Kryzysowe wydarzenia z lat 2008–2012 oraz reformy instytucjonalne podjęte w reakcji na to załamanie pokazują, że w ramach UE dochodzi do wyraźnego i trwałego podziału na dwie grupy krajów. Do jednej z nich należą te państwa członkowskie strefy euro, które zdecydowały o dalszym pogłębieniu integracji i podejmują działania zmierzające do zacieśnienia wzajemnej współpracy i koordynacji polityki. Obecnie w ramach istniejącej Unii Europejskiej powstaje więc „nowa Unia" „twardy rdzeń" – złożony z państw, które łączy wspólna waluta, a także zbliżone cele polityczne, w tym mniej lub bardziej silne przekonanie o konieczności utrzymania europejskiego modelu społecznej gospodarki rynkowej. Można oczekiwać, że wśród tych krajów będzie obowiązywać zwiększony zakres solidarności finansowej, ekonomicznej i politycznej. Z drugiej zaś strony istnieją jeszcze pozostałe państwa członkowskie UE, które nie planują dołączenia do strefy euro i wykazują mniej lub bardziej widoczną wstrzemięźliwość wobec dalszych inicjatyw integracyjnych.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Najbardziej spektakularnym przykładem dystansowania się od głównego nurtu integracji jest oczywiście Wielka Brytania, która w drodze negocjacji zagwarantowała sobie wiele wyłączeń z wspólnej polityki europejskiej – nie należy zarówno do strefy euro, jak i do strefy Schengen, nie bierze udziału w prowadzeniu wspólnej polityki azylowej i migracyjnej UE, a także nie uczestniczy w polityce spraw wewnętrznych i wymiaru sprawiedliwości[62]. Warto jednak zauważyć, że podobne tendencje można zaobserwować także w innych krajach.

Rysunek 8.10. Współzależność i koordynacja polityki w Unii Europejskiej



Źródło: opracowanie własne.

8.11.3. Czy Polska powinna dołączyć do strefy euro?

Proces stopniowego różnicowania się państw członkowskich UE na dwie grupy – tzw. „rdzeń" składający się z państw członkowskich strefy euro oraz kraje pozostałe – stawia przed Polską wielkie wyzwanie. Jak należy zareagować na ten wyłaniający się trwały podział wśród państw członkowskich UE? Czy Polska powinna

[62] Mimo to w referendum przeprowadzonym 24 czerwca 2016 roku Brytyjczycy stosunkiem głosów 51,9% do 48,1% wypowiedzieli się za wyjściem Wielkiej Brytanii z Unii Europejskiej.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

dołączyć do krajów „rdzenia", stając się w ten sposób częścią obszaru pogłębionej integracji i czerpiąc korzyści ze zwiększonego zakresu solidarności finansowej oraz bezpieczeństwa a także z możliwości wywierania większego wpływu na kierunek polityki europejskiej? Czy powinna raczej pozostać poza strefą euro, zachowując w ten sposób większy poziom autonomii w polityce wewnętrznej, lecz ryzykując jednocześnie pozostaniem na peryferiach Europy i godząc się z utratą korzyści w związku ze zmniejszonym zakresem solidarności europejskiej, a także tracąc wpływ na bieg spraw unijnych? Zdania wśród polskich ekonomistów są w tej kwestii podzielone[63] i wynikają głównie z ich podejścia do rezygnacji z waluty narodowej i wprowadzenia euro. Lecz jest to zbyt wąska perspektywa, niepozwalająca na właściwe ujęcie problemu. Warto bowiem pamiętać, że uczestnictwo w „rdzeniu" to nie tylko konieczność wprowadzenia wspólnej waluty, ale także harmonizacja prawa w wielu dziedzinach i ściślejsza koordynacja polityki narodowej, również w zakresie polityki zagranicznej i bezpieczeństwa. To również dostęp do dodatkowych instrumentów wsparcia finansowego i potencjalnie silniejszy wpływ na politykę innych państw członkowskich. Decyzja dotycząca tego, czy nasz kraj ma dołączyć do strefy euro, czy też pozostać przy własnej walucie, może więc przesądzić na wiele lat o perspektywach rozwoju i miejscu, jakie Polska będzie zajmowała w Europie.

Literatura cytowana

Alesina A., Angeloni I., Etro F., *The Political Economy of International Unions*, "NBER Working Paper" 2001, no. 8645.

Alesina A., Barro R.J., *Currency unions*, "Quarterly Journal of Economics" 2002, vol. 117, no. 2.

Alesina A., Barro R.J., Tenreyro S., *Optimal Currency Areas*, "NBER Working Paper" 2002, no. 9072.

Alesina A., Angeloni I., Etro F., *International Unions*, "American Economic Review" 2005, vol. 95, no. 3.

Arreaza A., Sorensen B.E., Yosha O., *Consumption smoothing through fiscal policy in OECD and EU countries*, "NBER Working Paper" 1998, no. 6372.

[63] Systematyczny przegląd argumentów „za" i „przeciw" wejściu Polski do strefy euro można znaleźć w raportach Narodowego Banku Polskiego (NBP 2004, 2009, 2014). Na ten temat pisali także, między innymi, Bąk (2014), Borowski (2008), Brzoza-Brzezina i in. (2014), Domański i in. (2008), Kaźmierczak (2013), Łon (2007), Osiatyński (2011), Rosati (2013b), Sławiński (2008), Tomkiewicz (2013) i Wojtyna (2015).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Astrubaldi P., Sørensen B.E., Yosha O., *Channels of interstate risk sharing. The United States 1963–1990*, "Quarterly Journal of Economics" 1984, vol. 111, no. 4.

Atkeson A., Bayoumi T., *Do private capital markets insure regional risk? Evidence from the United States and Europe*, "Open Economies Review" 1993, vol. 4, issue 3.

Axelrod R., *The Evolution of Cooperation*, Basic Books, New York 1984.

Baldacci E., Gupta S., Mulas-Granados C., *How effective is fiscal policy response in systemic banking crises?*, "IMF Working Paper" 2009, no. 160.

Bayoumi T., Eichengreen B., *Shocking aspects of European monetary integration*, w: *Adjustment and Growth in the European Monetary Union*, red. F. Torres, F. Givazzi, Cambridge University Press and CEPR, Cambridge 1992.

Bayoumi T., Masson P., *Fiscal flows in the United States and Canada: Lessons for monetary union in Europe*, "European Economic Review" 1995, vol. 39, issue 2.

Bąk H., *Wybrane problem funkcjonowania strefy euro – wnioski dla Polski*, w: *Wybrane aspekty funkcjonowania Polski w Unii Europejskiej*, red. G. Wojtkowska-Łodej, H. Bąk, Oficyna Wydawnicza SGH, Warszawa 2014.

Bąk H., Maciejewski S., *Asymmetric shocks and international risk sharing in the European Monetary Union and the European Union*, "Bank i Kredyt" 2015, no. 46(6).

Benassy-Quéré A., Qoeuré P., Jacquet P., Pisani-Ferry J., *Economic Policy. Theory and Practice*, Oxford University Press, Oxford–New York 2010.

Blanchard O.J., *Portugal, Italy, Spain and Germany. The implications of a suboptimal currency area*, WEL–MIT, New York 2006.

Borowski J., *Potential Benefits of Poland's EMU Accession*, "Focus on Transition" 2003, no. 1.

Borowski J., *Czy warto przyjąć euro w Polsce?*, „Forum Obywatelskiego Rozwoju", z. 4, Warszawa 2008.

Borowski J., *Integracja monetarna*, PWE, Warszawa 2011.

Brouwer J., Paap R., Viaene J.M., *The trade and FDI effects of EMU enlargement*, "Journal of International Money and Finance" 2008, no. 27(2).

Brzoza-Brzezina M., Makarski K., Wesołowski G., *Would it have paid to be in the eurozone?*, "Economic Modelling" 2014, vol. 41.

Buiter W.H., Corsetti G., Roubini N., *Excessive deficits: sense and nonsense in the Treaty of Maastricht*, "Economic Policy" 1993, issue 1, przedruk w: *The Political Economy of Monetary Union*, red. P. de Grauwe, Edward Elgar Publishing, Cheltenham 2001.

Buti M., Sapir A., *Economic Policy in EMU*, Clarendon Press, Oxford 1998.

Buti M., Turrini A., *Three waves of convergence. Can Eurozone countrias start growing together again?*, VOX, CEPR's Policy Portal, 17 April, 2015.

Calmfors L., *What Remains of the Stability Pact and What Next?*, Swedish Institute for European Policy Studies, report no. 8, November 2005.

Cieślik A., Michałek J., Mycielski J., *Analiza skutków handlowych przystąpienia Polski do Europejskiej Unii Monetarnej przy użyciu uogólnionego modelu grawitacyjnego*, WNE, seminarium BISE, NBP, maszynopis, Warszawa, sierpień 2008.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Cohen B., *The future of the euro. Let's get real*, "Review of International Political Economy" 2012, vol. 19, no. 4.

Daras T., Hagemajer J., *The long run-effects of the Poland's accession to the eurozone. Simulation using POLDYN – a dynamic computable general equilibrium model*, Narodowy Bank Polski, Warszawa 2008.

Darvas Z., Szapáry G., *Business Cycle Synchronization in the Enlarged EU: Comovements in the* (*Soon-To-Be*) *New and Old Members*, "Magyar Nemzeti Bank Working Paper" 2004, no. 1.

De Grauwe P., *Economics of Monetary Union*, Oxford University Press, Oxford 2000.

De Grauwe P., *The Challenge of Enlargement of Euroland*, Workshop on EMU: Current State and Future Prospects, University of Crete, Rethymno, 24–31 August, 2003.

DNB, *Going on for twelve: results of the 12th DNB-eurosurvey*, De Nederlandsche Bank, North Holland, September 2001.

Domański S.R., Kazimierczak A, Żyżyński J., *Polska wobec perspektywy wstąpienia do strefy euro. Za i przeciw szybkiej integracji walutowej*, ekspertyza wykonana dla Kancelarii Prezydenta RP, maszynopis, Warszawa 2008.

Drummond P. i in., *Adopting the Euro in Central Europe – Challenges of the Next Step in European Integration*, " IMF Occasional Paper" 2005, no. 234.

Duarte R., Marques C.R., *The dynamic effects of shocks to wages and prices in the United States and the Euro Area*, "Empirical Economics" 2013, no. 44(2).

Dziawgo D., *Credit rating Grecji w kontekście kryzysu finansowego*, "Annales Universitatis Mariae Curie-Skłodowska" 2012, t. XLVI, nr 4.

Eichengreen B., *Should the Maastricht Treaty Be Saved?*, "Princeton Studies in International Finance" 1992, no. 74.

Eichengreen B., *European Monetary Unification: Theory, Practice, and Analysis*, MIT Press, Cambridge 1997.

Eichengreen B., Bayoumi T., *Shocking Aspects of European Monetary Unification*, w: *Growth and Adjustment in the European Monetary Union*, red. F. Torres, F. Givazzi, Cambridge University Press, Cambridge 1997.

Eichengreen B., Steiner K., *Is Poland at Risk of a Boom-and-Bust Cycle in the Run-Up to Euro Adoption?*, paper prepared for the NBP Conference on Euro Adoption, Warsaw, September 2008.

European Commission, *One Market, One Money: an evaluation of the potential benefits and costs of forming an economic and monetary union*, "European Economy" 1990, no. 44.

European Monetary Institute, *Convergence Report*, *Box 1. Price stability*, March 1998.

Feldstein M., *Distinguished Lecture on Economics in Government: Thinking about International Economic Coordination*, "The Journal of Economic Perspectives" 1988, vol. 2, no. 2.

Feldstein M., *The Political Economy of the European Economic and Monetary Union: Political Sources of an Economic Liability*, "Journal of Economic Perspectives" 1997, no. 11(4).

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Feldstein M., *Europe Can't Handle the Euro*, "Wall Street Journal", 8 February, 2000.

Frankel J., *No Single Currency Regime is Right for all Countries or at All Times*, "Princeton Essays in International Finance" 1999, no. 215.

Frankel J., *Should Central European Countries Join the Euro? A Review and Update of Trade Estimates and Consideration of Endogenous OCA Criteria*, paper prepared for the NBP Conference on Euro Adoption, Warsaw 2008.

Frankel J., Rose A., *The Endogeneity of the Optimum Currency Area Criterion*, "Economic Journal" 1998, no. 108.

Frenkel J., Goldstein M., Masson P., *International coordination of economic policies: scope, methods and effects*, "NBER Working Papers" 1988, no. 26670.

Furceri D., Zdzienicka A., *The Euro Area Crisis: Need for a Supranational Risk Sharing Mechanism?*, "IMF Working Paper" 2013, no. WP/13/98.

Gros D., Alcidi C., *Economic Policy Coordination in the Euro Area under the European Semester*, CEPS, Brussels, 10 December, 2015.

IFM, *Euro und Mittelstand, Schriften zur Mittelstandforschung*, Institut für Mittelstandforschung, Bonn 1998.

IMF, *Towards a Fiscal Union for the Euro Area*, IMF Staff Discussion Note, SDN/13/09, 2013.

IMF, *Gross domestic product*, "IMF World Economic Outlook", October 2015.

Jonung L., Drea E., *It Can't Happen, It's a Bad Idea, It Won't Last: U.S. Economists on the EMU and the Euro 1989–2002*, "Econ Journal Watch" 2009, vol. 7, no. 1.

Kawalec S., Pytlarczyk E., *Kontrolowana dekompozycja strefy euro, aby uratować Unię Europejską i jednolity rynek,* Warszawa 2012 (tekst dostępny na stronie: www.liberte.pl).

Kaźmierczak A., *Dylematy integracji Polski ze strefą euro*, w: *Wprowadzenie euro w Polsce – za i przeciw*, Wydawnictwo Sejmowe, Warszawa 2013.

Kenen P., *The Theory of Optimum Currency Areas: An Eclectic View*, w: *Monetary Problems of the International Economy*, red. R. Mundell, A. Swoboda, The University of Chicago Press, Chicago 1969.

Koo R.C., *The Holy Grail of Macroeconomics-Lessons from Japan's Great Recession*, Wiley (Asia), Singapore 2009.

Koo R.C., *The world in balance sheet recession: causes, cure, and politics*, "Real-world Economics Review" 2011, no. 58.

Krugman P., *Increasing Returns and Economic Geography*, "Journal of Political Economy" 1991a, vol. 99.

Krugman P., *Geography and Trade*, MIT Press, Cambridge 1991b.

Krugman P., *Lessons of Massachusetts for EMU*, w: *Adjustment for Growth in the European Monetary Union*, red. F. Torres, F. Givazzi, Cambridge University Press, New York 1993.

Krugman P., *The Making of a Euromess*, "The New York Times", 15 February, 2010.

Krugman P., *End this depression now!*, W.W. Norton and Company, New York 2012.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Krugman P., Venables A., *Integration, specialization and adjustment*, "NBER Working Paper Series" 1993, no. 4559.

Łon E., *Dlaczego Polska nie powinna wchodzić do strefy euro?*, Poznań 2007, mimeo.

Marszałek P., *Coordination of monetary and fiscal policy*, "The Poznań University of Economics Review" 2003, vol. 3, no. 2.

Matkowski Z., Próchniak M., *Zbieżność poziomów dochodu i wahań cyklicznych w krajach Europy Środkowej i Wschodniej w stosunku do Unii Europejskiej*, w: *Wzrost gospodarczy w krajach transformacji. Konwergencja czy dywergencja?*, red. R. Rapacki, PWE, Warszawa 2009.

McKinnon R.I., *Optimum Currency Areas*, "American Economic Review" 1963, vol. 53.

McCallum J., *National Borders Matter: Canada-U.S. Regional Trade Patterns*, "The American Economic Review" 1995, vol. 85, no. 3.

McKinnon R.I., *Optimum Currency Areas and Key Currencies: Mundell I versus Mundell II*, "Journal of Common Market Studies" 2004, vol. 42, no. 4.

Meade J.E., *The Balance of Payments Problems of a Free Trade Area*, "Economic Journal" 1957, vol. 67.

MFW, *Euro Area Policies. 2012 Article IV Consultation IMF*, " IMF Country Report" 2012, no. 12/181.

Miller M., Sutherland A., *The Walters Critique of the EMS – a Case of Inconsistent Expectations?*, "The Manchester School" 1991, vol. LIX.

Mongelli F.P., Wyplosz C., *The euro at ten: unfulfilled threats and unexpected challenges*, 5th ECB Central Banking Conference, The euro at ten: lessons and challenges, 13–14 November, 2008.

Mundell R.A., *A Theory of Optimum Currency Areas*, "American Economic Review" 1961, no. 51.

Mundell R.A., *A Plan for a European Currency*, Paper Prepared for Discussion at the American management Association Conference on Future of the International Monetary System, New York, 10–12 December, 1969.

Mundell R.A., *Uncommon Arguments for Common Currencies*, w: *The Economics of Common Currencies*, red. H.G. Johnson, A.K. Swoboda, Allen and Unwin, London 1973.

NBP, *Raport na temat korzyści i kosztów przystąpienia Polski do strefy euro*, Warszawa 2004.

NBP, *Raport na temat pełnego uczestnictwa Rzeczypospolitej Polskiej w trzecim etapie Unii Gospodarczej i Walutowej*, Warszawa 2009.

NBP, *Ekonomiczne wyzwania integracji Polski ze strefą euro*, Warszawa 2014.

NBS, *The Effects of Euro Adoption on the Slovak Economy*, National Bank of Slovakia Research Department, March 2006, Bratislava.

O'Rourke K.H., Taylor A.M., *Cross of Euros*, "Journal of Economic Perspectives" 2013, vol. 27, no. 3.

Osiatyński J., *Warunki gotowości Polski do wejścia do strefy euro*, „Ekonomista" 2011, nr 5.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Próchniak M., Rapacki R., *Konwergencja typu beta i sigma w krajach transformacji w latach 1990–2005*, w: *Wzrost gospodarczy w krajach transformacji. Konwergencja czy dywergencja?*, red. R. Rapacki, PWE, Warszawa 2009.

Rogoff K., *The Purchasing Power Parity Puzzle*, "Journal of Economic Literature" 1996, vol. 34(2).

Rosati D.K., *The Balassa-Samuelson Effect in the EU Candidate Countries*, w: *Trade, Integration and Transition*, red. R. Grawe, A. Inotai, The World Bank, Institute for World Economics Hungarian Academy of Sciences, Budapest 2002.

Rosati D.K., *Gdy nowe państwa członkowskie UE zabiegają o szybką akcesję, Polska nie ustaliła nawet kalendarza wejścia do strefy euro*, „The Wall Street Journal Polska", dodatek do „Dziennika Polska–Europa–Świat", 24 lipca 2007.

Rosati D.K., *Dlaczego Polska powinna przyjąć euro*, „The Wall Street Journal Polska", dodatek do „Dziennika Polsa–Europa–Świat", 2 czerwca 2008.

Rosati D.K., *Euro – ekonomia i polityka*, Wyższa Szkoła Handlu i Prawa im. R. Łazarskiego, Warszawa 2009.

Rosati D.K., *Kierunki zmian w systemie zarządzania fiskalnego strefą euro*, „Ekonomista" 2010, nr 5.

Rosati D.K., *Will the Eurozone survive? Lessons from the sovereign debt crisis in Europe*, w: *L'Europe après Lisbonne. Quels défis pour l'UE après l'entrée en vigueur du traité de Lisbonne*, UCL – Presses Universitaires de Louvain, Louvain-la-Neuve, 2011.

Rosati D.K., *Czy strefa euro przetrwa kryzys?*, w: *Kryzys finansów publicznych. Przyczyny, mechanizm, drogi wyjścia*, red. D.K. Rosati, L. Oręziak, Wyższa Szkoła Handlu i Prawa im. R. Łazarskiego, Warszawa 2013a.

Rosati D.K., *Czy Polska powinna przystąpić do strefy euro*?, „Gospodarka Narodowa" 2013b, R. LXXXIII/XXIV.

Rosati D.K., *Asymmetric shocks in the Euro Area: convergence or divergence?*, Warsaw School of Economics, December 2016, mimeo.

Rose A., *One Money, One Market: The Effect of Common Currencies on Trade*, "Economic Policy" 2000, vol. 15.

Rose A., *EMU, Trade and Business Cycle Synchronization: What Do We Know and What Does it Mean for Poland?*, paper prepared for the NBP Conference on Euro Adoption, Warsaw, September 2008.

Roubini N., Mihm S., *Ekonomia kryzysu*, Wolters Kluwer, Warszawa 2011.

Scitovski T., *Economic Theory and Western European Integration,* Stanford 1958.

Shambaugh J.C., *The Euro's Three Crises*, "Brookings Papers on Economic Activity" 2012, Spring.

Sławiński A., *Znaczenie czynników ryzyka towarzyszących wchodzeniu Polski do ERM2 i do strefy euro*, „Ekonomista" 2008, nr 1.

Sørensen B.E., Yosha O., *International risk sharing and European monetary unification*, "Journal of International Economics" 1998, vol. 45.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237

Šuster M., *Euro introduction in Slovakia*, Narodna Banka Slovenska, paper presented at a conference in Vilnius, 1 April, 2009.

Tesar L.L., Werner I.M., *Home bias and globalization of securities markets*, "NBER Working Paper" 1992, no. 4218.

Tobin J., *Agenda for International Coordination of Macroeconomic Policies*, w: *International Monetary Cooperation: Essays in Honor of Henry C. Wallich*, red. P.A. Volcker i in., "Essays in International Finance" 1987, no. 169.

Tomkiewicz J., *Wejście do strefy euro a konkurencyjność gospodarki*, w: *Wprowadzenie euro w Polsce – za i przeciw*, Wydawnictwo Sejmowe, Warszawa 2013.

Von Hagen J., *Fiscal policy and international risk sharing*, w: *International macroeconomics*, red. G.D. Hess, E. van Wincoop, Cambridge University Press, Cambridge 2000.

Wallich H.C., *Institutional Cooperation in the World Economy*, w: *The World Economic System: Performance and Prospects*, red. J. Frenkel, M. Mussa, Auburn House, Dover 1984.

Werning I., Fahri E., *Fiscal Unions*, "NBER Working Paper Series" 2012, no. 18280.

Wojtyna A., *Nie odkładajmy tematu euro*, „Rzeczpospolita", 17 maja 2015.

Plik zabezpieczony watermarkiem jawnym i niejawnym: 8098123A35373237