BARBARA DAŃSKA-BORSIAK

Dynamiczne modele panelowe w badaniach ekonomicznych

WYDAWNICTWO
UNIWERSYTETU ŁÓDZKIEGO • ŁÓDŹ 2011

RECENZENT
Paweł Miłobędzki

REDKATOR WYDAWNICTWA UŁ
Bożena Tkacz

REDAKTOR TECHNICZNY
Jolanta Kędzierska

KOREKTORZY
Bogusława Kwiatkowska, Bożena Tkacz

OKŁADKĘ PROJEKTOWAŁA
Barbara Grzejszczak

Praca naukowa jako projekt badawczy MNiSW nr N111093833 finansowana
ze środków na naukę w latach 2007–2009

© Copyright by Barbara Dańska-Borsiak 2011

Wydawnictwo Uniwersytetu Łódzkiego
2011

Wydanie I. (dodruk)
Zam. 4806/2011

90-131 Łódź, ul. Lindleya 8
www.wydawnictwo.uni.lodz.pl
e-mail: ksiegarnia@uni.lodz.pl
tel. (42) 665 58 63, faks (42) 665 58 62

ISBN 978-83-7525-549-2
https://doi.org/10.18778/7525-549-2

SPIS TREŚCI

WSTĘP 7

ROZDZIAŁ 1. DANE PANELOWE W MODELOWANIU EKONOMETRYCZNYM – KLASYFIKACJA, ROZWÓJ HISTORYCZNY METODOLOGII 13
1.1. Klasyfikacja danych i modeli czasowo-przekrojowych i panelowych 14
1.2. Dane panelowe w modelowaniu makroekonomicznym (makropanele) i mikroekonomicznym (mikropanele). Źródła danych statystycznych 17
1.3. Zalety wykorzystania danych panelowych w praktyce ekonometrycznej 19
1.4. Początki rozwoju metodologii estymacji modeli panelowych 22
1.5. Współczesne kierunki badań 29
1.6. Sposoby uwzględnienia dynamiki modelowanych zjawisk 32
1.7. Rola warunków początkowych w procesie estymacji dynamicznych modeli panelowych 35

ROZDZIAŁ 2. ZARYS PROBLEMATYKI ESTYMACJI MODELI STATYCZNYCH NA PODSTAWIE DANYCH PANELOWYCH I PRZEKROJOWO-CZASOWYCH 39
2.1. Modele panelowe ze zmiennymi sztucznymi (modele FE) 40
2.1.1. Założenia konstrukcyjne i metody estymacji 40
2.1.2. Testowanie istotności efektów grupowych 44
2.2. Modele panelowe z dekompozycją składnika losowego (modele RE) 44
2.2.1. Założenia konstrukcyjne i metody estymacji 45
2.2.2. Testowanie istotności efektów grupowych 48
2.3. Model FE czy model RE 49
2.3.1. Wskazówki ogólne 49
2.3.2. Test Hausmana 50
2.4. Model Hausmana i Taylora 51
2.5. Metody estymacji modeli wykorzystujących dane czasowo-przekrojowe 55
2.5.1. Modele z heteroskedastycznością grupową składnika losowego 56
2.5.2. Modele z heteroskedastycznością grupową i korelacją przekrojową składnika losowego 58
2.5.3. Modele z heteroskedastycznością grupową, korelacją przekrojową i autokorelacją składnika losowego 63
2.6. Panelowe modele wielorównaniowe 66
2.6.1. Metody estymacji modeli prostych o równaniach pozornie niezależnych (SUR) 66
2.6.2. Metody estymacji modeli wielorównaniowych o równaniach łącznie współzależnych 68

ROZDZIAŁ 3. PROBLEMY ESTYMACJI I TESTOWANIA DYNAMICZNYCH MODELI PANELOWYCH 73
3.1. Panelowe modele statyczne a dynamiczne – własności estymatorów 73

3.2. Metody estymacji dynamicznych modeli panelowych 78
3.2.1. Estymator metody zmiennych instrumentalnych Andersona i Hsiao 81
3.2.2. Estymator GMM pierwszych różnic Arellano i Bonda (FDGMM) 87
3.2.3. Skorygowany estymator wariancji Windmeijera 95
3.2.4. Estymator Arellano i Bover 97
3.2.5. Estymator Ahna i Schmidta 102
3.2.6. Systemowy estymator GMM Blundella i Bonda 106
3.2.7. Skorygowany estymator wewnątrzgrupowy Kivieta 111
3.3. Testowanie hipotez 115
3.4. Porównanie własności estymatorów 118

ROZDZIAŁ 4. PRZYKŁADY ZASTOSOWAŃ DYNAMICZNYCH MODELI PANELOWYCH W BADANIACH MIKROEKONOMICZNYCH 123
4.1. Inwestycje w firmach 124
4.2. Produkcja i produktywność w firmach 127
4.3. Modelowanie zatrudnienia w przedsiębiorstwach 131
4.4. Innowacje a zatrudnienie w przedsiębiorstwach 134
4.5. Kształtowanie się wynagrodzeń w firmach 136
4.6. Struktura właścicielska i kapitałowa firmy 138
4.7. Mobilność pracowników i transfer kwalifikacji 144
4.8. Problem błędów pomiaru zmiennych 146

ROZDZIAŁ 5. ZASTOSOWANIE DYNAMICZNYCH MODELI PANELOWYCH W ANALIZACH SEKTOROWYCH 151
5.1. Podstawy ekonomiczne dla analiz produktywności 152
5.2. Model TFP według sektorów przemysłu w Wielkiej Brytanii 157
5.3. Wpływ rynków finansowych na TFP w sektorach 160
5.4. Wpływ technologii informacyjno-telekomunikacyjnych (ICT) na wyniki w sektorach . 161
5.5. Produktywność czynników produkcji w sekcji D w Polsce 163
5.6. Szacowanie łącznej produktywności czynników produkcji (TFP) 168
5.6.1. Szacowanie TFP na podstawie modelu wydajności 168
5.6.2. Szacowanie TFP przy założeniu doskonałych rynków czynników produkcji . . 173
5.7. Modelowanie kształtowania się łącznej produktywności czynników produkcji 175

ROZDZIAŁ 6. ZASTOSOWANIE DYNAMICZNYCH MODELI PANELOWYCH W BADANIACH REGIONALNYCH 181
6.1. Podstawy ekonomiczne teorii wzrostu i konwergencji 182
6.2. Modelowanie wzrostu gospodarczego w latach 90. XX w. 185
6.3. Współczesne modelowanie wzrostu gospodarczego 186
6.4. Determinanty TFP w regionach 189
6.5. Analiza β-konwergencji województw 193
6.5.1. Postać ogólna modelu β-konwergencji i specyfikacja zmiennych 194
6.5.2. Wyniki estymacji modelu β-konwergencji 196
6.5.3. Próba uwzględnienia kapitału ludzkiego w modelu wzrostu województw w Polsce 204
6.6. Poziom wykształcenia społeczeństw a wskaźnik demokracji 208

ZAKOŃCZENIE 211

DODATEK I. TESTY PIERWIASTKÓW JEDNOSTKOWYCH DLA DANYCH PANELOWYCH 215

DODATEK II. UOGÓLNIONA METODA MOMENTÓW (GMM) DLA MODELI DANYCH CZASOWYCH 227

WYKAZ OZNACZEŃ 237

BIBLIOGRAFIA 239

INDEKS NAZWISK Z POWOŁAŃ NA BIBLIOGRAFIĘ 251

INDEKS RZECZOWY 255

WSTĘP

We współczesnych badaniach ekonomicznych zauważalny jest wzrost zapotrzebowania na nowoczesne metody analityczne. Należą do nich metody analiz danych czasowo-przekrojowych oraz odpowiednie modele ekonometryczne. Dane te dotyczyć mogą obiektów o różnorodnym charakterze – sektorów gospodarki, firm, regionów *etc*. Szczególną klasę danych łączących wymiar czasowy i przekrojowy stanowią dane panelowe (patrz podrozdz. 1.1). Zawierają one znacznie więcej obserwacji dla obiektów (nawet kilkaset) niż dla okresów (kilka). Zastosowanie danych panelowych umożliwia, w porównaniu z szeregami czasowymi lub danymi przekrojowymi, dokonanie bardziej wnikliwych, szczegółowych analiz ekonomicznych. Możliwe jest bowiem wówczas wyjaśnienie różnic w zachowaniu różnych obiektów w danym okresie oraz różnic w zachowaniu wybranego obiektu w poszczególnych okresach próby. Wykorzystanie danych panelowych do estymacji modeli ekonometrycznych powoduje jednocześnie wiele komplikacji metodologicznych, niewystępujących w modelowaniu na podstawie danych czasowych lub przekrojowych. Wynikają one z wielowymiarowości danych.

Szczególną klasą modeli ekonometrycznych są dynamiczne modele panelowe. Charakteryzują się one istotnym wpływem na zmienną objaśnianą czynnika czasu, który może być uwzględniony w modelu poprzez wprowadzenie zmiennych z opóźnieniami czasowymi, w szczególności opóźnionej zmiennej endogenicznej. Poprawne oszacowanie modeli tej klasy wymaga zastosowania całkowicie odrębnych metod estymacji. Różnią się one zarówno od metod standardowo stosowanych w klasycznym modelowaniu dynamicznym, jak i od metod stosowanych do szacowania statycznych modeli panelowych. Wynika to z faktu, że opóźniona zmienna endogeniczna, występująca w roli zmiennej objaśniającej modelu, jest skorelowana z tzw. efektami grupowymi, które są wyróżniane w procesie konstrukcji modeli panelowych jako składowe odzwierciedlające efekt działania stałych w czasie czynników specyficznych dla danego obiektu.

Tematyka niniejszej monografii koncentruje się na modelach z mierzalną zmienną objaśnianą, co jest odzwierciedleniem wieloletnich zainteresowań naukowych autorki, rozwijanych na gruncie metodologicznym i empirycznym,

których efektem była m. in. wcześniejsza publikacja książkowa (D a ń s k a [2000]) traktująca o modelach statycznych. Zagadnienia dotyczące dynamicznych modeli panelowych innych klas (m. in. modeli dla zmiennych jakościowych i uciętych, paneli niezbilansowanych) omówione zostaną w kolejnej publikacji książkowej.

Głównym celem badawczym zrealizowanym w niniejszej monografii jest analiza własności estymatorów dynamicznych modeli panelowych pod kątem ich przydatności do analiz empirycznych. Dokonano przy tym przeglądu zagadnień o charakterze metodologicznym i praktycznym, związanych z wykorzystaniem danych panelowych do konstrukcji, estymacji i wnioskowania na podstawie modeli dynamicznych.

W trakcie badań zrealizowano również następujące cele cząstkowe, które służą osadzeniu prezentowanej metodologii w szerszym kontekście:

– przedstawienie zalet i wad zastosowania danych panelowych w modelowaniu ekonometrycznym i zwrócenie uwagi na zasadnicze różnice, występujące między modelami dynamicznymi i statycznymi, estymowanymi na podstawie danych panelowych;

– zarysowanie historycznego rozwoju metodologii i zaprezentowanie współcześnie proponowanych, alternatywnych metod estymacji dynamicznych modeli panelowych oraz dokonanie krytycznego porównania ich własności;

– przedstawienie korzyści wynikających z zastosowania dynamicznych modeli panelowych jako narzędzia analiz empirycznych w różnorodnych obszarach badawczych;

– wskazanie możliwości zastosowania modeli tej klasy do analiz przekrojowo-czasowych w Polsce. Cel ten zrealizowany został poprzez konstrukcję i estymację własnych modeli: modelu opisującego kształtowanie się łącznej produktywności czynników produkcji według działów sekcji „przetwórstwo” oraz modelu konwergencji województw. Tematyka badawcza modeli własnych odzwierciedla wieloletnie zainteresowania naukowe autorki.

Realizacja tak określonych celów: głównego i cząstkowych czyni niniejszą monografię użyteczną dla wszystkich osób prowadzących analizy na podstawie danych panelowych. Może ona być pomocna w rozpropagowaniu na gruncie polskim użytecznego narzędzia wnioskowania, jakim są modele panelowe, jak również stanowić bazę dla rozwoju na gruncie krajowym zarówno nowych metod estymacji, jak i twórczego ich zastosowania.

Metody estymacji dynamicznych modeli panelowych rozwijają się szczególnie intensywnie od początku lat 90. XX w. W literaturze anglojęzycznej istnieje wiele publikacji związanych z tą tematyką, wciąż pojawiają się też nowe opracowania, w których ulepsza się już istniejące metody bądź proponuje nowe. W Polsce brak jest opracowań w wyczerpujący sposób przedstawiających teoretyczne aspekty modelowania dynamicznego na podstawie danych panelowych. Lukę tę wydaje się wypełniać niniejsza publikacja. W sposób syntetyczny

przedstawiono w niej różnorodne propozycje, od najwcześniejszych do najnowszych. Wyodrębniono przy tym i szczególnie dokładnie omówiono grupę metod estymacji, które charakteryzują się największą uniwersalnością (są to w zasadzie propozycje bazujące na uogólnionej metodzie momentów). Przedstawiono też metody mające nieco węższe zastosowanie w praktyce, które jednak w pewnych szczególnych przypadkach przewyższają metody uniwersalne pod względem własności estymatorów. Wskazano też wady i zalety alternatywnych metod estymacji w zależności od własności szeregów czasowych obserwacji dla obiektów, co pomóc może badaczom w wyborze właściwej metody estymacji dla konkretnych przypadków.

Oprócz problemów metodologicznych drugim nurtem niniejszej monografii jest prezentacja przykładów praktycznego zastosowania dynamicznych modeli panelowych. Dane panelowe są stosowane w modelowaniu ekonometrycznym w badaniach na szczeblu makro-, mezo- i mikroekonomicznym. Różnice między zbiorami danych wykorzystywanymi w tych analizach, dotyczące głównie zależności miedzy wymiarem czasowym a przekrojowym, są istotne szczególnie w przypadku dynamicznych modeli panelowych, gdyż przekładają się na własności stosowanych estymatorów. W niniejszej pracy omówione są reprezentatywne analizy światowe prowadzone w skali mikro-, mezo- i makroekonomicznej, ze szczególnym zwróceniem uwagi na stosowane metody estymacji, właściwe dla poszczególnych prób. Analizy mikroekonomiczne dotyczą różnych aspektów funkcjonowania przedsiębiorstw, analizy mezoekonomiczne – analiz na szczeblu sekcji i działów działalności gospodarczej, a analizy makroekonomiczne – mają charakter regionalny.

Poniżej omówione zostaną krótko treści i cele merytoryczne poszczególnych rozdziałów.

W rozdz. 1 przedstawiona jest klasyfikacja danych czasowo-przekrojowych, przykłady ogólnie dostępnych źródeł takich danych i oprogramowania oraz zalety wykorzystania danych czasowo-przekrojowych w praktyce ekonometrycznej. Omówiony jest także rozwój historyczny metod estymacji dynamicznych modeli panelowych. Dwa ostatnie podrozdziały dotyczą sposobów uwzględnienia dynamiki modelowanych zjawisk oraz roli warunków początkowych w procesie estymacji i wnioskowania.

Rozdział 2 poświęcony jest problematyce szacowania i testowania statycznych modeli panelowych, tzn. modeli szacowanych na podstawie danych, w których liczba obserwacji dla obiektów N przekracza często wielokrotnie liczbę obserwacji w czasie T. Ich konstrukcja zakłada niezmienność parametrów przy zmiennych objaśniających względem czasu i obiektów, a heterogeniczność obiektów lub zróżnicowanie modelowanego zjawiska w czasie uwzględnia się poprzez zróżnicowanie względem obiektów (i/lub czasu) składnika losowego albo, w zależności od założeń, wyrazu wolnego. W literaturze anglojęzycznej modele takie nazywa się odpowiednio *random effects models* albo *fixed effects*

models. Zaprezentowany jest też model Hausmana i Taylora, który dopuszcza, aby część zmiennych objaśniających była skorelowana z efektami grupowymi, a pozostałe zmienne nie wykazywały takiej korelacji. W rozdz. 2 przedstawiono też problematykę szacowania i testowania modeli bazujących na typowych danych czasowo-przekrojowych, w których liczba obserwacji w czasie jest większa niż liczba obiektów. W końcowej części rozdziału zaprezentowano zarys metod estymacji wielorównaniowych modeli statycznych o równaniach pozornie niezależnych i o równaniach łącznie współzależnych.

Zasadniczym celem rozdz. 3 jest zaprezentowanie alternatywnych metod estymacji dynamicznych modeli panelowych oraz ich krytyczne porównanie. Przedstawiono w nim również testy statystyczne służące do weryfikacji jakości estymacji modeli. Prezentację metod estymacji poprzedzono szczegółowym uzasadnieniem przyczyn utraty zgodności i nieobciążoności estymatorów stosowanych dla modeli statycznych, która ma miejsce w przypadku modeli dynamicznych. Uzasadnienie to, jak również opis problemów związanych z szacowaniem efektów długookresowych i krótkookresowych w modelach panelowych, zawarto w podrozdz. 3.1. Podrozdział 3.2 poświęcony jest prezentacji metod estymacji dynamicznych modeli panelowych bazujących na uogólnionej metodzie momentów i metodzie zmiennych instrumentalnych. Są to metody mające największe znaczenie w praktyce, zaś pojawiające się w literaturze przedmiotu propozycje alternatywne wydają się mieć mniejsze znaczenie. Wyjątkiem jest propozycja stosowania estymatora wewnątrzgrupowego z poprawką korygującą jego obciążenie. Ze względu na dobre własności tej metody w małych próbach, została ona przedstawiona w pkt 3.2.7. Testy statystyczne, stosowane do oceny jakości dynamicznych modeli panelowych, opisano w podrozdz. 3.3. Końcowy podrozdz. 3.4 poświęcony jest porównaniu własności estymatorów proponowanych dla dynamicznych modeli panelowych.

Treści zawarte w trzech ostatnich rozdziałach pracy służą wykazaniu uniwersalności narzędzia analitycznego, jakim są modele panelowe. Prezentowane są tam przykłady zastosowania dynamicznych modeli panelowych w analizach empirycznych z obszaru mikro-, mezo- i makroekonomii.

W rozdz. 4 przedstawiono przykłady badań światowych z obszaru ekonomii przemysłowej i finansów przedsiębiorstw, prowadzonych na podstawie dynamicznych modeli panelowych. Przykłady dobrano tak, aby obrazowały korzyści badawcze płynące z zastosowania mikropaneli dla analizy ważnych aspektów funkcjonowania przedsiębiorstw: inwestycji, produktywności, zatrudnienia, innowacji, czynników finansowych czy transferu kwalifikacji.

Rozdział 5 zawiera przykłady zastosowania dynamicznych modeli panelowych w analizach funkcjonowania jednostek klasyfikacji działalności gospodarczej. Oscylują one wokół zagadnień związanych z produktywnością, a w szczególności łączną produktywnością czynników produkcji (TFP). Obok analiz zaczerpniętych z literatury światowej przedstawiono wyniki własnych badań

dotyczących możliwości oszacowania i modelowania TFP w działach sekcji „przetwórstwo” w Polsce. Konstrukcja i estymacja dynamicznych modeli panelowych, opisujących kształtowanie się TFP, wydaje się ważnym osiągnięciem badawczym. Istniejące publikacje, w których analizowano łączną produktywność czynników produkcji, wymienione w podrozdz. 5.1, koncentrują się bowiem na ogół na alternatywnych metodach oszacowania TFP. Unikalność prezentowanych tu badań polega na tym, że oprócz oszacowania wartości TFP według działów podjęto próbę modelowania czynników wpływających na kształtowanie się tej zmiennej w wymiarze przekrojowo-czasowym.

Badania prezentowane w rozdz. 6 koncentrują się wokół zagadnień wzrostu, konwergencji i produktywności w regionach. Także w tym rozdziale, obok przykładów badań światowych prezentowane są analizy własne. Prace dotyczące wzrostu i konwergencji regionalnej, obejmujące regiony Polski wymienione, są w podrozdz. 6.5. Jednakże w opracowaniach tych albo posługiwano się metodami innymi niż panelowe, albo zasięg terytorialny badań był szerszy niż obszar Polski. Wydaje się, że przedstawiona przez autorkę możliwość zastosowania dynamicznych modeli panelowych do badania β-konwergencji województw jest odpowiedzią na istniejące zapotrzebowanie, wyrażone *explicite* w pracy: M a l a g a, K l i b e r [2007] (s. 6): „[...] W konsekwencji zrezygnowaliśmy także z ekonometrycznych metod estymacji, lub ogólniej, szacowania parametrów zastosowanych modeli wzrostu. Jednym z powodów był także brak dostatecznie długich i wiarygodnych szeregów czasowych [...] w układzie województw”. Znaczenie analiz prezentowanych w niniejszej publikacji polega na wskazaniu możliwości rozwiązania tego typu problemów.

Cechą charakterystyczną makropaneli jest na tyle duży wymiar czasowy próby T, że wprowadza się założenie, iż $T \to \infty$. Ze względu na to założenie, do analizy makroekonomicznych danych panelowych zaczęto stosować podejście znane z analiz szeregów czasowych, w szczególności rozwinęły się metody badania stacjonarności w panelach, regresji pozornych i kointegracji. Przegląd testów pierwiastków jednostkowych dla danych panelowych zamieszczono w dodatku I.

Spośród metod proponowanych do estymacji dynamicznych modeli panelowych największe znaczenie praktyczne mają, jak wspomniano powyżej, metody bazujące na uogólnionej metodzie momentów i metodzie zmiennych instrumentalnych. Wydaje się, że o ile metoda zmiennych instrumentalnych jest szeroko znana, to uogólniona metoda momentów jest znacznie mniej rozpowszechniona. Możliwości zastosowania tej metody do estymacji modeli szeregów czasowych przedstawiono w dodatku II, z zamiarem ułatwienia Czytelnikowi śledzenia treści rozdz. 3.

Całość monografii dopełnia zakończenie, w którym podsumowane zostały treści zamieszczone w publikacji, a także zarysowane możliwe kierunki dalszych badań. Zawiera ona ponadto obszerną bibliografię tematu.

Powstanie pracy nie byłoby możliwe bez życzliwości kierownika Katedry Ekonometrii Przestrzennej Uniwersytetu Łódzkiego, prof. Bogdana Sucheckiego, oraz prof. Jadwigi Sucheckiej. W tym miejscu autorka chce podziękować za wsparcie i cenne uwagi krytyczne, które wpłynęły na poprawę jakości publikacji. Specjalne podziękowania składa na ręce prof. Pawła Miłobędzkiego, którego uwagi były niezmiernie ważne dla ostatecznego kształtu pracy.

ROZDZIAŁ 1

DANE PANELOWE W MODELOWANIU EKONOMETRYCZNYM – KLASYFIKACJA, ROZWÓJ HISTORYCZNY METODOLOGII

Wykorzystanie danych czasowo-przekrojowych w modelowaniu ekonometrycznym zyskuje obecnie coraz większą popularność. Dla zapewnienia dobrych własności estymatorów konieczne jest jednak wówczas stosowanie specyficznych metod estymacji, niejednokrotnie bardziej skomplikowanych niż metody używane do estymacji modeli na podstawie danych „jednowymiarowych" (szeregów czasowych lub danych przekrojowych). Mimo wynikających stąd problemów natury metodologicznej i technicznej (konieczność wykorzystania wysoce specjalistycznego oprogramowania) liczba analiz prowadzonych na podstawie danych czasowo-przekrojowych rośnie. Fakt ten wydaje się świadczyć o tym, że wnioski uzyskiwane na podstawie różnego rodzaju danych, łączących wymiar czasowy i przekrojowy, są unikalne i niemożliwe do uzyskania na podstawie danych „jednowymiarowych".

W niniejszym rozdziale przedstawiona jest klasyfikacja danych łączących wymiar czasowy i przekrojowy, przykłady ogólnie dostępnych źródeł takich danych i oprogramowanie służące do estymacji modeli ekonometrycznych na ich podstawie. Przedstawione są również korzyści wynikające z zastosowania danych czasowo-przekrojowych w praktyce ekonometrycznej. Ich znaczenie dla ekonomicznych analiz ilościowych jest na tyle duże, że przeważyło trudności wynikające z natury takich danych i przyczyniło się do rozwoju odpowiednich metod estymacji. Omówiony w podrozdz. 1.4 rozwój historyczny metod estymacji datuje się od końca XIX w., kiedy to powstały pionierskie prace rozważające te zagadnienia. W podrozdz. 1.5 przedstawiono najbardziej zaawansowane, współczesne konstrukcje i metody pochodzące z pierwszej dekady XXI w. Szczególny nacisk jest tu położony na zagadnienia dotyczące dynamicznych modeli panelowych, co ma uzasadnienie w tematyce ogólnej niniejszej pracy. Dlatego dwa ostatnie podrozdziały dotyczą sposobów uwzględnienia dynamiki modelowanych zjawisk oraz roli warunków początkowych w procesie estymacji i wnioskowania.

1.1. Klasyfikacja danych i modeli czasowo-przekrojowych i panelowych

Dane czasowo-przekrojowe powstają z połączenia w jeden zbiór szeregów czasowych obserwacji, pochodzących z różnych obiektów. Obiektami tymi mogą być zarówno jednostki przestrzeni geograficznej, jak i firmy, gospodarstwa domowe lub inne obiekty dowolnego typu.

Określenia „obiekt", „przekrój", „jednostka" rozumieć należy bardzo ogólnie. Obiektami, czyli jednostkami przekroju, mogą być np. państwa będące członkami tej samej organizacji gospodarczej, regiony administracyjno-gospodarcze w obrębie jednego lub kilku państw, powiązane ze sobą makroregiony gospodarcze, ale także przedsiębiorstwa lub działy przemysłu. Obiekty takie funkcjonują w ścisłym powiązaniu ze sobą, z drugiej jednak strony pozostają one odrębnymi jednostkami. Naturalne więc wydaje się, że na zachowanie każdego obiektu wpływają zarówno czynniki indywidualne, wynikające wyłącznie z jego specyfiki, działające na funkcjonowanie tylko tego jednego obiektu, jak i inne czynniki, determinujące działalność wszystkich obiektów w jednakowy sposób i w jednakowym stopniu. Obecność tych drugich, „wspólnych" czynników, uzasadnia łączenie obserwacji pochodzących z różnych obiektów w jeden szereg i prowadzenie analiz ekonometrycznych na podstawie połączonych danych (ang. *pooled data*).

Dane panelowe (ang. *panel data*) stanowią szczególną podgrupę danych, łączących wymiar czasowy i przekrojowy. Od typowych danych czasowo-przekrojowych (ang. *time-series – cross-section data*) różnią się tym, że liczba obserwowanych obiektów N jest bardzo duża w stosunku do liczby punktów w czasie T. Jeśli wszystkie obiekty są obserwowane przez cały okres próby, to takie dane nazywa się panelem zbilansowanym (ang. *balanced panel*, *complete panel*), a jeśli dla pewnych obiektów szereg czasowy jest dłuższy lub krótszy niż dla pozostałych, to panel jest niezbilansowany (ang. *unbalnced panel*, *incomplete panel*). W anglojęzycznej literaturze przedmiotu spotyka się niekiedy rozróżnienie między *panel data* – danymi, dla których liczba punktów w czasie jest bardzo mała, i *longitudinal data* – danymi, zawierającymi nieco więcej obserwacji w czasie. Oba wymienione rodzaje danych łączy to, że liczba obiektów jest większa, z reguły znacznie, niż liczba okresów. Wydaje się jednak, że w najnowszej literaturze, szczególnie w kontekście wykorzystania danych do estymacji modeli dynamicznych, termin *longitudinal data* przestaje być używany. Dane, łączące wymiar przekrojowy i czasowy, w których liczba obiektów jest większa niż liczba obserwacji w czasie, nazywa się danymi panelowymi (ang. *panel data*), przy czym wprowadza się rozróżnienie między *micro panels* i *macro panels*. Szerzej na temat różnic między tymi dwoma rodzajami danych i wynikających stąd konsekwencji napisano w dalszej części niniejszego rozdziału. Dane, łączące wymiar przekrojowy i czasowy, mogą też mieć inną

strukturę, tzn. liczba obiektów jest mniejsza niż liczba obserwacji w czasie. Dane takie nazywa się danymi czasowo-przekrojowymi. W literaturze anglojęzycznej określa się je jako (*typical*) *time-series – cross-section data*.

Typowym przykładem bazy indywidualnych danych panelowych jest często wspominany w literaturze zbiór informacji Uniwersytetu Michigan nazywany *Panel Study of Income Dynamics*. Baza ta zawiera ok. 5000 zmiennych, a dane dotyczą ok. 6000 rodzin i 15 000 gospodarstw jednoosobowych. Obserwacje są zbierane corocznie od roku 1968 do chwili obecnej. Źródłami zagregowanych danych panelowych mogą też być np.: Eurostat, GUS, roczniki statystyczne, a źródłami danych indywidualnych – dane dziekanatu, baza danych firmy ubezpieczeniowej itp. Więcej informacji na temat dostępnych baz danych panelowych zawarto w podrozdz. 1.2.

Modele ekonometryczne estymowane w oparciu o dane panelowe, nazywane modelami panelowymi (ang. *panel data models*), są zwykle zorientowane na analizę przekrojową, a zadaniem ich jest wyodrębnienie różnic pomiędzy obiektami, które są nierozerwalnie związane z czynnikami specyficznymi dla poszczególnych obiektów. Heterogeniczność obiektów jest tu więc integralną częścią lub nawet centralnym zagadnieniem analizy, zaś efekty czasowe są traktowane jako skokowe zmiany stanu. Są one zazwyczaj modelowane jako efekty specyficzne dla okresu, w którym wystąpiły i nie są przenoszone na następne okresy.

W modelach ekonometrycznych, szacowanych na podstawie danych panelowych zakłada się z reguły, że na kształtowanie się zmiennej objaśnianej wpływają, oprócz zmiennych objaśniających, niemierzalne, stałe w czasie i specyficzne dla danego obiektu czynniki, zwane efektami grupowymi i/lub niemierzalne, stałe względem obiektów specyficzne dla danego okresu czynniki, zwane efektami czasowymi. Uwzględnienie w modelach panelowych efektów grupowych oraz czasowych spowodowało konieczność opracowania specyficznych metod estymacji, odrębnych zarówno od metod stosowanych do estymacji na podstawie szeregów czasowych lub danych przekrojowych, jak i od metod stosowanych w analizach prowadzonych na podstawie danych czasowo-przekrojowych[1].

Szczególnymi przypadkami modeli szacowanych na podstawie danych panelowych są, podobnie jak w przypadku szeregów czasowych lub danych przekrojowych, modele zmiennych jakościowych (ang. *discrete-response models*), w których zakres zmienności zmiennej objaśnianej jest ograniczony – modele ucięte (ang. *truncated models*) oraz modele cenzurowane (ang. *censored models*). Metody estymacji takich modeli są zbliżone do metod stosowanych dla odpowiednich modeli danych przekrojowych lub modeli szeregów czasowych.

---

[1] Do estymacji modeli ekonometrycznych na podstawie danych czasowo-przekrojowych, w których nie zakłada się występowania efektów grupowych, stosuje się zazwyczaj uogólnioną metodę najmniejszych kwadratów z uwzględnieniem zjawiska heteroskedastyczności grupowej i korelacji przekrojowej.

Osobną kategorię, specyficzną dla modeli panelowych tworzą modele szacowane na podstawie danych niekompletnych (ang. *incomplete panels*). Wyróżnia się tutaj dwie grupy: modele z rozkładem opóźnień szacowane na podstawie krótkich paneli (ang. *short panels*) oraz modele szacowane na podstawie danych, w których brakuje informacji o pewnych obiektach w określonych okresach. Problemem, który może się pojawić w przypadku estymacji modeli z rozkładem opóźnień na podstawie krótkich paneli, jest to, że maksymalny rząd opóźnienia zmiennej przekracza wymiar czasowy próby. Estymację takich modeli umożliwia zastosowanie estymatora minimalnej odległości (ang. *minimum distance estimator*). Modele szacowane na podstawie danych niekompletnych podzielić można, w zależności od przyczyn ich niekompletności, na dwie grupy: 1) brak obiektów wynika z przyczyn losowych lub z rotacyjności próby (ang. *rotating panels*), 2) obiekty w próbie zmieniają się w czasie, gdyż powtarzanie obserwacji dla tych samych obiektów jest niemożliwe (ang. *pseudopanels*, *repeteated cross-sectional data*). Metody estymacji modeli bazujących na niekompletnych danych są bezpośrednią adaptacją metod stosowanych przypadku kompletności danych, o ile rotacja jest kontrolowana, model jest statyczny, a składniki losowe z różnych obserwacji nie zależą od siebie. Metody te bazują na UMNK i MNW. W przypadku modeli dynamicznych powstaje dodatkowy problem związany z tym, że różne założenia dotyczące warunków początkowych implikują różne schematy dołączania nowych obiektów do próby. Jeśli niekompletność próby wynika z przyczyn losowych, zaleca się ograniczenie próby do podzbioru, dla którego obserwacje są kompletne, lub zastosowanie podobnych procedur, jak w przypadku paneli rotacyjnych[2].

Modele szacowane na podstawie danych panelowych lub czasowo-przekrojowych stosowane są w praktyce ekonometrycznej coraz powszechniej. Wzrost częstotliwości wykorzystania takich danych wiąże się m. in. z dostępem do odpowiedniego oprogramowania komputerowego. Pakiety komputerowe posiadające odpowiednie moduły to np.: STATA, LIMDEP, EViews, Gretl, Soritec (w tym ostatnim pakiecie dostępne są tylko modele FE). Spośród wymienionych, jedynie program Gretl jest programem należącym do oprogramowań Powszechnej Licencji Publicznej (GNU), czyli o powszechnym, darmowym dostępie dla wszystkich użytkowników[3]. Wydaje się, że programami posiadającymi najbardziej rozbudowane moduły dla modeli panelowych są: STATA (szczególnie godna polecenia dla modeli dynamicznych) i LIMDEP (szczególnie dla modeli z ograniczonym zakresem zmienności zmiennej objaśnianej).

---

[2] Więcej informacji na temat założeń i metod estymacji panelowych modeli zmiennych jakościowych, modeli z ograniczonym zakresem zmienności zmiennej objaśnianej i modeli bazujących na danych niekompletnych znaleźć można np. w: H s i a o [2003], B a l t a g i [2008].

[3] Pełna wersja instalacyjna programu dostępna jest na stronach internetowych: http://gretl.sourceforge.net i http://www.kufel.torun.pl (wersja polska).

1.2. Dane panelowe w modelowaniu makroekonomicznym (makropanele) i mikroekonomicznym (mikropanele). Źródła danych statystycznych

W podrozdz. 1.1 wspomniano, że w związku z dynamicznie rosnącą liczbą praktycznych zastosowań modeli panelowych w różnego rodzaju analizach ekonomicznych i wynikającym z tej różnorodności zróżnicowaniem próby pod względem wymiaru czasowego i przekrojowego, w najnowszych opracowaniach (np. B a l t a g i [2008], s. 1) pojawił się podział danych panelowych na makropanele i mikropanele. Mikropanele to dane, charakteryzujące się bardzo dużą liczbą obserwacji przekrojowych, sięgającą setek lub nawet tysięcy obiektów; liczba obserwacji w czasie jest natomiast bardzo niewielka – rzadko przekracza dziesięć, zazwyczaj jest to kilka obserwacji. Typowym przykładem mikropanelu są dane pochodzące z gospodarstw domowych lub inne dane indywidualne. Makropanele charakteryzują się natomiast bardziej ograniczoną liczbą obserwacji przekrojowych (N waha się od kilkunastu do kilkudziesięciu) i dłuższym horyzontem czasowym próby, sięgającym nawet $T = 30$ okresów. Są to zazwyczaj dane, wykorzystywane w badaniach makroekonomicznych, dotyczące krajów (np. Unii Europejskiej, OECD) lub regionów, dla których dostępne są stosunkowo długie szeregi czasowe dla wielu zmiennych makroekonomicznych. Modele panelowe estymowane na podstawie makro- i mikropaneli wymagają stosowania różnych metod ekonometrycznych. Potrzeba taka wynika z różnych wymiarów tych danych i związanych z nimi pożądanych własności asymptotycznych estymatorów. W przypadku mikropaneli szczególnie ważne są własności asymptotyczne dla $N \to \infty$ i skończonego T, a w przypadku makropaneli – dla $N \to \infty$ i $T \to \infty$, lub własności estymatorów w małych próbach. Ponadto w przypadku makropaneli o szczególnie długim wymiarze czasowym mogą pojawiać się problemy związane z niestacjonarnością, niewystępujące w mikropanelach, w których wymiar T jest niewielki. Problemem, który jest typowy dla makropaneli, jest natomiast zależność między obiektami (ang. *cross-unit dependence*), powodująca zjawisko korelacji przekrojowej. Nie występuje on zazwyczaj w modelach szacowanych na podstawie mikropaneli, gdyż obiekty są elementami próby losowej, a zatem mało prawdopodobna jest ich wzajemna korelacja.

Przegląd metod estymacji pod kątem ich przydatności do estymacji jednorównaniowych modeli dynamicznych na podstawie mikropaneli zawarty jest w artykule S. B o n d a [2002]. Metody, które bazują na GMM, i które są omówione w rozdz. 3 niniejszej publikacji, według autora są szczególnie użyteczne w przypadku modeli zawierających endogeniczne lub z góry ustalone zmienne objaśniające. Endogeniczność zmiennych objaśniających jest zjawiskiem naturalnym: założenie o ścisłej ich egzogeniczności oznaczałoby, że przeszłe wartości zmiennych nie wpływają na wartości bieżące, co byłoby trudne do przyjęcia w modelowaniu konsumpcji, dochodu, inwestycji *etc.*,

a więc zjawisk modelowanych często na podstawie mikropaneli. Autor zwraca też uwagę na problemy własności estymatorów w przypadku małych prób (jeśli dodatkowo wymiar przekrojowy próby jest niewielki) oraz warunków początkowych w kontekście mikropaneli. Wpływ obserwacji początkowej na każdą kolejną obserwację jest tu szczególnie istotny ze względu na krótkość szeregu czasowego. Podkreślając przydatność metod opartych na GMM do estymacji modeli dynamicznych na podstawie mikropaneli, Bond [2002] zwraca uwagę na możliwe obciążenie tych estymatorów w przypadku, gdy instrumenty są słabe i gdy stosuje się metodę GMM pierwszych różnic Arellano i Bonda [1991] dla szeregów o wysokim stopniu trwałości (ang. *highly persistent time--series*). Ilustracją jego rozważań są dwa przykłady praktycznego wykorzystania mikropaneli do analiz dynamicznych. Własności różnych estymatorów są analizowane dla autoregresyjnego modelu stopy inwestycji dla firm brytyjskich oraz modelu produkcji dla firm z USA. Dokładniejsze omówienie obu modeli jest zawarte w rozdz. 4.

Własności estymatorów bazujących na GMM w kontekście modeli dynamicznych szacowanych na podstawie makropaneli są rozważane przez S. Bonda, A. Hoeffler i J. Temple'a [2001]. Autorzy koncentrują się na modelach wzrostu i konwergencji, które stanowią szczególnie liczną grupę wśród modeli estymowanych na podstawie danych makropanelowych. W przypadku tego typu danych, podobnie jak w przypadku mikropaneli, często pojawiającymi się problemami są endogeniczność i błędy pomiaru zmiennych, jak również problem pominiętych zmiennych. Źródłem problemów w przypadku wykorzystania makropaneli może też być mała próba oraz wysoki stopień trwałości szeregów. Autorzy wskazują systemową GMM Blundella i Bonda [1998] jako najlepszą metodę estymacji modeli wzrostu i konwergencji. Rozważania są zilustrowane przykładem zastosowania estymatora GMM pierwszych różnic i systemowego estymatora GMM z alternatywnymi zbiorami instrumentów do szacowania modelu wzrostu PKB i konwergencji.

Dostęp do baz danych, zarówno makro-, jak i mikropanelowych jest w obecnej dobie stosunkowo łatwy, zwłaszcza do baz międzynarodowych oraz mikropanelowych baz amerykańskich i zachodnioeuropejskich. Przykłady takich baz wraz z adresami witryn internetowych są podane i opisane w: Baltagi [2008] (podrozdz. 1.1, s. 1–6); i Hsiao [2003] (podrozdz. 1.1, s. 1–3). Pogląd na temat ogromnych rozmiarów baz danych mikropanelowych dać może przykład danych z gospodarstw domowych, zgromadzonych na Uniwersytecie Michigan (USA), znany jako *Panel Study of Income Dynamics* (PSID). Dane zaczęto zbierać w 1968 r. i pochodziły one od 4800 rodzin. Do 2003 r. w PSID zgromadzono dane o 65 000 jednostkach, a długość szeregów czasowych dochodzi aż do 36 lat. Bazę PSID znaleźć można na stronie http://psidonline.isr.umich.edu. Inną ważną bazą danych zgromadzoną w USA jest *National Longitudinal Surveys* (NSL). Gromadzenie danych rozpoczęto

w połowie lat 60. XX w. Baza składa się z pięciu rozłącznych baz danych wzdłużnych, obejmujących odrębne segmenty siły roboczej: mężczyzn w wieku 45–59 lat (w roku 1966), mężczyzn w wieku 14–24 (w 1966), kobiet w wieku 30–44 (w 1967), kobiet w wieku 14–24 (w 1968) i młodzieży obu płci w wieku 14–21 lat (w 1979). W roku 1986 bazę NSL rozszerzono tak, że objęła dzieci kobiet uczestniczących w NLS w roku 1979. Baza danych sięga tysięcy zmiennych opisujących podażową stronę rynku pracy. Sumaryczne statystyki dla bazy NSL zawarte są w pracy B. H. B a l t a g i'ego [2008] (tab. 1.1, s. 2), a tabele i dane na stronie internetowej Bureau of Labor Statistics: http://www.bls.gov/nls/. Wiele krajów europejskich przeprowadza coroczne lub częstsze badania ankietowe. Wymienić tu można: *British Household Panel Survey* (BHPS): http://www.iser.essex.ac.uk/survey/bhps, *German Socio-Economic Panel*: http://www.diw.de/en/soep, *The Netherlands Socio-Economic Panel* (SEP): http://www.nwo.nl/nwohome.nsf/pages/NWOP_57SDAK. W 1994 r. Eurostat rozpoczął koordynację i łączenie baz danych dotyczących wskaźników społeczno-ekonomicznych gromadzonych w poszczególnych krajach członkowskich, inicjując pracę *National Data Collection Units* (NDU). Dane indywidualne oraz pochodzące z gospodarstw domowych z terenu Unii Europejskiej zgromadzone są w bazie *European Community Household Panel* (ECHP). Zawiera ona zmienne demograficzne, zmienne opisujące siłę roboczą, dochody, dane o zdrowiu i wykształceniu, warunkach mieszkaniowych *etc*. i dostępna jest na stronie internetowej Eurostatu: http://epp.eurostat.ec.europa.eu/portal/page/portal/microdata/echp.

Międzynarodowe dane makropanelowe są dostępne m. in. na stronach Penn World Table (http://pwt.econ.upenn.edu), Banku Światowego (*World Development Indicators*: www.worldbank.org/data), Międzynarodowego Funduszu Walutowego (www.imf.org), OECD (www.oecd.org). Istnieją również bazy danych regionalnych, udostępniane przez odpowiednie urzędy statystyczne. Na przykład dane dla wszystkich regionów NUTS-2, krajów Unii Europejskiej są dostępne na stronie http://epp.eurostat.ec.europa.eu, a roczne i (w bardziej ograniczonym zakresie) kwartalne dane dla Polski, w układzie makroregionów (NUTS-1), województw (NUTS-2) i (wybrane dane) powiatów (NUTS-3) są dostępne na stronie GUS: www.stat.gov.pl/bdr.

1.3. Zalety wykorzystania danych panelowych w praktyce ekonometrycznej

Wykorzystanie danych panelowych lub czasowo-przekrojowych w procesie konstrukcji i estymacji modeli ekonometrycznych dostarcza korzyści w czterech głównych aspektach:

– powoduje wzrost liczby stopni swobody i redukuje problem współliniowości danych;

– ułatwia identyfikowalność modeli ekonomicznych i dokonanie wyboru między konkurencyjnymi hipotezami ekonomicznymi;

– umożliwia eliminację lub redukcję obciążenia estymatorów;

– dostarcza podstaw w skali mikro dla analiz prowadzonych na podstawie danych zagregowanych.

Wymienione sytuacje, w których wykorzystanie danych panelowych przynosi największe korzyści, zostaną omówione nieco szerzej poniżej. Więcej na temat zalet wykorzystania danych panelowych w praktyce ekonometrycznej, a także trudności w estymacji wynikających z dwuwymiarowości danych znaleźć można m. in. w: H s i a o [2003] (podrozdz. 11.2, s. 311–316); B a l t a g i [2008] (podrozdz. 1.2, s. 6–10).

Mała liczba stopni swobody i współliniowość szeregów czasowych często powodują niemożność zbadania wpływu, wywieranego przez poszczególne zmienne objaśniające na zmienną objaśnianą. Informacje pochodzące z próby czasowej lub przekrojowej nie są wystarczające, by sprostać wymaganiom modelu. W takiej sytuacji zachodzi konieczność wydłużenia horyzontu próby lub zredukowania wymagań modelu poprzez narzucenie określonych restrykcji. Innym rozwiązaniem jest połączenie obserwacji w czasie i względem obiektów w jeden zbiór, co oprócz zwiększenia liczby stopni swobody skutkuje wzbogaceniem informacji o indywidualnych zachowaniach poszczególnych obiektów, umożliwiając ostatecznie bardziej kompletną interpretację modelu.

W naukach społeczno-ekonomicznych często formułowane są alternatywne, sprzeczne ze sobą hipotezy badawcze. Jednym z przykładów może być wpływ przynależności do związków zawodowych i umów zbiorowych na płace pracowników. Sprzeczność dotyczy interpretacji oceny parametru przy zmiennej zero-jedynkowej, wprowadzanej do modelu w charakterze zmiennej wskaźnikowej (1 – przynależność pracownika do związku zawodowego). Zwolennicy hipotezy o istotności wpływu umów zbiorowych na płace interpretują tę ocenę jako efekt przynależności do związków, podczas gdy jej przeciwnicy – jako przybliżoną miarę kwalifikacji pracownika, argumentując, że wynegocjowane przez związki podwyżki płac są w długim okresie kompensowane przez pracodawców poprzez zatrudnianie wyżej wykwalifikowanych pracowników. Wykorzystanie danych panelowych, przy założeniu stałych (w krótkim okresie) kwalifikacji danego pracownika, umożliwia obserwowanie różnic w płacach osób zapisujących się do związków i rezygnujących z przynależności. Innymi przykładami mogą być hipotezy dotyczące właściwej polityki krótkookresowej prowadzącej do zmniejszenia bezrobocia, wpływu wykształcenia, płci i stanu cywilnego na płace, związków przyczynowych między zmiennymi. Możliwości weryfikacji wymienionych hipotez na podstawie modeli panelowych przedstawione zostały m. in. w pracach: F r e e m a n, M e d o f f [1981]; C h a m b e r-

l a i n, G r i l l i c h e s [1975]; F u j i k i, H s i a o, S h e n [2002]; J a n g, T h o r n - b e r r y [1998]; A l l i s o n [2000].

Jednym z najważniejszych problemów podczas konstrukcji modelu ekonometrycznego jest problem doboru zmiennych objaśniających i natury zależności miedzy tymi zmiennymi a czynnikami nieuwzględnionymi w modelu, ale wpływającymi na zmienną objaśnianą poprzez składnik losowy (ang. *specification problem*). Przykładem takich zmiennych może być dochód i styl życia. Do modelu wprowadza się zazwyczaj łatwiej mierzalną pierwszą z nich, przy czym zmienna ta reprezentuje również styl życia. Zmiany dochodu, które nie wiążą się ze zmianą stylu życia są wtedy błędnie interpretowane jako efekty zmian obu wielkości. Dzięki konstrukcji modeli panelowych zmienna „dochód" odzwierciedlałaby rzeczywisty wpływ dochodu na zmienną objaśnianą, a styl życia byłby potraktowany jako część składowa efektu grupowego. Jeśli pominięte zmienne są skorelowane ze zmiennymi objaśniającymi, a założenia modelu nie dopuszczają takiej korelacji, to estymatory parametrów modelu są obciążone. Obciążenie to można zminimalizować lub zredukować po uprzednim określeniu typu korelacji między zmiennymi objaśniającymi a składnikiem losowym. C. H s i a o [2003] (rozdz. 11, s. 313) wyróżnia cztery typy takiej korelacji. Korelacja pierwszego typu ma miejsce w przypadku pominięcia zmiennych skorelowanych ze zmiennymi objaśniającymi, niezależnie od tego, czy pominięcie wynika z błędu specyfikacji, czy z braku danych. Drugi typ korelacji jest związany z dynamiczną strukturą modelu i spowodowany korelacją opóźnionej zmiennej objaśnianej ze składnikiem losowym. Trzeci typ spowodowany jest współzależnością modelu i dotyczy sytuacji, gdy łącznie współzależne zmienne objaśniane są skorelowane ze składnikami losowymi. Korelacja czwartego typu jest spowodowana błędami pomiaru zmiennych objaśniających.

Określenie przyczyny korelacji między zmiennymi objaśniającymi a składnikiem losowym dostarcza informacji niezbędnych przy wyborze właściwego estymatora zgodnego. Zastosowanie danych panelowych daje możliwość eliminacji zmiennych będących przyczyną obciążenia poprzez obliczenie odchyleń od średnich grupowych lub czasowych (korelacja pierwszego typu) lub pierwszych różnic zmiennych (korelacja drugiego typu) (por. H s i a o [2003], rozdz. 3, 4, 7, 8). W przypadku modelu współzależnego (korelacja trzeciego typu) możliwe jest zastosowanie zmiennych instrumentalnych, przy czym czasowo-przekrojowy charakter danych umożliwia wykorzystanie jako instrumentów również zmiennych uwzględnionych w modelu, oczyszczonych ze stałych efektów, powodujących korelację (por. H s i a o [2003], podrozdz. 5.3, 5.4). Błędy pomiaru zmiennych objaśniających mogą spowodować korelację tych zmiennych ze składnikiem losowym (korelacja czwartego typu). Różnego rodzaju transformacje zmiennych (w tym obliczanie pierwszych różnic) powodują dające się przewidzieć zmiany ocen parametrów strukturalnych. Na tej podstawie określić można znaczenie błędów pomiaru i wyznaczyć zgodne oceny parametrów (por. H s i a o [2003], podrozdz. 10.5). J. M. W o o l d r i d g e [2002]

(s. 421) zauważa: „w wielu przypadkach, jedynym powodem, dla którego korzysta się z danych panelowych jest możliwość dopuszczenia, aby nieobserwowalne efekty θ_i były skorelowane ze zmiennymi objaśniającymi". J. L e e [2002] (s. 16) dodaje: „możliwość usunięcia z modelu niezmiennego w czasie, nieobserwowalnego efektu θ_i jest największą zaletą wykorzystania danych panelowych".

Wykorzystanie danych zagregowanych wymaga przyjęcia założenia o istnieniu podmiotu reprezentatywnego. Tymczasem, jeśli w skali mikro jednostki są heterogeniczne, to własności szeregów czasowych danych zagregowanych mogą się znacznie różnić od własności szeregów danych zdezagregowanych, a co za tym idzie – oceny w skali makro oparte o dane zagregowane mogą być mylące. Ponadto prognozy w skali makro sporządzone na podstawie danych zagregowanych mogą być mniej trafne niż sporządzone na podstawie danych zdezagregowanych (por. np. F u j i k i, H s i a o, S h e n [2002]). Wnioskowanie o całej populacji z uwzględnieniem heterogeniczności obiektów w skali mikro jest możliwe dzięki zastosowaniu modeli panelowych ze zmiennymi parametrami.

Przedstawione powyżej korzyści, płynące z zastosowania modeli panelowych, są niewątpliwie powodem rosnącej liczby ich aplikacji. Z drugiej strony, specyficzna natura danych, łączących wymiar czasowy i przekrojowy, może być przyczyną poważnych komplikacji natury metodologicznej. Dlatego istotnym pytaniem, na jakie należy odpowiedzieć decydując się na korzystanie z modelu panelowego, jest: czy model taki dostarczy dodatkowych informacji. Jeśli cechy charakteryzujące badaną zbiorowość, wpływające na jej mobilność, takie jak np. ceny paliwa lub struktura przestrzenna, pozostają względnie stałe, to model estymowany na podstawie indywidualnych danych zbieranych w czasie nie będzie miał przewagi informacyjnej nad modelem estymowanym na podstawie danych przekrojowych. Jeżeli natomiast pewne cechy zmieniają się znacząco, np. wprowadzane są pewne nowe regulacje, to model panelowy może dostarczyć bardzo istotnych informacji o zmianach.

1.4. Początki rozwoju metodologii estymacji modeli panelowych

Łączenie szeregów czasowych dotyczących różnych obiektów i konstruowanie modeli szacowanych na ich podstawie ma długą tradycję. Informacje o pionierskich pracach, w których rozważano te zagadnienia, datujących się z pierwszego ćwierćwiecza XX w. lub nawet końca wieku XIX, znaleźć można w książce M. N e r l o v e'a [2002] (rozdz. 1, s. 1–114), oraz referacie M. Nerlove'a [2000]. W monografii z dziedziny astronomii, opublikowanej w 1861 r. G. B. Airy[4] zauważył występowanie pewnych efektów zależnych od dnia,

[4] Teksty źródłowe prac Airy'ego i pozostałych, cytowanych w tym akapicie są autorce niedostępne, a informacje o nich przytoczone zostały za: N e r l o v e [2000].

w którym dokonano obserwacji. Efekty te, nazwane przez niego *constant errors*, są we współczesnej terminologii losowymi efektami grupowymi. R. A. Fisher w 1925 r. stworzył koncepcję modeli, które zgodnie ze współczesną terminologią nazywane byłyby modelami FE (*fixed effects*) i RE (*random effects*). Kolejne prace, dotyczące konsekwencji traktowania efektów grupowych jako składowej składnika losowego (w modelach RE) lub specyficznego dla obiektów wyrazu wolnego (w modelach FE), autorstwa H. E. Danielsa i C. Eisenharta opublikowane zostały w latach 1939 i 1947.

W drugiej połowie XX w. nastąpił szybki rozwój metodologii modelowania na podstawie danych łączących wymiar czasowy i przekrojowy. Pierwsze zaawansowane metodologicznie prace, w których zwrócono uwagę na istotność wpływu czynników specyficznych dla poszczególnych obiektów (lub okresów) na zmienną objaśniającą, pochodzą z przełomu lat 50. i 60.

C. Hildreth [1950][5] rozważał model wielorównaniowy, szacowany na podstawie próby czasowo-przekrojowej, w którym wyróżniono stałe w czasie efekty grupowe i stałe względem obiektów efekty czasowe. Dopuszczał on możliwość, aby część zmiennych, zarówno endogenicznych, jak i z góry ustalonych, była stała w czasie lub stała względem obiektów. W dalszej części Hildreth zauważył, że efekty grupowe i czasowe można traktować jako losowe (będące częścią składnika losowego) lub nielosowe (będące parametrami, podlegającymi estymacji), ale stwierdził, że decyzja o sposobie ich potraktowania jest trudna. Hildreth zapisał też funkcję wiarygodności dla modelu jednorównaniowego, w którym założył, że efekty są składowymi składnika losowego, składnik losowy $u_{it} = \alpha_i + \lambda_t + \varepsilon_{it}$ ma rozkład normalny, jest homoskedastyczny, a jego składowe nie wykazują autokorelacji. Stwierdził jednak, że metoda największej wiarygodności jest niemożliwa do zastosowania, ponieważ układ równań normalnych otrzymanych na podstawie funkcji wiarygodności jest „wysoce nieliniowy" (ang. *highly nonlinear*) względem parametrów.

Wpływ efektów grupowych i czasowych na zmienną objaśnianą obserwowali w latach późniejszych m. in.: E. Kuch [1959] w analizie inwestycji, I. Hoch [1962], Y. Mundlak [1961] i M. Nerlove [1965] w badaniach dotyczących funkcji produkcji. Wydaje się, że właśnie duża użyteczność danych panelowych w modelowaniu produkcji zainicjowała rozwój metod estymacji modeli ekonometrycznych na podstawie danych panelowych. Klasyczny model szacowany na podstawie danych panelowych wykorzystywany we współczesnej ekonometrii od drugiej połowy XX w. ma w najbardziej ogólnym przypadku postać:

---

[5] Tekst źródłowy: Hildreth C. [1950], *Combining cross-section data and time-series*, „Cowles Comission Discussion Paper", No. 347, 15 May jest autorce niedostępny. Prezentowane tu informacje przytoczono za: Nelrove [2000].

$$y_{it} = \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + u_{it} = \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + (\alpha_i + \lambda_t + \varepsilon_{it}) \;\; i = 1, ..., N, \; t = 1, ..., T \qquad (1.1)$$

gdzie: $\mathbf{x}_{it}^{\mathrm{T}} = [x_{1it}, ..., x_{Kit}]$ – wektor obserwacji na zmiennych objaśniających; $\boldsymbol{\beta} = [\beta_1, ..., \beta_K]^{\mathrm{T}}$ – wektor parametrów strukturalnych, $u_{it} = \alpha_i + \lambda_t + \varepsilon_{it}$ – składnik losowy, w którym wyróżnia się stałe w czasie składowe α_i, właściwe dla i-tego obiektu, zwane efektami grupowymi, stałe względem obiektów składowe λ_t właściwe dla okresu t (efekty czasowe) oraz „właściwy" składnik losowy ε_{it}. Uwzględnianie w modelu obu rodzajów efektów (grupowych i czasowych) jest opcjonalne i zależy od celu badania. Dość często spotyka się modele, w których występują efekty jednego typu, z reguły efekty grupowe. Szczegółowa specyfikacja założeń oraz omówienie konstrukcji modelowych i metod estymacji wynikających z tych założeń są treścią rozdz. 2, a w szczególności podrozdz. 2.1–2.4.

I. Hoch [1962] szacował parametry modelu funkcji produkcji Cobba–Douglasa na podstawie danych rocznych z lat 1946–1951, z 63 farm w stanie Minnesota. Zastosowanie do modelowania procesu produkcyjnego funkcji Cobba–Douglasa, postaci:

$$y_{it} = \alpha + \beta_1 x_{1it} + ... + \beta_k x_{kit} + u_{it}, \quad i = 1, ..., N, \; t = 1, ..., T \qquad (1.2)$$

gdzie y_{it} jest logarytmem produkcji, x_{kit} są logarytmami nakładów, wymagałoby założenia, że pominięte zmienne są niezależne od zmiennych x_{kit} i mają identyczne, niezależne rozkłady. Przyjęcie postaci (1.2) uniemożliwiałoby ponadto oszacowanie różnic między obiektami, takich jak efekty różnic w sposobie zarządzania, lub różnic o charakterze technicznym. Nie byłoby też możliwe oszacowanie wpływu czynników zmieniających się w czasie – w modelu Hocha mogą to być np. warunki atmosferyczne. I. Hoch [1962] zauważył, że ponieważ dane mają wymiar przekrojowy i czasowy, to możliwe jest uwzględnienie wspomnianych wyżej czynników, jeśli składnik losowy modelu (1.2) zapisany zostanie w postaci:

$$u_{it} = \alpha M_i + \lambda P_t + \varepsilon_{it} \qquad (1.3)$$

gdzie M_i są czynnikami specyficznymi dla farm, a P_t – specyficznymi dla czasu. Ponieważ jednak M_i i P_t są nieobserwowalne, trzeba wprowadzić podstawienie: $\alpha_i = \alpha M_i$ i $\lambda_t = \lambda P_t$. Podstawienie takie prowadzi do modelu panelowego postaci (1.1). W modelu Hocha α_i i λ_t są następnie traktowane jako parametry podlegające estymacji. Zaproponowana specyfikacja jest więc, zgodnie z wprowadzonym później nazewnictwem, modelem FE. Model (1.2)–(1.3) oszacowany został klasyczną MNK. Autor testował dwa różne zbiory zmiennych objaśniających: 1) nakłady pracy, wartość nieruchomości, wartość maszyn, nakłady na nawozy,

oraz 2) nakłady pracy, nakłady na nawozy, wydatki bieżące, wartość środków trwałych. Model estymowany był w trzech wariantach: a) przy założeniu, że $\alpha_i = \lambda_t = 0$, b) $\alpha_i = 0$, $\lambda_t \neq 0$, c) $\alpha_i \neq 0$ i $\lambda_t \neq 0$. Jakość modelu, mierzona wartością współczynnika determinacji, wzrosła z 0,71, w wariancie 1a), do 0,88, w wariancie 2c). Zauważa się znaczny spadek oceny parametru przy nakładach pracy, towarzyszący wprowadzaniu efektów: czasowego (wariant b) i czasowego i grupowego (wariant c). Suma elastyczności w przypadku c), tzn. gdy $\alpha_i \neq 0$ i $\lambda_t \neq 0$, jest równa 0,832 i jest znacząco niższa w porównaniu z przypadkiem a), w którym jest ona równa 0,954. Jeśli oceny α_i są interpretowane jako efekty skali, fakt ten można interpretować jako wzrost efektywności farm, wynikający z ich rozmiaru. Wydaje się, że opisany spadek sumy elastyczności potwierdza ogólną tendencję do uzyskiwania zawyżonych ocen parametrów strukturalnych, jeśli ignoruje się istniejące różnice między obiektami.

Pierwszą pracą, która w bardziej pełny sposób poruszała zarówno aspekty teoretyczne, jak i praktyczne modelowania na podstawie danych panelowych jest praca P. Balestry i M. Nerlove'a [1966], dotycząca w części aplikacyjnej modelowania popytu na gaz. Ponieważ można ją uznać za „kamień milowy" w rozwoju metodologii estymacji modeli panelowych, a ponadto estymowany w niej model jest modelem dynamicznym, poniżej zostanie ona omówiona szczegółowo.

W analizie popytu na gaz P. Balestra i M. Nerlove [1966] dążyli do poprawnego określenia i estymacji parametrów funkcji popytu na dobro, którego konsumpcja jest technologicznie związana z eksploatacją pewnych urządzeń. Funkcja ta ma postać dynamiczną, a estymacji parametrów dokonano na podstawie rocznych danych panelowych z 36 stanów USA. Pierwszym etapem była konstrukcja modelu popytu. Założono, że popyt na gaz zależy m. in. od liczby urządzeń gazowych używanych w danym okresie przez gospodarstwa domowe. Przyjęto założenie o bardzo niskiej krótkoterminowej cenowej elastyczności popytu, gdyż zainstalowanie urządzenia gazowego zniechęca do korzystania z innego rodzaju paliwa. Zatem przyrost popytu na gaz G^* określono jako funkcję ceny gazu p i przyrostu popytu na paliwa ogółem F^*. Przyrost popytu na paliwa ogółem F^* i przyrost popytu na gaz G^* w okresie t wyrażono analogicznymi wzorami:

$$F_t^* = F_t - (1-r)F_{t-1} \tag{1.4.a}$$

$$G_t^* = G_t - (1-r_g)G_{t-1} \tag{1.4.b}$$

gdzie: F_t – popyt na paliwa w okresie t; r – współczynnik zużycia; G_t – popyt na gaz w okresie t; r_g – współczynnik zużycia dla urządzeń gazowych.

Przyjmując liniową postać funkcji popytu, otrzymano równanie odzwierciedlające dynamikę popytu:

$$G_t^* = \beta_0 + \beta_1 p_t + \beta_2 F_t^* = \beta_0 + \beta_1 p_t + \beta_2 \left[F_t - (1-r)F_{t-1}\right] + (1-r_g)G_{t-1} \qquad (1.5)$$

Wartości współczynników r i r_g nie muszą być znane, gdyż można je traktować jako parametry modelu, natomiast konsumpcja paliw ogółem jest funkcją kilku innych zmiennych. Można ją aproksymować jako:

$$F_t = \gamma_0 + \gamma_1 N_t + \gamma_2 Y_t \qquad (1.6)$$

gdzie N_t i Y_t oznaczają odpowiednio liczbę ludności i dochód na osobę. Model popytu przyjmuje wtedy postać:

$$G_t = \alpha_0 + \alpha_1 p_t + \alpha_2 \Delta N_t + \alpha_3 N_{t-1} + \alpha_4 \Delta Y_t + \alpha_5 Y_{t-1} + \alpha_6 G_{t-1} + u_{it} \qquad (1.7)$$

gdzie u_{it} jest składnikiem losowym.

Aby zapewnić jednoznaczną identyfikowalność parametru r, przyjęto następujące warunki ograniczające:

$$\begin{aligned} N_t^* &= N_t - (1-r)N_{t-1} \\ Y_t^* &= Y_t - (1-r)Y_{t-1} \end{aligned} \qquad (1.8)$$

Ostatecznie więc model popytu na gaz przyjmuje postać:

$$G_t = \alpha_0 + \alpha_1 p_t + \alpha_2 N_t^* + \alpha_4 Y_t^* + \alpha_6 G_{t-1} + u_{it} \qquad (1.9)$$

Dostępne dane pochodziły z lat 1950–1962. Okres próby podzielono na dwa, technologicznie różne okresy: lata 1950–1956, będące okresem rozwoju, i lata 1957–1962, w których wszystkie stany miały jednakowo łatwy dostęp do gazu. W artykule P. Balestry i M. Nerlove'a [1966] omówiono tylko wyniki empiryczne uzyskane na podstawie drugiej podpróby.

Na wstępie oszacowano tzw. model *pooled*, a więc zastosowano klasyczną MNK do oszacowania modelu bez efektów grupowych na podstawie wszystkich obserwacji. Zarówno w przypadku równania (1.7), jak i równania (1.9) oceny współczynników przy opóźnionych zmiennych endogenicznych były większe od 1, co oznaczałoby ujemną wartość współczynnika zużycia dla urządzeń gazowych. Natomiast uzyskane oszacowanie współczynnika zużycia dla urządzeń różnych typów było merytorycznie akceptowalne, wynosiło bowiem ok. 11%.

Ponieważ urządzenia gazowe w rozważanym okresie były na ogół nowe, założono, że współczynnik przy zmiennej G_{t-1} jest bliski 1 i zmienna ta została przeniesiona na lewą stronę równań modeli (1.7) i (1.9). W ten sposób zmienną

endogeniczną modelu stał się przyrost popytu na gaz. Rozwiązanie to nie poprawiło jednak wyników estymacji. Ponadto przeciętna elastyczność przyrostu popytu na gaz dla estymowanych modeli wahała się pomiędzy 0,58 a 0,69 i rosła w czasie (elastyczność dla okresu 1957–1962 była wyższa niż dla lat 1950–1956).

B a l e s t r a i N e r l o v e [1966] postawili więc hipotezę, że połączenie danych przekrojowych i czasowych sprawiło, iż uwidoczniły się pewne „inne efekty", specyficzne dla poszczególnych stanów. W celu uwzględnienia tych efektów w modelu, próbowano wprowadzić 36 zmiennych zero-jedynkowych odpowiadających poszczególnym stanom (co oznacza konstrukcję modelu postaci (1.1) z efektami nielosowymi, tzw. modelu FE), z góry poddając w wątpliwość efektywność tej metody w przypadku modelu dynamicznego. Rzeczywiście, oszacowania parametrów przy opóźnionych zmiennych endogenicznych spadły do poziomu, przy którym współczynnik zużycia dla urządzeń gazowych wynosiłby 30%. Wartość tę uznano za nieprawdopodobną.

Następnym krokiem była próba uwzględnienia w modelu czynników specyficznych dla poszczególnych stanów, przy założeniu, że są one losowe. Składnik losowy u_{it} modelu (1.9) rozdzielono na dwie niezależne części: specyficzną dla danego stanu i wspólną średnią, tzn. $u_{it} = \alpha_i + \varepsilon_{it}$, co odpowiada konstrukcji modelu postaci (1.1) z efektami losowymi (modelu RE).

Do estymacji modelu (1.9) jako modelu z efektami losowymi próbowano najpierw zastosować metodę największej wiarygodności (MNW). Model (1.9), z założeniem $u_{it} = \alpha_i + \varepsilon_{it}$, zapisano macierzowo w postaci:

$$\mathbf{y} = \mathbf{Z\alpha} + \mathbf{y}_M \boldsymbol{\gamma} + \mathbf{u} = \mathbf{X\beta} + \mathbf{u} \tag{1.10}$$

gdzie $\mathbf{Z}$ jest macierzą obserwacji na zmiennych egzogenicznych, $\mathbf{y}$ – wektorem obserwacji na zmiennej endogenicznej, $\mathbf{y}_M$ – wektorem obserwacji na opóźnionych wartościach tej zmiennej, $\mathbf{u}$ – wektorem składników losowych.

Pierwszym krokiem do znalezienia estymatora wektora $\boldsymbol{\beta}$ było oszacowanie nieznanej macierzy wariancji-kowariancji składnika losowego $\boldsymbol{\Omega}$. Odpowiednie procedury estymacyjne dla modeli, w których nie występują opóźnione zmienne endogeniczne, zostały opracowane przez A. Z e l l n e r a [1962] i L. G. T e s l e r a [1964]. Jednakże ich zastosowanie nie było możliwe, ponieważ obie te procedury bazują na estymatorach KMNK. Estymatory te w przypadku modeli, w których występują opóźnione zmienne endogeniczne, nie są zaś zgodne.

Zastosowane MNW wymaga znalezienia estymatora wektora parametrów $\boldsymbol{\beta}$ i estymatorów składowych σ^2 i ρ macierzy wariancji-kowariancji składnika losowego modelu (1.10), postaci:

$$\mathrm{E}(\mathbf{u}\mathbf{u}^{\mathrm{T}}) = \mathbf{\Omega} = \sigma^2 \begin{bmatrix} \mathbf{A} & 0 & \dots & 0 \\ 0 & \mathbf{A} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \mathbf{A} \end{bmatrix}, \text{ gdzie } \mathbf{A} = \begin{bmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \vdots & \vdots & \vdots & & \vdots \\ \rho & \rho & \rho & \dots & 1 \end{bmatrix} \quad (1.11)$$

Będą one posiadały wszystkie (asymptotyczne) własności dobrych estymatorów. Zakładając, że składnik losowy **u** ma wielowymiarowy rozkład normalny z wartością oczekiwaną zero i macierzą wariancji-kowariancji $\mathbf{\Omega}$, wykazano, że logarytm funkcji wiarygodności $L(\boldsymbol{\beta}, \sigma^2, \rho)$ osiąga maksima, gdy:

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= \left(\mathbf{X}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{y} \\ \hat{\rho} &= \frac{\sum_{i=1}^{N}\left\{\left[\sum_{t=1}^{T} u_{it}\right]^2 - \sum_{t=1}^{T} u_{it}^2\right\}}{(T-1)\sum_{i=1}^{N}\sum_{t=1}^{T} u_{it}^2} \\ \hat{\sigma}^2 &= \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} u_{it}^2}{NT} \end{aligned} \quad (1.12)$$

gdzie przez u_{it} oznaczono elementy wektora **u**. Równania (1.12) określają estymatory największej wiarygodności poszczególnych parametrów. Następnie zauważono, że ρ, jako udział części wariancji, która spowodowana jest różnicami między obiektami w wariancji całkowitej, jest liczbą z przedziału [0, 1]. Dla dowolnej takiej wartości wzory (1.12) określają wartości parametrów $\boldsymbol{\beta}$ i σ^2, przy których funkcja wiarygodności $L(\boldsymbol{\beta}, \sigma^2, \rho)$ osiąga wartość maksymalną $\hat{L}(\rho)$. Gdyby więc funkcja L osiągała maksimum dla $\rho \notin [0,1]$, to powyższa procedura nie mogłaby być stosowana.

Zastosowanie opisanej metody do modelu popytu na gaz nie dało zadowalających rezultatów, gdyż dla 20 wybranych wartości $\rho \in [0, 1]$ funkcja wiarygodności nie osiągała maksimum. Dlatego poszukiwano alternatywnych estymatorów w klasie estymatorów asymptotycznie równoważnych estymatorom największej wiarygodności. Zastąpienie macierzy $\mathbf{\Omega}$ jej zgodnym estymatorem $\mathbf{S}_m$ prowadzi do estymatorów UMNK wektora $\boldsymbol{\beta}$ o takich samych własnościach granicznych, jak estymatory największej wiarygodności. Do wyznaczenia $\mathbf{S}_m$ potrzebne są zgodne estymatory wektora $\boldsymbol{\beta}$, dla określenia których zastosowano dwustopniową uogólnioną metodę najmniejszych kwadratów, przy założeniu, że wartość obserwacji początkowej jest znana. Estymatory te wykorzystane zostały następnie do znalezienia zgodnych estymatorów σ^2 i ρ (a więc i macierzy $\mathbf{S}_m$) ze wzoru (1.12).

Zauważono, że model jednorównaniowy (1.10) można potraktować jako „quasi-wielorównaniowy”, w którym opóźnione zmienne zależne są określane w równaniach, będących opóźnionymi wersjami samych siebie. Analogia taka wynika z faktu, że niezgodność estymatorów KMNK dla panelowych modeli dynamicznych jest spowodowana korelacją opóźnionych zmiennych endogenicznych ze składnikiem losowym $u_{it} = \alpha_i + \varepsilon_{it}$.

Dowolne równanie modelu (1.10) można zapisać w formie zredukowanej jako:

$$y_{it} = \sum_{k=1}^{K} \sum_{\tau=0}^{\infty} \lambda_{\tau k} z_{i(t-\tau)k} + w_{it} \quad \text{dla } i - 1, ..., N, t = 1, ..., T \tag{1.13}$$

Ponieważ zmienne z_{itk}, tzn. elementy macierzy **Z** są zmiennymi nielosowymi, to są one niezależne od u_{it},. Z kolei składniki losowe równania (1.13) w_{it} są zależne tylko od $u_{it} = \alpha_i + \varepsilon_{it}$. Dlatego możliwe jest estymowanie parametrów $\lambda_{\tau k}$ metodą najmniejszych kwadratów. Po podstawieniu estymatorów $\hat{\lambda}_{\tau k}$, dowolne równanie modelu (1.10) przyjmuje postać:

$$y_{it} = \sum_{k=1}^{K} \alpha_k z_{itk} + \sum_{\tau=1}^{\theta} \sum_{k=1}^{K} \gamma_\tau \hat{\lambda}_{0k} z_{i(t=\tau)k} + ... + u_{it} \tag{1.14}$$

Parametry α_k i γ_τ mogą być już estymowane UMNK i służyć do skonstruowania estymatora $\mathbf{S}_m$, a następnie estymatora wektora $\boldsymbol{\beta}$.

Zastosowanie tej procedury do modelu popytu na gaz dało oszacowania parametrów zgodne z oczekiwaniami. Ocena parametru przy opóźnionym popycie na gaz wyniosła 0,9546, skąd wynika współczynnik zużycia 4,5%. P. Balestra i M. Nerlove [1966] pokazali też, że czynniki specyficzne dla regionów są bardzo istotne dla kształtowania się popytu, ponieważ stanowią one aż 75% wariancji resztowej.

1.5. Współczesne kierunki badań

Początki modelowania na podstawie danych panelowych są związane z badaniami mikroekonomicznymi. Badania takie, np. dotyczące budżetów gospodarstw domowych, są prowadzone na podstawie danych, których wymiar przekrojowy *N* sięga setek lub nawet tysięcy, a wymiar czasowy *T* jest bardzo mały – najczęściej nie przekracza 10. Dane tego typu nazywane są w literaturze mikropanelami (ang. *micropanels*). Używane jest też anglojęzyczne określenie *longitudinal data*, które przetłumaczyć można jako dane wzdłużne. W ostatnim

okresie, w związku z coraz większą dostępnością danych finansowych i makroekonomicznych oraz procesami integracji gospodarek, pojawia się coraz większa liczba badań empirycznych, prowadzonych na podstawie panelowych danych makroekonomicznych, tzw. makropaneli. Charakteryzują się one z reguły niezbyt dużymi wartościami *N* i *T*. Przykładami tego typu baz danych są Datastream, International Financial Statistics, Penn World Table, bazy krajowych urzędów statystycznych. Makropanele, a czasem również mikropanele, charakteryzują się tym, że obserwacje w czasie dotyczą tych samych obiektów – krajów, firm, gospodarstw. Innym rodzajem danych są tzw. pseudopanele. Ich cechą charakterystyczną jest to, że składają się z powtarzanych w czasie obserwacji na określonej liczbie obiektów (firm, gospodarstw), które mogą się wymieniać – dany obiekt nie musi być obserwowany we wszystkich okresach. Ponieważ obserwacje przekrojowe nie muszą dotyczyć tego samego obiektu we wszystkich okresach, konieczne jest w takim przypadku zastosowanie procedur grupowania. Estymacja prowadzona jest przy założeniu nielosowych efektów grupowych na podstawie średnich grupowych, o ile liczba obserwacji w grupie jest wystarczająca. W przeciwnym wypadku należy uwzględnić w procesie estymacji możliwe błędy pomiaru zmiennych. Pionierskim opracowaniem, dotyczącym pseudopaneli jest artykuł A. Deatona [1985]. Przegląd najnowszych badań na ten temat zawarty jest w pracy M. Verbeeka [2008].

Wykorzystanie danych panelowych umożliwia weryfikację wielu hipotez ekonomicznych, które nie mogą być sprawdzone na podstawie pojedynczych szeregów czasowych ani prób przekrojowych. Wpływa też pozytywnie na jakość estymacji. W procesie estymacji panelowych modeli statycznych powstaje jednak szereg problemów, niespotykanych przy korzystaniu z danych czasowych lub przekrojowych. W przypadku panelowych modeli liniowych, jeśli wymiar czasowy próby jest krótki, a standardowe założenie o braku autokorelacji składnika losowego niespełnione, pojawia się problem identyfikacyjny, związany z rozróżnieniem między dynamiką zjawisk a zależnością stanową. Rozróżnienie to pojawiło się w artykule T. W. Andersona i C. Hsiao [1982], a szersze rozważania zawarte są w pracy M. Arellano [2003] (rozdz. 5). Innym standardowym założeniem modeli panelowych jest założenie o stałości parametrów strukturalnych względem obiektów i czasu. Jest ono nieco rozluźniane w przypadku modeli z efektami nielosowymi (ang. *fixed effects model*), ponieważ modele te mają zróżnicowany względem obiektów wyraz wolny. P. A. V. B. Swamy [1970] przedstawił propozycję konstrukcji i estymacji modelu, w którym założenie o stałości parametrów strukturalnych jest uchylone. Model Swamy'ego jest uogólnieniem modelu z efektami losowymi (ang. *random effects model*). Założenia i metody estymacji modeli z losowymi parametrami są omówione szczegółowo w pracy C. Hsiao i M. H. Pesarana [2008]. Odrębną grupę modeli szacowanych na podstawie danych panelowych stanowią tzw. panelowe modele przestrzenne. Charakteryzują się one tym, że

korelacja przekrojowa jest uwzględniona poprzez konstrukcję i wprowadzenie do modelu macierzy zależności przestrzennych. Konstrukcja taka jest możliwa w przypadku, gdy analizowanymi obiektami są jednostki z przestrzeni geograficznej (kraje, regiony *etc.*). Konstrukcja i estymacja modeli tej klasy omówiona jest przez L. Anselina, J. Le Gallo i H. Jayeta [2008].

Dodatkowe trudności powstają, jeśli na podstawie danych panelowych szacuje się model dynamiczny. Zasadniczym przedmiotem uwagi w literaturze światowej poświęconej modelowaniu na podstawie danych panelowych są, jak się wydaje, dynamiczne modele liniowe. Przyczyny komplikacji, proponowane w literaturze metody estymacji takich modeli oraz odniesienia do literatury są omówione szczegółowo w kolejnych rozdziałach niniejszej pracy. Heterogeniczność parametrów strukturalnych przysparza w przypadku modeli dynamicznych dodatkowych problemów. Pierwszą pracą, w której rozważano ten problem, był artykuł Robertsona i Symonsa [1992]. Autorzy opisywali własności różnych estymatorów w sytuacji, gdy rzeczywiste wartości parametrów strukturalnych różnią się względem obiektów, a konstrukcja modelu zakłada ich identyczność dla wszystkich obiektów. Za pomocą eksperymentów Monte Carlo Robertson i Symons [1992] wykazali, że jeśli właściwą postacią modelu jest statyczny model z heterogenicznymi parametrami, a postacią estymowaną – dynamiczny model z parametrami homogenicznymi dla wszystkich obiektów, to parametr autoregresyjny jest przeszacowany, a parametry przy pozostałych zmiennych objaśniających – niedoszacowane. Jeśli dodatkowo zmienne objaśniające są procesami błądzenia losowego, to przy $T \to \infty$ $p\lim \hat{\gamma} = 1$, zaś $p\lim \hat{\beta} = 0$. Wyniki te są niezależne od wymiaru czasowego i przekrojowego próby. Pesaran i Smith [1995] stosowali do estymacji modelu dynamicznego z heterogenicznymi parametrami strukturalnymi różne techniki, w celu uzyskania zgodnych estymatorów średnich wartości parametrów. Dla małych wartości T estymatory każdej z zastosowanych metod (regresji czasowej na podstawie średnich grupowych, regresji przekrojowej na podstawie średnich czasowych, modeli FE i RE, obliczania średnich grupowych ocen parametrów na podstawie osobnych modeli dla każdego z obiektów) nie były zgodne. Jedynie dla dużych wartości N i T estymatory średnich grupowych wartości parametrów uzyskane z regresji przekrojowej były zgodne. Maddala, Trost, Li i Joutz [1997] proponują zastosowanie do estymacji modeli panelowych z heterogenicznymi parametrami tzw. *shrinkage estimators*. Mają one, zdaniem autorów, przewagę zarówno nad estymatorami stosowanymi dla indywidualnych modeli szeregów czasowych, jak i nad estymatorami stosowanymi standardowo dla modeli panelowych. *Shrinkage estimators* są polecane zwłaszcza dla celów prognostycznych. Wynikom Maddali i in. [1997] wydają się zaprzeczać m. in. wyniki prac: Baltagi i Griffina [1997], Baltagi, Bressona, Griffina i Pirotte'a [2003]. Wyniki empiryczne prezentowane w obu tych artykułach

dotyczą popytu na benzynę, odpowiednio w 18 krajach OECD i w 21 regionach Francji. B a l t a g i i G r i f f i n [1997] stwierdzili, że najbardziej poprawne wyniki uzyskano na podstawie modeli o homogenicznych parametrach. Dotyczy to ocen elastyczności krótkookresowych i jakości prognoz, mierzonej RMSE. Podobne rezultaty uzyskali B a l t a g i, B r e s s o n, G r i f f i n i P i r o t t e [2003]. Wyniki uzyskane na podstawie modeli szeregów czasowych dla poszczególnych regionów były niewiarygodne i dawały najsłabsze prognozy, estymatory *shrinkage* charakteryzowały się niską efektywnością, a standardowe procedury panelowe dały najlepsze prognozy. Wśród alternatywnych rozwiązań problemu heterogeniczności parametrów strukturalnych w panelowych modelach dynamicznych proponowanych w literaturze wymienić można zastosowanie podejścia Bayesowskiego, sugerowane przez H s i a o, P e s a r a n a i T a h m i s c i o g l u [1999], lub zastosowanie podejścia EGARCH, sugerowane przez L a h i r i i L i u [2006].

Pojawiają się też prace, dotyczące metod estymacji modeli panelowych z jakościową zmienną objaśnianą, statycznych i dynamicznych. Ich przegląd znaleźć można np. w B a l t a g i [2008] (s. 237–268), H s i a o [2003] (s. 188–224), lub A r e l l a n o, H o n o r é [2001] (s. 3270–3290). Procedury stosowane w przypadku panelowych modeli VAR, z uwzględnieniem sytuacji, gdy wymiar czasowy T jest bardzo mały, są przedstawione w pracy B i n d e r a, H s i a o i P e s a r a n a [2005]. Problem niestacjonarności w modelach panelowych jest rozważany m. in. przez M o o n a i P e r r o n a [2007], P e d r o n i e g o [2004, 2007] i w innych, wcześniejszych pracach.

Liczba pozycji literaturowych na temat makro- i mikropaneli, paneli niezbilansowanych, modeli nieliniowych, modeli dynamicznych, skutków nieuwzględnienia heterogeniczności parametrów, modeli wielorównaniowych, modeli niestacjonarnych i innych zagadnień, związanych ze stosowaniem w procesie estymacji danych panelowych, wciąż rośnie. Wśród najnowszych pozycji, w których dokonano bardzo szerokiego przeglądu tych zagadnień, wymienić można: B a l t a g i [2008]; M a t y a s, S e v e s t r e [2008]; H s i a o [2003]; W o o l d r i d g e [2002].

1.6. Sposoby uwzględnienia dynamiki modelowanych zjawisk

W podrozdz. 1.2 niniejszego rozdziału wspomniano, że zbiory danych panelowych mogą różnić się między sobą m. in. pod względem wymiaru czasowego T. Jeśli wymiar ten jest stosunkowo duży, to liczba posiadanych obserwacji umożliwia analizę dynamiki modelowanych zjawisk. Analiza taka jest w niektórych przypadkach kluczowa, np. jeśli model ma służyć do prognozowania. Dynamika czasowa może być uwzględniona w modelu szacowanym na podstawie danych panelowych na kilka różnych sposobów.

Jedną z prostszych metod jest wprowadzenie do zbioru zmiennych objaśniających zmiennej czasowej. Na przykład dla odzwierciedlenia trendu czasowego można przyjąć $x_{itk} = t$.

Innym sposobem jest wykorzystanie specyfikacji modelu dwuczynnikowego (ang. *two-way model*) postaci:

$$y_{it} = \alpha + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + (\mu_i + \lambda_t + \varepsilon_{it}) \text{ dla } i = 1, ..., N, t = 1, ..., T \tag{1.15}$$

gdzie składnik losowy zawiera, oprócz efektów grupowych μ_i i „właściwego" składnika losowego ε_{it}, zmienne w czasie, a stałe względem obiektów, efekty czasowe λ_t. Efekty te można w dalszym ciągu analizy traktować jako losowe lub nielosowe[6], ale w każdym z tych przypadków interpretuje się je jako efekt działania czynnika czasu, specyficzny tylko dla danego okresu, a niezależny względem obiektów.

Dynamikę czasową uwzględnić również można poprzez konstrukcję modelu, w którym parametry zmieniają się w czasie, a więc modelu postaci (por. F r e e s [2005]):

$$y_{it} = \alpha + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta}_t + \varepsilon_{it} \text{ dla } i = 1, ..., N, t = 1, ..., T \tag{1.16}$$

Ocena parametru β_{kt} (k-tej współrzędnej wektora $\boldsymbol{\beta}_t$) przy zmiennej objaśniającej x_{kit} odzwierciedla wówczas wpływ zmiennej x_k na zmienną y, specyficzny dla okresu t, a przeciętny dla wszystkich obiektów.

Przekrojowo-czasowy charakter danych umożliwia też uwzględnienie trendu czasowego przez analizę zmian w zmiennej endogenicznej. Przykładem może być model:

$$\Delta y_{it} = \mu_i + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + \varepsilon_{it} \tag{1.17}$$

gdzie $\Delta y_{it} = y_{it} - y_{i,t-1}$. Wadą tego sposobu jest utrata n obserwacji (pierwszych obserwacji dla wszystkich obiektów), spowodowana obliczaniem przyrostów. Przy stosunkowo dużych rozmiarach próby nie stanowi to jednak problemu.

Model postaci:

$$y_{it} = \mu_i + \gamma y_{i,t-1} + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + \varepsilon_{it} \tag{1.18}$$

jest uogólnieniem ostatniego przypadku. Parametr γ, zwany parametrem autoregresyjnym, mierzy siłę związku między bieżącą a opóźnioną wartością zmiennej endogenicznej.

---

[6] Omówienie klasyfikacji modeli panelowych oraz metod ich estymacji znajduje się w rozdz. 2 niniejszej publikacji.

Modelem autoregresyjnym jest również model postaci:

$$y_{it} = \mu_i + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + \varepsilon_{it}, \quad \text{gdzie} \quad \varepsilon_{it} = \rho\varepsilon_{i,t-1} + \eta_{it} \tag{1.19}$$

Cechę autoregresji posiada tu składnik losowy, a nie, jak w przypadku modelu (1.18), zmienna endogeniczna. Między modelami (1.18) i (1.19) występuje jednak podobieństwo. Widoczne ono się staje, jeśli obliczy się na podstawie modelu (1.19) wartość $y_{it} - \rho y_{i,t-1}$ i dokona prostego przekształcenia. Otrzymuje się wówczas:

$$y_{it} = \mu_i^* + \rho y_{i,t-1} + \left(\mathbf{x}_{it} - \rho\mathbf{x}_{i,t-1}\right)^{\mathrm{T}}\boldsymbol{\beta} + \eta_{it} \tag{1.20}$$

Wartości $\mu_i^* = \mu_i\ (1 - \rho)$ i ρ z równania (1.20) równe są odpowiednio wartościom μ_i i γ z równania (1.18). Widoczną różnicą jest postać wektora zmiennych objaśniających.

Powyżej wskazano kilka możliwości uwzględnienia aspektu czasowego w konstrukcji modeli szacowanych na podstawie danych panelowych. Jedną z częściej stosowanych wydaje się zastosowanie modelu postaci (1.18). Poprawne oszacowanie jego parametrów wymaga zastosowania specyficznych metod estymacji, innych niż dla panelowych modli statycznych. O ile w przypadku klasycznych modeli ekonometrycznych można stosować jednakowe metody estymacji dla modeli statycznych i dynamicznych, to w przypadku danych panelowych wprowadzenie do modelu opóźnionej zmiennej objaśnianej wpływa negatywnie na własności estymatorów modeli statycznych. W literaturze przedmiotu proponuje się różne metody estymacji panelowych modeli dynamicznych oparte na uogólnionej metodzie momentów (GMM) lub metodzie największej wiarygodności (MNW). Największą popularność zyskały, jak się wydaje, metoda GMM pierwszych różnic (FDGMM) opracowana przez Arellano i Bonda [1991] oraz systemowa metoda GMM (SGMM) zaproponowana w pracach Arellano i Bover [1995] i Blundella i Bonda [1998]. Metody te są omówione w rozdz. 3 niniejszej pracy.

Literatura dotycząca różnych aspektów modelowania na podstawie modeli dynamicznych jest bogata. Hsiao [2003] (s. 73–78) wykazał, że dla modelu dynamicznego z efektami losowymi (ang. *random effects*) zgodność estymatorów MNW i UMNK zależy od założeń o wartościach początkowych i od tego, w jaki sposób N i T dążą do nieskończoności. Arellano i Honoré [2001] (rozdz. 53, s. 3229–3296) zwracają szczególną uwagę na konsekwencje ścisłej egzogeniczności zmiennych objaśniających dla identyfikowalności parametrów strukturalnych, w porównaniu z przypadkiem, gdy zakłada się słabą egzogeniczność tych zmiennych. Propozycje metod estymacji panelowych modeli dynamicznych, alternatywnych w stosunku do GMM przedstawiane były m. in.

w: Chamberlain [1982, 1984]; Keane, Runkle [1992]; Arellano [1990]. Wśród najnowszych opracowań wymienić warto publikacje: Kniesner, Li [2002] oraz Park, Sickles, Simar [2007]. G. Chamberlain [1982, 1984] rozważał dynamiczny model panelowy z arbitralną strukturą zakłóceń (ang. *arbitrary error structure*). Model panelowy traktowany był jako regresja wielowymiarowa T równań z restrykcjami. Do jego estymacji zaproponowano efektywny estymator minimalnej odległości, odporny na autokorelację resztową. W publikacji z 1984 r., jako alternatywa dla tego estymatora zaproponowany został estymator asymptotycznie mu równoważny, bazujący na zastosowaniu trójstopniowej MNK (3MNK) do $(T - 2)$ równań na przyrostach. Kontynuacją idei Chamberlaina była praca M. Arellano [1990], który przedstawił możliwości testowania różnych restrykcji nałożonych na kowariancje. Testy te dotyczą w szczególności modeli z efektami losowymi oraz modeli pierwszych różnic z procesami: białego szumu, średniej ruchomej oraz autoregresyjnym. Jeśli restrykcje te są prawdziwe, to 3MNK jest nieefektywna i w takiej sytuacji Arellano [1990] zaproponował estymator UMNK, który jest asymptotycznie efektywny w tym sensie, że jego rozkład graniczny jest taki sam, jak rozkład estymatora Chamberlaina. Potrzebę modelowania momentów wyższego rzędu w przypadku procesów dynamicznych przy wykorzystaniu danych panelowych podkreślali C. Meghir i F. Windmeijer [1999]. Praca ta dotyczyła procesów typu ARCH z multiplikatywnymi efektami grupowymi. T. Kniesner, Q. Li [2002] rozważali modele semiparametryczne, w których postać funkcyjna zależności między zmiennymi objaśniającymi a zmienną objaśnianą nie jest znana. Zaproponowali oni dwustopniową, semiparametryczną, $\sqrt{N}$ -zgodną procedurę estymacji. Estymator semiparametryczny dla dynamicznego modelu panelowego zaproponowali również B. Park, R. Sickles, L. Simar [2007]. Jest to efektywny estymator dla stochastycznego modelu granicznego (ang. *stochastic frontier model*).

1.7. Rola warunków początkowych w procesie estymacji dynamicznych modeli panelowych

Estymacja panelowych modeli dynamicznych pociąga za sobą problemy niespotykane w modelowaniu modeli statycznych. Problemy te spowodowane są nie tylko tym, że własności wielu estymatorów różnią się w przypadku statycznym i dynamicznym. Ponadto w przypadku dynamicznym poprawne sformułowanie funkcji wiarygodności wymaga nie tylko określenia natury efektów grupowych (losowe czy nielosowe), ale również przyjęcia założeń, dotyczących tzw. warunków początkowych, czyli założeń o procesie, generującym dane sprzed okresu próby. Dla skończonych wartości T interpretacja modelu zależy od założeń dotyczących wartości początkowej y_{i0}. Ponadto od tych założeń

zależą statystyczne własności estymatorów metody największej wiarygodności i uogólnionej metody najmniejszych kwadratów. Niezależne od nich są natomiast własności metody zmiennych instrumentalnych i uogólnionej metody momentów.

Rozwój badań dotyczących konsekwencji różnych warunków początkowych nastąpił w latach 80. XX w. Idee, które poprzedzały ten rozwój, przedstawione zostały w pracach G. S. Maddali [1975] i A. Trognona [1978]. Maddala [1975] podkreślił ważność początkowej obserwacji w procesie estymacji modeli panelowych za pomocą metody największej wiarygodności. Odszedł on od założeń przyjmowanych we wcześniejszych eksperymentach, że wartości początkowe są *fixed*, natomiast przyjął, że rozkład obserwacji początkowych jest określany przez scenariusz eksperymentu Monte Carlo. Rozwiązania brzegowe uzyskane w wyniku tak zaprojektowanego eksperymentu różniły się od uzyskiwanych we wcześniejszych eksperymentach. W późniejszej publikacji Maddala napisał: „założenia czynione odnośnie do y_{i0} są ważne, ale najrozsądniejszym wyjściem jest przyjąć założenie, że rozkład y_{i0} jest implikowany przez model" (Maddala [1987], s. 316). Trognon [1978] wykazał, że jeśli egzogeniczna zmienna x jest generowana przez proces autoregresyjny pierwszego rzędu:

$$x_t = x_{t-1} + w_t \tag{1.21}$$

gdzie w jest białym szumem o wartości oczekiwanej 0 i skalarnej macierzy wariancji-kowariancji, to w wyniku maksymalizacji funkcji wiarygodności, która traktuje opóźnioną zmienną objaśnianą jak zmienną z góry ustaloną (a zatem tak samo, jak każdą inną zmienną objaśniającą x_{it}), a wartości początkowe jako *fixed*, otrzymuje się niezgodne rozwiązania brzegowe dla stosunkowo dużego zakresu rzeczywistych wartości parametrów. W szczególności Trognon [1978] wykazał istnienie wartości krytycznej $\gamma^* = [(T-3)^2 - 8]/(T+1)^2$ takiej, że dla $\gamma < \gamma^*$ można uzyskać zgodne estymatory metody największej wiarygodności parametru γ przy zmiennej $y_{i,t-1}$, a dla $\gamma > \gamma^*$ jest to niemożliwe.

Kolejnymi ważnymi opracowaniami, w których rozważana była rola warunków początkowych i implikacje płynące z przyjęcia różnych założeń dotyczących wartości y_{i0}, są prace Andersona i Hsiao [1981, 1982][7] oraz Sevestre i Trognona [1985]. W tej ostatniej pracy rozważany był panelowy model AR(1) (niezawierający dodatkowych zmiennych objaśniających). Autorzy wykazali, że asymptotyczne własności estymatorów UMNK i UMNK z estymacją zależą, dla skończonej liczby T i $N \to \infty$, od wartości oczekiwanych

[7] Rozważania zawarte w obu tych pracach, dotyczące założeń o wartościach początkowych i proponowanych estymatorów metody zmiennych instrumentalnych, są omawiane bardziej szczegółowo w rozdz. 3. W tym miejscu zauważmy jedynie, że zastosowanie metody zmiennych instrumentalnych często powoduje, iż zgodność estymatora nie zależy od wartości początkowej y_{i0}.

$\mathrm{E}\left(y_{i0}^2\right)$ i $\mathrm{E}\left(y_{i0}\alpha_i\right)$[8]. Sevestre i Trognon [1985] udowodnili również następującą zależność między estymetorami: wewnątrzgrupowym $\left(\hat{\gamma}(0)_{\mathrm{W}}\right)$, uogólnionej MNK $\left(\hat{\gamma}(\lambda)_{\mathrm{UMNK}}\right)$, *pooled* $\left(\hat{\gamma}(1)_{\mathrm{KMNK}}\right)$ i miedzygrupowym $\left(\hat{\gamma}(\infty)_{\mathrm{B}}\right)$ dla panelowych modeli AR(1):

$$p\lim \hat{\gamma}(0)_{\mathrm{W}} < \gamma < p\lim \hat{\gamma}(\lambda)_{\mathrm{UMNK}} < p\lim \hat{\gamma}(1)_{\mathrm{KMNK}} < p\lim \hat{\gamma}(\infty)_{\mathrm{B}},$$

gdzie: $\hat{\gamma}(\lambda)$ jest estymatorem uzyskanym przez rozwiązanie układu równań normalnych UMNK dla dowolnej wartości λ, $\lambda = \dfrac{\xi}{\eta}$; ξ i η są pierwiastkami charakterystycznymi macierzy $\mathbf{\Omega}$ (gdzie $(\mathrm{D}^2(\mathbf{u}) = \sigma^2\,\mathbf{\Omega})$: $\xi = 1-\rho+T\rho$ (pierwiastek T-krotny), $\eta = 1-\rho$ (pierwiastek $N(T-1)$-krotny), $\rho = \sigma_\alpha^2/\sigma^2$, $\sigma^2 = \sigma_\alpha^2 + \sigma_\varepsilon^2$.

Ridder i Wansbeek [1990] (s. 566–571) zauważyli, że estymator $\hat{\gamma}(\lambda)_{\mathrm{UMNK}}$ jest niezgodny, nawet jeśli został on uzyskany na podstawie zgodnych oszacowań parametru ρ, zwanego współczynnikiem korelacji wewnątrzgrupowej (ang. *intraclass correlation*). Powodem jest korelacja między opóźnioną zmienną objaśnianą a transformowanym składnikiem losowym.

Wśród nowszych opracowań wagę warunków początkowych podkreśla Nerlove [2000]. Rozważa on model z jedną zmienną egzogeniczną postaci:

$$y_{it} = \beta x_{it} + \gamma y_{i,t-1} + u_{it} = \beta x_{it} + \gamma y_{i,t-1} + \alpha_i + \varepsilon_{it},\ i = 1, ..., N,\ t = 1, ..., T \qquad (1.22)$$

o którym zakłada, że składnik losowy u_{it} spełnia założenia modelu RE. Model (1.22), uwzględniając założenia RE, zapisać można w postaci:

$$y_{it} = \gamma^t y_{i0} + \sum_{j=0}^{t-1} \gamma^j \beta x_{i,t-j} + \frac{1-\gamma^t}{1-\gamma}\alpha_i + \nu_{it},\ \text{gdzie}\ \nu_{it} = \sum_{j=0}^{t-1} \gamma^j \varepsilon_{i,t-j} \qquad (1.23)$$

Zgodnie z równaniem (1.23) y_{it} jest sumą czterech składników, z których pierwszy zależy od wartości początkowej y_{i0}, drugi – od opóźnionych wartości zmiennej egzogenicznej, trzeci składnik zależny jest od efektów grupowych, czwarty zaś – od opóźnionych wartości właściwego składnika losowego ε_{it}. Ostatni składnik (1.23), ν_{it}, jest procesem autoregresyjnym o wartości początkowej $\nu_{i0} = 0$ i $\nu_{it} = \gamma\nu_{i,t-1} + \varepsilon_{it}$. Założenie o stacjonarności ($|\gamma| < 1$) nie jest wymagane, o ile T jest skończone. Na podstawie (1.23) widać, że y_{i0} jest

[8] Autorzy przyjmowali trzy alternatywne założenia dotyczące nielosowości y_{i0} lub losowości i rozkładu y_{i0}. Ridder i Wansbeek [1990] rozszerzyli rozważania na przypadek modelu ze stacjonarną zmienną egzogeniczną.

skorelowane zarówno z α_i jak i z opóźnionymi wartościami zmiennej x; ponadto ε_{i0} jest zmienną losową o wartości oczekiwanej 0 i wariancji σ_ε^2, o jednakowych, niezależnych rozkładach dla każdego i. Zatem wartość początkową można zapisać jako:

$$y_{i0} = f\left(x_{i0}, x_{i1}, ..., \alpha_i, \varepsilon_{i0}\right) \tag{1.24}$$

Przy założeniu, że efekty grupowe α_i są nielosowe (model FE), a opóźnione wartości zmiennej x dane, y_{i0} są również nielosowe i nieskorelowane ze składnikiem losowym v_{it} równania (1.23) dla każdego t. Natomiast przy założeniu, że α_i są losowe (model RE), wartości początkowe nie są egzogeniczne. Wartości te są wówczas skorelowane z α_i, jako że są one częścią składnika losowego – trzecim i czwartym składnikiem (1.23).

Założenie, że wartości początkowe są nielosowe, było przyjmowane w literaturze w początkach modelowania dynamicznego. Dopiero Anderson i Hsiao [1981] i [1982] zainicjowali rozważania, dotyczące różnych założeń o wartościach początkowych i konsekwencji wynikających z tych założeń.

W literaturze lat 80. zakłada się na ogół, że wartości początkowe są zmiennymi losowymi o jednakowych, niezależnych rozkładach, charakteryzowanych przez ich wariancje i korelacje z efektami grupowymi. Poza tym nie muszą one być generowane przez ten sam proces, który generuje pozostałe wartości y_{it}. Własności różnych estymatorów zależą od procesu generującego y_{i0}. Proces ten może być modelowany jednocześnie z modelem „podstawowym" (1.22). Rozwiązanie takie zaproponowali jako pierwsi Bhargava i Sargan [1983]. Założyli oni, że wartości początkowe są generowane przez ten sam proces, który generuje obserwacje w próbie. Potraktowali oni ponadto panelowy model dynamiczny jako układ równań łącznie współzależnych. Dokładniej, jeśli dane panelowe pochodzą z N obiektów i T okresów, to model łącznie współzależny jest złożony z $T + 1$ równań (osobne równanie dla każdego okresu próby $t = 1, ..., T$ i dla wartości początkowej $t = 0$), z których każde szacowane jest na podstawie N-elementowej próby. Autorzy rozważali model, w którym występują zarówno zmienne objaśniające skorelowane ze składnikiem losowym, jak i zmienne nieskorelowane, a do jego estymacji zastosowali metodę największej wiarygodności.

Próbę łącznego modelowania procesów, generujących wartości początkowe i obserwacje zawarte w próbie, podjęto w późniejszych latach, w pracach Ridder a i Wansbeeka [1990] oraz Arellano i Honoré [2001].

Warto w tym miejscu podkreślić, że estymatory panelowych modeli dynamicznych, bazujące na GMM, nie zależą od warunków początkowych. Ponieważ modelowanie procesów, generujących te warunki jest trudne, brak zależności o którym wspomniano wyżej, jest zapewne jedną z przyczyn popularności metod opartych na GMM.

ROZDZIAŁ 2

ZARYS PROBLEMATYKI ESTYMACJI MODELI STATYCZNYCH NA PODSTAWIE DANYCH PANELOWYCH I PRZEKROJOWO-CZASOWYCH

Przedmiotem rozważań tego rozdziału są modele statyczne, szacowane na podstawie danych łączących wymiar czasowy i przekrojowy.

Początkowe podrozdziały dotyczą problematyki szacowania i testowania modeli panelowych, to znaczy modeli szacowanych na podstawie danych, w których liczba obserwacji dla obiektów N przekracza często wielokrotnie liczbę obserwacji w czasie T. Konstrukcja takich modeli zakłada niezmienność parametrów przy zmiennych objaśniających względem czasu i obiektów. Heterogeniczność obiektów lub zróżnicowanie modelowanego zjawiska w czasie uwzględnia się w tego typu modelach poprzez zróżnicowanie względem obiektów (i/lub czasu) składnika losowego albo, w zależności od założeń, wyrazu wolnego. W literaturze anglojęzycznej modele takie nazywa się odpowiednio *random effects models* (lub *error--components models* lub *variance components models*), albo *fixed effects models* (lub *variable-intercept models*). U podstaw takiej konstrukcji modelu leży założenie, że źródłem heterogeniczności jest działanie na modelowane zjawisko czynników nieobserwowalnych, a zatem nieuwzględnionych w modelu. Efekty wpływu tych czynników mogą być stałe w czasie, a specyficzne dla obiektów (efekty sposobu zarządzania, zdolności, płeć), lub stałe dla obiektów, a zmieniające się w czasie (ceny, stopy procentowe). W modelach rozważanego typu zakłada się, że efekty działania pojedynczych pominiętych zmiennych są nieistotne, ale ich łączny wpływ jest istotny. Ponadto jeśli efekty te uwzględnia się w modelu jako składową składnika losowego, to zakłada się dodatkowo, że nie są one skorelowane ze zmiennymi objaśniającymi modelu. Jeśli zaś różnicuje się wyraz wolny, to założenie takie nie jest konieczne. Możliwe jest też dopuszczenie, aby część zmiennych objaśniających była skorelowana z efektami grupowymi, a pozostałe zmienne nie wykazywały takiej korelacji. Parametry takiego modelu można oszacować za pomocą estymatora H a u s m a n a i T a y l o r a [1981].

W niniejszym rozdziale przedstawiono też problematykę szacowania i testowania modeli bazujących na typowych danych czasowo-przekrojowych, w których liczba obserwacji w czasie jest większa niż liczba obiektów. Metody analizy prowadzonej na podstawie takich danych są zbliżone do metod stosowanych w analizach szeregów czasowych, gdyż wymiar czasowy jest na tyle długi,

że możliwe jest prowadzenie analizy w kontekście procesów stochastycznych z uwzględnieniem zjawisk heteroskedastyczności i (auto)korelacji. W końcowej części niniejszego rozdziału zaprezentowano zarys metod estymacji wielorównaniowych modeli statycznych o równaniach pozornie niezależnych i o równaniach łącznie współzależnych.

2.1. Modele panelowe ze zmiennymi sztucznymi (modele FE)

Struktura danych panelowych, charakteryzująca się dużą liczbą obiektów i krótkością szeregów czasowych, spowodowała konieczność opracowania specyficznych metod estymacji modeli, budowanych w oparciu o takie dane. Do szacowania jednorównaniowych modeli panelowych, w których nie występuje opóźniona zmienna endogeniczna stosuje się, w zależności od typu modelu, metody oparte na klasycznej MNK lub uogólnionej MNK. Zaznaczyć warto, że estymacja dynamicznych modeli panelowych wymaga stosowania odrębnych metod, które opisane są w rozdz. 3.

2.1.1. Założenia konstrukcyjne i metody estymacji

Statyczny model panelowy można zapisać jako:

$$y_{it} = \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + u_{it} = \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + (\alpha_i + \varepsilon_{it}), \quad i = 1, ..., N, t = 1, ..., T \tag{2.1}$$

gdzie i – numer obiektu (województwa, firmy *etc.*), t – numer okresu; $\mathbf{x}_{it}^{\mathrm{T}} = [x_{1it}, ..., x_{Kit}]$ – wektor obserwacji na zmiennych objaśniających; $\boldsymbol{\beta} = [\beta_1, ..., \beta_K]^{\mathrm{T}}$ – wektor parametrów strukturalnych[1]; $u_{it} = \alpha_i + \varepsilon_{it}$ – składnik losowy, w którym wyróżnia się stałą w czasie składową właściwą dla i-tego obiektu α_i, zwaną efektem grupowym; oraz „właściwy” składnik losowy ε_{it}. Model (2.1) nazywa się modelem z efektami grupowymi (ang. *one way model*)[2].

---

[1] Model z parametrami strukturalnymi, zmieniającymi się dla obiektów, zaproponowany został w pracy S w a m y'e g o [1970]. Jego konstrukcja i estymacja jest omówiona m. in. w: D a ń s k a [2000] (podrozdz. 1.1).

[2] Konstruowane są też modele, w których u_{it} rozdzielany jest na trzy składowe: $u_{it} = \alpha_i + \lambda_t + \varepsilon_{it}$. Stała względem obiektów składowa λ_t nazywana jest efektem czasowym. Tego typu modele określa się w literaturze anglojęzycznej jako *two way modes*. Ich estymacja przebiega analogicznie, jak estymacja modeli postaci (2.1), co zostało omówione dokładniej np. w: B a l t a g i [2008] (s. 35–50), a w języku polskim w: D a ń s k a [2000] (podrozdz. 2.1).

W procesie estymacji i wnioskowania efekty grupowe można traktować jako składową składnika losowego lub wyraz wolny, zróżnicowany względem obiektów. Niniejszy podrozdział dotyczy modeli drugiej z wymienionych kategorii, zwanych modelami ze zmiennymi sztucznymi lub modelami z efektami nielosowymi (ang. *fixed effects models*, FE). Przy założeniach, że: α_i są nielosowe, traktowane jako N nieznanych parametrów podlegających estymacji, x_{kit} są niezależne od ε_{it} dla wszystkich k, i, t, $\varepsilon_{it} \sim \mathrm{N}(0, \sigma_\varepsilon^2)$ dla wszystkich i i t, model FE zapisać można w postaci:

$$y_{it} = \alpha_i + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + \varepsilon_{it}, \quad i = 1, ..., N, \ t = 1, ..., T \tag{2.2}$$

Wyraz wolny α_i modelu (2.2) różni się względem obiektów i odzwierciedla nieobserwowalny i nieuwzględniony w równaniu regresji efekt specyficzny dla i-tego obiektu, a parametry przy zmiennych objaśniających są jednakowe dla wszystkich obiektów i okresów. Oznacza to, że efekty zmian zmiennych x_{kit} są identyczne we wszystkich okresach i dla wszystkich obiektów, ale przeciętny poziom zjawiska może być różny dla różnych obiektów.

Przy założeniu, że dane pogrupowane są po obiektach, tzn. pierwszych T obserwacji pochodzi z pierwszego obiektu, kolejnych T obserwacji – z drugiego obiektu *etc*., model (2.2) zapisać można w postaci macierzowej:

$$\mathbf{y} = \mathbf{Z}_\alpha \boldsymbol{\alpha} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.3}$$

gdzie: $\mathbf{y} = [y_{11} \ ... \ y_{1T} \ ... \ y_{N1} \ ... \ y_{NT}]^{\mathrm{T}}$ jest wektorem zmiennych endogenicznych o wymiarach ($NT \times 1$); $\mathbf{X}$ – macierzą zmiennych egzogenicznych o wymiarach ($NT \times K$); $\boldsymbol{\alpha}^{\mathrm{T}} = [\alpha_1, ..., \alpha_N]$, $\boldsymbol{\varepsilon}^{\mathrm{T}} = [\varepsilon_{11}, ..., \varepsilon_{1T}, ..., \varepsilon_{N1}, ..., \varepsilon_{NT}]$. Macierz $\mathbf{Z}_\alpha = \mathbf{I}_N \otimes \mathbf{i}_T$ jest macierzą zmiennych zero-jedynkowych. Dla przykładu, dla $N = 3$, $T = 2$, budowa macierzy $\mathbf{Z}_\alpha$ jest następująca[3]:

$$\mathbf{Z}_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

[3] Symbol $\otimes$ oznacza iloczyn Kroneckera macierzy.

Pomnożenie tej macierzy przez wektor **α** daje wektor efektów grupowych:

$$\mathbf{Z}_\alpha \boldsymbol{\alpha} = [\alpha_1 \quad \ldots \quad \alpha_1 \quad , \ldots, \quad \alpha_N \quad \ldots \quad \alpha_N]^T,$$

gdzie sekwencje identycznych elementów mają długość *T*.

Model FE spełnia założenia klasycznej metody najmniejszych kwadratów (KMNK). Jej zastosowanie w praktyce może jednak być utrudnione, ponieważ potencjalnie duża liczba zmiennych sztucznych (kolumn macierzy $\mathbf{Z}_\alpha$) może spowodować problemy techniczne z odwracalnością macierzy[4] ($[\mathbf{Z}_\alpha \mid \mathbf{X}]^T \cdot [\mathbf{Z}_\alpha \mid \mathbf{X}]$). Dlatego estymacja parametrów modelu (2.3) za pomocą KMNK musi być poprzedzona opisaną poniżej transformacją tego modelu.

Model średnich po czasie, uzyskany na podstawie (2.2), ma postać:

$$\bar{y}_{i.} = \alpha_i + \bar{\mathbf{x}}_{i.}^{\mathrm{T}} \boldsymbol{\beta} + \bar{\varepsilon}_{i.} \tag{2.4}$$

gdzie symbol $\bar{\mathbf{x}}_{i.}^{\mathrm{T}}$ oznacza wektor, którego elementami są tzw. średnie grupowe, tzn. średnie po czasie z obserwacji dla *i*-tego obiektu; średnie grupowe innych zmiennych oznaczono w analogiczny sposób. Średnia po obiektach na podstawie równania (2.4) ma postać[5]:

$$\bar{y}_{..} = \bar{\mathbf{x}}_{..}^{\mathrm{T}} \boldsymbol{\beta} + \bar{\varepsilon}_{..} \tag{2.5}$$

Odejmując stronami równania (2.2) i (2.4) otrzymuje się:

$$y_{it} - \bar{y}_{i.} = \left(\mathbf{x}_{it} - \bar{\mathbf{x}}_{i.}\right)^{\mathrm{T}} \boldsymbol{\beta} + \left(\varepsilon_{it} - \bar{\varepsilon}_{i.}\right) \tag{2.6}$$

Przekształcenia (2.4)–(2.6) są równoważne pomnożeniu stronami równania (2.3) przez tzw. operator wewnątrzgrupowy **Q**, zdefiniowany jako:

$$\mathbf{Q} = \mathbf{I}_N \otimes \left(\mathbf{I}_T - \frac{\mathbf{J}_T}{T} \right) \tag{2.7}$$

gdzie $\mathbf{I}_N$ oznacza macierz jednostkową stopnia *N*, $\mathbf{J}_T$ – macierz jedynkową stopnia *T*. Pomnożenie **Q** przez dowolną macierz przekształca ją w macierz, której elementami są odchylenia od średnich grupowych. Dla przykładu,

---

[4] Zmiennych sztucznych jest tyle, ile obiektów, a zatem macierz ta ma wymiary $((N + K) \times NT)$.

[5] Wykorzystuje się tutaj arbitralnie narzucone ograniczenie $\sum \alpha_i = 0$, wprowadzone w celu uniknięcia współliniowości zmiennych.

typowym elementem iloczynu **Qy** (gdzie **y** określone jest jak w równaniu (2.3)) jest:

$$\mathbf{Qy} = \left[\left(\mathbf{I}_T - \frac{\mathbf{J}_T}{T}\right)y_1 \quad \cdots \quad \left(\mathbf{I}_T - \frac{\mathbf{J}_T}{T}\right)y_N\right]^{\mathrm{T}} =$$

$$= [y_{11} - y_{1.} \quad \cdots \quad y_{1T} - y_{1.} \quad \cdots \quad y_{N1} - y_{N.} \quad \cdots \quad y_{NT} - y_{N.}]^{\mathrm{T}}.$$

Ponadto **Q** jest macierzą symetryczną i idempotentną (tzn. $\mathbf{Q} = \mathbf{Q}^{\mathrm{T}}$, $\mathbf{Q} = \mathbf{Q}^2$). Estymator KMNK modelu (2.6) lub równoważnego mu modelu

$$\mathbf{Qy} = \mathbf{QZ}_{\alpha}\boldsymbol{\alpha} + \mathbf{QX}\boldsymbol{\beta} + \mathbf{Q}\boldsymbol{\varepsilon} \tag{2.8}$$

jest łatwy do wyznaczenia, ponieważ wymaga odwrócenia macierzy $\mathbf{X}^{\mathrm{T}}\mathbf{QX}$ stopnia K. Estymator ten ma postać:

$$\hat{\boldsymbol{\beta}}_{\mathrm{W}} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{QX}\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Qy} \tag{2.9}$$

i nazywany jest estymatorem wewnątrzgrupowym lub estymatorem KMNK ze zmiennymi sztucznymi (ang. *within estimator, least squares dummy variables estimator* (LSDV), *covariance estimator*).

Oceny efektów grupowych wyznacza się z równania (2.4), zastępując wektor **β** wektorem ocen uzyskanych na podstawie wzoru (2.9):

$$\hat{\alpha}_i = \bar{y}_{i.} - \tilde{\alpha} - \bar{\mathbf{x}}_{i.}^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{\mathrm{W}}\,^{6} \tag{2.10}$$

Należy podkreślić, że w przypadkach, gdy N jest bardzo duże, tzn. próba zawiera informacje o bardzo wielu obiektach, stosowanie wyżej omówionych metod jest dyskusyjne, ponieważ wymaga wprowadzenia do równania regresji dodatkowych N zmiennych sztucznych, co powoduje znaczną utratę stopni swobody. Co więcej, jeśli jednocześnie T nie wzrasta, to estymator efektów grupowych $\hat{\alpha}_i$ traci zgodność. Konstrukcja modelu ze zmiennymi sztucznymi uniemożliwia ponadto oszacowanie wpływu wywieranego na zmienną y przez zmienne stałe w czasie, np. płeć lub wykształcenie. Jeśli bowiem zmienne te są uwzględnione w modelu jako zmienne objaśniające, to są z niego usuwane

---

[6] Niekiedy przyjmuje się, że wyjściowy model (2.2) zawiera wspólną stałą φ. Wyraz wolny jest równy $\varphi + \alpha_i$. Ocenę tej stałej wyznacza się wtedy z równania (2.5) jako: $\hat{\varphi} = \bar{y}_{..} - \bar{\mathbf{x}}_{..}^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{\mathrm{W}}$. Rezultaty statystyczne są w obu przypadkach identyczne, jednakże interpretacja parametrów przy zmiennych sztucznych jest inna. Różnice te dyskutowane są szczegółowo w pracy S u i t s a [1984].

podczas transformacji wewnątrzgrupowej (2.8). Problemów tych można uniknąć, jeśli założy się, że α_i są zmiennymi losowymi. Przyjęcie tego założenia prowadzi do modeli z dekompozycją składnika losowego.

2.1.2. Testowanie istotności efektów grupowych

Ponieważ efekty grupowe w modelach FE są szacowane jako parametry przy zmiennych zero-jedynkowych, najprostszym sposobem sprawdzenia, czy ich wyodrębnianie jest uzasadnione, jest przeprowadzenie testu *F* (Chowa).

Hipoteza zerowa głosi, że łączny wpływ zmiennych sztucznych jest statystycznie nieistotny, a więc, że wyodrębnianie efektów grupowych jest nieuzasadnione:

H_0: $\alpha_1 = \alpha_2 = ... = \alpha_N = 0$, zaś

H_1: $\alpha_1 \neq 0$, lub $\alpha_2 \neq 0$, lub ... $\alpha_N \neq 0$.

Statystyką służącą do weryfikacji prawdziwości tych hipotez jest:

$$F_0 = \frac{(RRSS - URSS)/(N-1)}{URSS/(NT - N - K)} \tag{2.11}$$

gdzie:

RRSS – suma kwadratów reszt tzw. modelu *pooled*, tzn. uzyskana w wyniku zastosowania KMNK do modelu bez wyróżnionych efektów grupowych,

URSS – suma kwadratów reszt uzyskana przy zastosowaniu estymatora wewnątrzgrupowego.

Statystyka F_0 ma rozkład *F* o (N –1) i (NT – N – K) stopniach swobody. Jeśli jej wartość przewyższa odpowiednią wartość krytyczną, to hipoteza zerowa jest odrzucana.

2.2. Modele panelowe z dekompozycją składnika losowego (modele RE)

W modelach FE efekty grupowe α_i traktowane są jako parametry przy zmiennych sztucznych, podlegające estymacji. Alternatywnie można przyjąć założenie, że są one zmiennymi losowymi. Pozwala to uniknąć problemów związanych ze zbyt dużą liczbą analizowanych obiektów, o których wspomniano poprzednio. Możliwe jest też wówczas uogólnienie wnioskowania na całą

populację, z której pochodzą obiekty (może to mieć miejsce np. w badaniach budżetów gospodarstw domowych).

2.2.1. Założenia konstrukcyjne i metody estymacji

Rozważmy model postaci:

$$y_{it} = \alpha + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + (\alpha_i + \varepsilon_{it}) = \alpha + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + u_{it}, \quad i = 1, ..., N, \ t = 1, ..., T \qquad (2.12)$$

o którym zakłada się, że:

– efekty grupowe α_i są zmiennymi losowymi o wartości oczekiwanej 0 i jednakowej dla każdego i wariancji σ_α^2;

– efekty α_i oraz składowa ε_{it} są niezależne od siebie wzajemnie i od zmiennych x_{kit};

– $\varepsilon_{it} \sim \mathrm{N}(0, \sigma_\varepsilon^2)$, $\alpha_i \sim \mathrm{N}(0, \sigma_\alpha^2)$ dla wszystkich i, t.

Składnik losowy modelu (2.12) zawiera więc dwie składowe – stały w czasie efekt grupowy α_i i „właściwy” składnik losowy ε_{it}, zmienny w czasie i po obiektach: $u_{it} = \alpha_i + \varepsilon_{it}$. Model taki nosi nazwę modelu z efektami losowymi lub modelu z dekompozycją składnika losowego (ang. *random effects model*, RE).

Model (2.12) zapisać można macierzowo w postaci:

$$\mathbf{y} = \alpha\,\mathbf{i}_{NT} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \qquad (2.13)$$

gdzie składnik losowy dekomponowany jest na dwie składowe: $\mathbf{u} = \mathbf{Z}_\alpha\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$, $\mathbf{i}_{NT}$ jest wektorem jedynkowym, a pozostałe oznaczenia są takie same, jak we wzorze (2.3). Macierz wariancji-kowariancji składnika losowego $\boldsymbol{\Omega}$ modelu (2.13) ma postać:

$$\boldsymbol{\Omega} = E(\mathbf{u}\mathbf{u}') = \sigma_\alpha^2(\mathbf{I}_N \otimes \mathbf{J}_T) + \sigma_\varepsilon^2(\mathbf{I}_N \otimes \mathbf{I}_T) \qquad (2.14)$$

Z założeń modelu (2.12) wynika, że dla każdego i, t wariancja składnika losowego u_{it} $\sigma_u^2 = \sigma_\alpha^2 + \sigma_\varepsilon^2$ jest stała, a kowariancja $\mathrm{cov}(u_{it}, u_{js}) = \sigma_\alpha^2$ dla $i = j, t \neq s$ i zero, w pozostałych przypadkach; oznacza to, że występuje korelacja w czasie między składnikami losowymi dotyczącymi tych samych obiektów. Macierz $\boldsymbol{\Omega}$ jest więc macierzą blokowo-diagonalną.

Dla przykładu, jeśli $N = 3$, $T = 2$, budowa macierzy $\boldsymbol{\Omega}$ jest następująca:

$$\mathbf{\Omega} = \begin{bmatrix} \sigma_\alpha^2+\sigma_\varepsilon^2 & \sigma_\alpha^2 & 0 & 0 & 0 & 0 \\ \sigma_\alpha^2 & \sigma_\alpha^2+\sigma_\varepsilon^2 & \sigma_\alpha^2 & 0 & 0 & 0 \\ 0 & \sigma_\alpha^2 & \sigma_\alpha^2+\sigma_\varepsilon^2 & \sigma_\alpha^2 & 0 & 0 \\ 0 & 0 & \sigma_\alpha^2 & \sigma_\alpha^2+\sigma_\varepsilon^2 & \sigma_\alpha^2 & 0 \\ 0 & 0 & 0 & \sigma_\alpha^2 & \sigma_\alpha^2+\sigma_\varepsilon^2 & \sigma_\alpha^2 \\ 0 & 0 & 0 & 0 & \sigma_\alpha^2 & \sigma_\alpha^2+\sigma_\varepsilon^2 \end{bmatrix},$$

gdyż:

$$\sigma_\alpha^2(\mathbf{I}_3 \otimes \mathbf{J}_2) = \sigma_\mu^2\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right) \text{ oraz } \sigma_\varepsilon^2(\mathbf{I}_3 \otimes \mathbf{I}_2) = \sigma_\alpha^2\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right).$$

Z przedstawionych powyżej założeń wynika, że do estymacji parametrów modelu (2.13) można zastosować uogólnioną metodę najmniejszych kwadratów (UMNK). Wymagałoby to jednak odwrócenia macierzy $\mathbf{\Omega}$, która jest macierzą stopnia *NT*, i z reguły ma na tyle duże wymiary, że jej odwrócenie jest utrudnione, nawet przy użyciu komputera[7]. Możliwych jest kilka sposobów postępowania. Jeden z nich zaproponowany został przez W a n s b e e k a i K a p t e y n a [1982]. Dowiedli oni, że macierz $\mathbf{\Omega}$ można zapisać w postaci[8]:

$$\mathbf{\Omega} = \sigma_1^2\mathbf{P} + \sigma_\varepsilon^2\mathbf{Q} \tag{2.15}$$

gdzie:

$\sigma_1^2 = T\sigma_\alpha^2 + \sigma_\varepsilon^2$, $\mathbf{P} = \mathbf{I}_N \otimes \frac{1}{T}\mathbf{J}_T$ jest tzw. operatorem międzygrupowym, tzn. operatorem, którego pomnożenie przez dowolną macierz daje macierz średnich grupowych. Ponadto $\mathbf{P}$ jest macierzą symetryczną i idempotentną, macierze $\mathbf{P}$ i $\mathbf{Q}$ są ortogonalne[9] (tzn. $\mathbf{PQ} = \mathbf{QP} = \mathbf{0}$), $\mathbf{P} + \mathbf{Q} = \mathbf{I}_{NT}$; oraz $\mathbf{yy}^\mathrm{T} = \mathbf{y}^\mathrm{T}\mathbf{Py} + \mathbf{y}^\mathrm{T}\mathbf{Qy}$, a więc całkowitą wariancję zmiennej *y* można przedstawić jako sumę wariancji międzygrupowej i wewnątrzgrupowej. Pomnożenie dowolnego wektora $\mathbf{y}$ przez $\sigma_\varepsilon\mathbf{\Omega}^{-1/2}$, gdzie:

$$\mathbf{\Omega}^{-1/2} = \mathbf{P}/\sigma_1 + \mathbf{Q}/\sigma_\varepsilon \tag{2.16}$$

---

[7] Jeśli $N = 20$, a $T = 5$, to $\mathbf{\Omega}$ jest macierzą o wymiarach (100 × 100).

[8] Ponadto zachodzi równość: $\mathbf{\Omega}^{-1} = \mathbf{P}/\sigma_1^2 + \mathbf{Q}/\sigma_\varepsilon^2$.

[9] $\mathbf{Q}$ jest operatorem wewnątrzgrupowym, określonym w pkt 2.1.1.

daje wektor $\mathbf{y}^*$, którego typowym elementem jest $y_{it} - \theta\,\bar{y}_i$, gdzie $\theta = 1 - (\sigma_\varepsilon / \sigma_1)$. Wansbeek i Kapteyn udowodnili, że model, otrzymany przez pomnożenie stronami równania (2.13) przez $\sigma_\varepsilon \mathbf{\Omega}^{-1/2}$, a więc model postaci:

$$\sigma_\varepsilon \mathbf{\Omega}^{-1/2}\mathbf{y} = \sigma_\varepsilon \mathbf{\Omega}^{-1/2}\alpha\,\mathbf{i}_{NT} + \sigma_\varepsilon \mathbf{\Omega}^{-1/2}\mathbf{X}\boldsymbol{\beta} + \sigma_\varepsilon \mathbf{\Omega}^{-1/2}\mathbf{u} \qquad (2.17)$$

można estymować za pomocą KMNK, zatem ich metoda pozwala uniknąć bezpośredniego stosowania UMNK. Nieznane składowe wariancji σ_ε^2 i σ_1^2 występujące we wzorze (2.15) zastępuje się ich nieobciążonymi oszacowaniami:

$$\begin{aligned} \hat{\sigma}_1^2 &= \frac{\mathbf{u}^{\mathrm{T}}\mathbf{P}\mathbf{u}}{\mathrm{Tr}(\mathbf{P})} = T\sum_{i=1}^{N}\bar{e}_{i.}^2 / N, \\ \hat{\sigma}_\varepsilon^2 &= \frac{\mathbf{u}^{\mathrm{T}}\mathbf{Q}\mathbf{u}}{\mathrm{tr}(\mathbf{Q})} = \frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=1}^{T}(e_{it} - \bar{e}_{i.})^2 \end{aligned} \qquad (2.18)$$

gdzie e_{it} oznaczają reszty modelu (2.13) pomnożonego stronami przez $\sigma_\varepsilon \mathbf{\Omega}^{-1/2}$. T. A m e m i y a [1971] zasugerował wykorzystanie reszt KMNK ze zmiennymi sztucznymi.

Innym sposobem oszacowania parametrów modelu RE jest utworzenie układu równań:

$$\begin{pmatrix} \mathbf{Q}\mathbf{y} \\ \mathbf{P}\mathbf{y} \end{pmatrix} = \begin{pmatrix} \mathbf{Q}[\mathbf{i}_{NT} & \mathbf{X}] \\ \mathbf{P}[\mathbf{i}_{NT} & \mathbf{X}] \end{pmatrix}\begin{bmatrix} \alpha \\ \boldsymbol{\beta} \end{bmatrix} + \begin{pmatrix} \mathbf{Q}\mathbf{u} \\ \mathbf{P}\mathbf{u} \end{pmatrix} \qquad (2.19)$$

Składnik losowy modelu (2.19) ma wartość oczekiwaną równą **0** i macierz wariancji-kowariancji daną wzorem:

$$\begin{pmatrix} \sigma_\varepsilon^2\mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \sigma_1^2\mathbf{P} \end{pmatrix} \qquad (2.20)$$

Zastosowanie UMNK do układu 2*NT* równań (2.19) jest równoważne zastosowaniu tej metody do modelu (2.13), a otrzymany w ten sposób estymator wektora $\boldsymbol{\beta}$ ma postać:

$$\hat{\boldsymbol{\beta}}_{\mathrm{UMNK}} = \left[\mathbf{W}_{XX} + \phi^2\mathbf{B}_{XX}\right]^{-1}\left[\mathbf{W}_{Xy} + \phi^2\mathbf{B}_{Xy}\right] \qquad (2.21)$$

gdzie:

$$\mathbf{W}_{XX} = \mathbf{X}'\mathbf{Q}\mathbf{X}, \quad \mathbf{B}_{XX} = \mathbf{X}'\left(\mathbf{P} - \bar{\mathbf{J}}_{NT}\right)\mathbf{X}, \quad \mathbf{B}_{Xy} = \mathbf{X}^{\mathrm{T}}\left(\mathbf{P} - \bar{\mathbf{J}}_{NT}\right)\mathbf{y}, \quad \phi^2 = \sigma_\varepsilon^2 / \sigma_1^2.$$

Wariancja estymatora (2.21) jest równa:

$$\mathrm{D}^2\left(\hat{\boldsymbol{\beta}}_{\mathrm{UMNK}}\right) = \sigma_\varepsilon^2\left(\mathbf{W}_{XX} + \phi^2\mathbf{B}_{XX}\right)^{-1} \tag{2.22}$$

Nieznane oceny składowych wariancji σ_ε^2 i σ_1^2 otrzymać można, jak poprzednio, na podstawie reszt KMNK ze zmiennymi sztucznymi.

Ponadto można wykazać, że estymator (2.21) jest ważoną średnią estymatora wewnątrzgrupowego $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$, danego wzorem (2.9), i estymatora międzygrupowego (ang. *between estimator*), danego wzorem

$$\hat{\boldsymbol{\beta}}_{\mathrm{B}} = \mathbf{B}_{XX}^{-1}\mathbf{B}_{Xy} \tag{2.23}$$

otrzymywanego przy zastosowaniu KMNK do modelu (2.4).

W literaturze przedmiotu podkreśla się, że oceny parametrów uzyskane na podstawie estymatora międzygrupowego, który wykorzystuje informacje o charakterze przekrojowym, są ocenami efektów długookresowych, a oceny na podstawie estymatora wewnątrzgrupowego, wykorzystującego informacje o charakterze czasowym – ocenami efektów krótkookresowych.

2.2.2. Testowanie istotności efektów grupowych

Jednym z możliwych do zastosowania testów, badających, czy zasadne jest wyodrębnainie efektów grupowych w modelu RE jest test mnożnika Lagrange'a, zaproponowany przez B r e u s c h a i P a g a n a [1980]. Test ten służy do badania czy model z dekompozycją składnika losowego jest statystycznie lepszy niż model, w którym nie wyróżnia się efektów grupowych (model *pooled*). Hipoteza zerowa głosi, że wyodrębnianie części składnika losowego, specyficznej dla obiektów, jest nieuzasadnione. Ma ona postać:

H_0: $\sigma_{\mu_i}^2 = 0,$

wobec:

H_1: $\sigma_{\mu_i}^2 \neq 0.$

Do weryfikacji prawdziwości hipotezy zerowej wykorzystuje się statystykę:

$$LM = \frac{NT}{2(T-1)}\left[\frac{\sum_i \left(\sum_t e_{it}\right)^2}{\sum_i \sum_t e_{it}^2} - 1\right]^2 \tag{2.24}$$

gdzie e_{it} są resztami modelu, w którym nie wyróżnia się efektów grupowych.

Statystyka *LM* ma rozkład χ^2 z jednym stopniem swobody. Hipoteza zerowa jest odrzucana, a więc efekty grupowe uznaje się za istotne, jeśli wartość statystyki *LM* przewyższa odpowiednią wartość krytyczną.

2.3. Model FE czy model RE

W podrozdz. 2.1 i 2.2 przedstawiono dwa podstawowe typy modeli panelowych, różniące się założeniami, przyjmowanymi odnośnie do natury efektów grupowych. W naturalny sposób pojawia się pytanie o sposób dokonania wyboru pomiędzy nimi, o ile odpowiednie testy wykazały występowanie specyfiki grupowej. Obecnie omówione zostaną ogólne, intuicyjne zalecenia, dotyczące sytuacji, w których każdy z modeli jest preferowany, jak również test statystyczny, porównujący własności estymatorów $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ i $\hat{\boldsymbol{\beta}}_{\mathrm{UMNK}}$.

2.3.1. Wskazówki ogólne

W praktycznych zastosowaniach pojawia się problem wyboru jednego ze wspomnianych wyżej modeli. Rozstrzygnięcia dokonać można kierując się jednym z kilku kryteriów.

- W zależności od relacji między liczbą obiektów i liczbą okresów: jeśli *T* jest stosunkowo duże, a *N* niewielkie, to możliwe jest wprowadzenie do modelu *N* zmiennych sztucznych odpowiadających poszczególnym obiektom, czyli zastosowanie modelu FE. Jeśli *T* jest bardzo małe (a *N* bardzo duże), to duża liczba zmiennych sztucznych może spowodować utratę zgodności estymatora. Wówczas zalecane jest uwzględnienie różnic między obiektami poprzez różnicowanie części składnika losowego, a więc wybór modelu RE.
- W zależności od natury analizowanych obiektów: jeśli bada się obiekty tego samego rodzaju (państwa, bardzo duże firmy, gałęzie przemysłu) i istotne

jest oszacowanie efektów grupowych dla tych konkretnych obiektów, to właściwszy jest model FE. Rozkład zmiennej y rozważany jest wtedy warunkowo, przy ustalonych wartościach α_i. Jeśli natomiast rozważane są obiekty losowo wybrane z pewnej populacji (gospodarstwa domowe, małe firmy), oszacowanie konkretnych efektów grupowych dla tych obiektów jest mniej istotne, gdyż badanie zazwyczaj koncentruje się na charakterystykach całej populacji. Ich oszacowanie umożliwia model RE.

- Zbadanie, czy efekty grupowe nie są skorelowane ze zmiennymi objaśniającymi. Występowanie takiej korelacji oznacza, że również składnik losowy u_{it}, którego efekt α_i jest składową, jest skorelowany ze zmienną objaśniającą. Estymator $\hat{\boldsymbol{\beta}}_{\mathrm{UMNK}}$ traci więc zgodność. Estymator $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ ma natomiast pożądane własności, gdyż przekształcenie wewnątrzgrupowe usuwa z modelu stałe w czasie efekty α_i. Dla sprawdzenia, czy wspomniana korelacja występuje, można zastosować test Hausmana.

2.3.2. Test Hausmana

Omawiany test jest najpopularniejszym testem statystycznym, stosowanym do rozstrzygania o tym, który z estymatorów: $\hat{\boldsymbol{\beta}}_{\mathrm{UMNK}}$ czy $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ posiada własności dobrego estymatora, a więc do dokonania wyboru pomiędzy modelem RE a modelem FE. Test został zaproponowany przez J. A. H a u s m a n a [1978].

Idea konstrukcji tego testu opiera się na porównaniu własności dwóch estymatorów, z których jeden jest zgodny, niezależnie od tego, która z hipotez jest prawdziwa, a drugi – tylko jeśli prawdziwa jest H_0. Hipoteza zerowa zakłada, że efekty grupowe α_i są nieskorelowane ze zmiennymi objaśniającymi:

H_0: $\mathrm{E}(u_{it} \mid \mathbf{x}_{it}) = \mathbf{0}$,

wobec

H_1: $\mathrm{E}(u_{it} \mid \mathbf{x}_{it}) \neq \mathbf{0}$.

Przy założeniu prawdziwości H_0, oba estymatory: $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ i $\hat{\boldsymbol{\beta}}_{\mathrm{UMNK}}$ są zgodne, ale $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ jest nieefektywny. Jeśli H_0 jest nieprawdziwa, to $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ jest zgodny, gdyż przekształcenie wewnątrzgrupowe $\mathbf{Q}$ usuwa efekty grupowe, będące przyczyną korelacji; estymator $\hat{\boldsymbol{\beta}}_{\mathrm{UMNK}}$ jest natomiast niezgodny. Innymi słowy, hipoteza zerowa głosi, że poprawną specyfikacją jest model RE, a hipoteza alternatywna – że poprawna jest specyfikacja modelu FE.

Konstrukcja statystyki empirycznej testu Hausmana opiera się na tym, że jeśli H_0 jest niesłuszna, to estymatory: $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ i $\hat{\boldsymbol{\beta}}_{\mathrm{UMNK}}$ mają różne granice stochastyczne. Statystyka weryfikująca hipotezę zerową ma postać:

$$m = \hat{\mathbf{q}}_1^{\mathrm{T}} \left[\mathrm{D}^2(\hat{\mathbf{q}}_1) \right]^{-1} \hat{\mathbf{q}}_1 \tag{2.25}$$

gdzie $\hat{\mathbf{q}}_1 = \hat{\boldsymbol{\beta}}_{\mathrm{UMNK}} - \hat{\boldsymbol{\beta}}_{\mathrm{W}}$.

Statystyka m ma rozkład χ^2 z K stopniami swobody, gdzie K oznacza liczbę zmiennych objaśniających modelu. Jeśli $m > \chi^2_{\mathrm{kryt}}$, to odrzucamy hipotezę zerową na rzecz H_1. W przeciwnym przypadku nie ma podstaw do odrzucenia H_0.

Praktycznym problemem, jaki może się pojawić przy obliczaniu statystyki m, jest to, że macierz $\left[\mathrm{D}^2(\hat{\mathbf{q}}_1) \right]$ może, w skończonych próbach, nie być dodatnio określona, a więc może nie być możliwe jej odwrócenie.

2.4. Model Hausmana i Taylora

O zmiennych objaśniających w modelach FE, RE zakłada się, że ich wartości zmieniają się w czasie i po obiektach. H a u s m a n i T a y l o r [1981] opracowali metodę estymacji modeli, w których występują stałe w czasie zmienne objaśniające. Opiera się ona na metodzie zmiennych instrumentalnych (IV). Zasadniczą ideę metody Hausmana i Taylora zaprezentować można na przykładzie modelu z czterema zmiennymi objaśniającymi, postaci:

$$y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + \phi_1 z_{1i} + \phi_2 z_{2i} + \alpha_i + \varepsilon_{it} \tag{2.26}$$

w którym wszystkie zmienne objaśniające są ściśle egzogeniczne w sensie braku korelacji ze składnikiem losowym ε_{it}: $\mathrm{E}(x_{kit}\varepsilon_{it}) = \mathrm{E}(z_{kt}\varepsilon_{it}) = 0$, $k = 1, 2$. Zmienne z_{1i} i x_{1it} są ponadto nieskorelowane z efektami grupowymi α_i, zaś zmienne z_{2i} i x_{2it} są z nimi skorelowane.

Estymacja parametrów ϕ_1 i β_1 za pomocą IV nie nastręcza problemów, ponieważ zmienne z_{1i} i x_{1it} są egzogeniczne, mogą być instrumentami dla samych siebie. Parametr β_2 można oszacować, zastępując zmienną x_{2it} zmienną instrumentalną $(x_{2it} - \bar{x}_{2i.})$, gdzie $\bar{x}_{2i.}$ oznacza średnią po czasie policzoną z wartości x_{2it} dla $t = 1, ..., T$. Zmienna ta jest nieskorelowana z efektami grupowymi. W celu znalezienia odpowiedniego instrumentu dla zmiennej z_{2i} zauważyć należy, że korelacja tej zmiennej z α_i jest związana ze zmiennością międzygrupową, a zatem odpowiedni instrument powinien również wykazywać zmienność międzygrupową, być stały w czasie i nieskorelowany z α_i. Zmienną spełniającą te warunki jest średnia $\bar{x}_{1i.}$

Powyżej przedstawione rozumowanie można uogólnić i zapisać bardziej formalnie. Model rozważany przez H a u s m a n a i T a y l o r a [1981] ma postać:

$$y_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\varphi} + \alpha_i + \varepsilon_{it} \tag{2.27}$$

gdzie $\mathbf{Z}_i$ jest wektorem zmiennych objaśniających, stałych w czasie. $\mathbf{X}_{it}$ i $\mathbf{Z}_i$ dekomponowane są następnie na dwa zbiory:

$$\mathbf{X}_{it} \equiv \mathbf{X} = [\mathbf{X}_1 \vdots \mathbf{X}_2] \text{ oraz } \mathbf{Z}_i \equiv \mathbf{Z} = [\mathbf{Z}_1 \vdots \mathbf{Z}_2],$$

gdzie wymiary macierzy są następujące: $\mathbf{X}_1$ ma wymiary ($NT \times k_1$), $\mathbf{X}_2$ – ($NT \times k_2$), $\mathbf{Z}_1$ – ($NT \times g_1$), $\mathbf{Z}_2$ – ($NT \times g_2$). Zakłada się, że $\mathbf{X}_1$ i $\mathbf{Z}_1$ są egzogeniczne, tzn. nieskorelowane z ε_{it} ani z α_i, zaś $\mathbf{X}_2$ i $\mathbf{Z}_2$ – endogeniczne, ze względu na korelację z efektami α_i (nie są one natomiast skorelowane ze składnikiem losowym ε_{it}). Oznacza to połączenie założeń modelu FE (dla zmiennych $\mathbf{X}_2$ i $\mathbf{Z}_2$) i modelu RE (dla zmiennych $\mathbf{X}_1$ i $\mathbf{Z}_1$), dotyczących korelacji między efektami grupowymi a zmiennymi objaśniającymi.

O modelu (2.27) zakłada się też, że macierz $\boldsymbol{\Omega}$ wariancji-kowariancji składnika losowego $\mathbf{u} = \boldsymbol{\alpha} + \boldsymbol{\varepsilon}$ modelu (2.27) ma strukturę taką, jak w modelu RE, tzn.:

$$\boldsymbol{\Omega} = \mathrm{E}(\mathbf{uu}') = \sigma_\alpha^2(\mathbf{I}_N \otimes \mathbf{J}_T) + \sigma_\varepsilon^2(\mathbf{I}_N \otimes \mathbf{I}_T) \tag{2.28}$$

gdzie σ_α^2 i σ_ε^2 są wariancjami poszczególnych składowych.

Przekształcenie wewnątrzgrupowe $\mathbf{Q}$ usunęłoby z modelu (2.27) zarówno efekty grupowe, jak i zmienne $\mathbf{Z}_i$, zatem jego zastosowanie, oprócz usunięcia źródła obciążenia estymatora, skutkowałoby brakiem możliwości oszacowania parametrów $\boldsymbol{\varphi}$. Hausman i Taylor [1981] zaproponowali zatem pomnożenie modelu (2.27) przez macierz $\boldsymbol{\Omega}^{-1/2}$, określoną wzorem (2.16)[10]. Tak przekształcony model ma postać:

$$\boldsymbol{\Omega}^{-1/2} y_{it} = \boldsymbol{\Omega}^{-1/2}\mathbf{X}_{it}\boldsymbol{\beta} + \boldsymbol{\Omega}^{-1/2}\mathbf{Z}_i\boldsymbol{\varphi} + \boldsymbol{\Omega}^{-1/2}(\alpha_i + \varepsilon_{it}) \tag{2.29}$$

Powyższe przekształcenie nazywane jest w literaturze anglojęzycznej *quasi time-demeaning* (np. Wooldridge [2002], s. 287), gdyż polega ono na odejmowaniu od każdej obserwacji, zarówno na zmiennej objaśnianej, jak i na zmiennych objaśniających, ułamka średniej po czasie.

Nieznane elementy macierzy $\boldsymbol{\Omega}^{-1/2}$ we wzorze (2.29) można oszacować wykorzystując oceny $\hat{\sigma}_\alpha^2$ i $\hat{\sigma}_\varepsilon^2$, uzyskane na podstawie regresji wewnątrzgru-

[10] Zachodzi twierdzenie mówiące, że określona w powyższy sposób macierz nieosobliwa $\boldsymbol{\Omega}^{-1/2}$ przekształca macierz $\boldsymbol{\Omega}$ w macierz skalarną – por. Hausman, Taylor [1981], Proposition 2.1, s. 1381.

powej w następujący sposób. Niech $\hat{\mathbf{d}}_i$ będzie wektorem średnich grupowych o wymiarach ($NT \times 1$), oszacowanym na podstawie estymatora wewnątrzgrupowego $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$, tzn.:

$$\hat{\mathbf{d}}_i = y_{i\cdot} - \mathbf{X}_{i\cdot}\hat{\boldsymbol{\beta}}_{\mathrm{W}} = \mathbf{Z}_i\boldsymbol{\varphi} + \boldsymbol{\alpha}_i + \left[\mathbf{P} - \mathbf{X}_{i\cdot}\left(\tilde{\mathbf{X}}_{it}^{\mathrm{T}}\tilde{\mathbf{X}}_{it}\right)^{-1}\tilde{\mathbf{X}}_{it}^{\mathrm{T}}\right]\varepsilon_{it} \tag{2.30}$$

gdzie $\tilde{\mathbf{X}}_{it} = \mathbf{Q}\mathbf{X}_{it}$. Zakładając, że ostatni człon wyrażenia (2.30) jest składnikiem losowym o zerowej wartości oczekiwanej, możliwe jest uzyskanie zgodnego estymatora tego modelu za pomocą dwustopniowej metody najmniejszych kwadratów (2MNK). Warunkiem jest, aby $k_1 \geq g_2$, a więc liczba zmiennych egzogenicznych zmieniających się w czasie była co najmniej równa liczbie stałych w czasie zmiennych skorelowanych z efektami grupowymi α_i. Estymator wektora $\boldsymbol{\varphi}$ ma postać:

$$\hat{\boldsymbol{\varphi}}_{2\mathrm{MNK}} = (\mathbf{Z}_i^{\mathrm{T}}\mathbf{P}_{\mathrm{A}}\mathbf{Z}_i)^{-1}\mathbf{Z}_i^{\mathrm{T}}\mathbf{P}_{\mathrm{A}}\hat{\mathbf{d}}_i \tag{2.31}$$

gdzie $\mathbf{P}_{\mathrm{A}} = \mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}$, zaś macierz instrumentów $\mathbf{A}$ ma postać $\mathbf{A} = [\mathbf{X}_1 \vdots \mathbf{Z}_1]$. H a u s m a n i T a y l o r [1981] wykazali, że estymator (2.31) jest zgodny, lecz nieefektywny. Zgodne estymatory składowych wariancji można otrzymać na podstawie $\hat{\boldsymbol{\varphi}}_{2\mathrm{MNK}}$ i $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ jako:

$$\begin{aligned} \hat{\sigma}_\varepsilon^2 &= \frac{\tilde{\mathbf{y}}^{\mathrm{T}}\mathbf{P}_{\tilde{\mathbf{X}}}^{'}\tilde{\mathbf{y}}}{N(T-1)}, \\ \hat{\sigma}_\alpha^2 &= \frac{\left(\mathbf{y}_{i\cdot} - \mathbf{X}_{i\cdot}\hat{\boldsymbol{\beta}}_{\mathrm{W}} - \mathbf{Z}_i\hat{\boldsymbol{\varphi}}_{2\mathrm{MNK}}\right)^{-1}\left(\mathbf{y}_{i\cdot} - \mathbf{X}_{i\cdot}\hat{\boldsymbol{\beta}}_{\mathrm{W}} - \mathbf{Z}_i\hat{\boldsymbol{\varphi}}_{2\mathrm{MNK}}\right)}{N} - \frac{\hat{\sigma}_\varepsilon^2}{T} \end{aligned} \tag{2.32}$$

gdzie $\tilde{\mathbf{y}} = \mathbf{Q}\mathbf{y}$, $\tilde{\mathbf{X}} = \mathbf{Q}\mathbf{X}$, $\mathbf{P}_{\tilde{\mathrm{X}}}^{'} = \mathbf{I} - \mathbf{P}_{\mathrm{X}}$, $\mathbf{P}_{\mathrm{X}}$ określone jest analogicznie jak $\mathbf{P}_{\mathrm{A}}$.

Model postaci (2.29) należy następnie oszacować za pomocą IV, ze zbiorem instrumentów określonych jako: $\mathbf{A}_0 = [\mathbf{Q}, \mathbf{X}_1, \mathbf{Z}_1]$ lub, równoważnie[11]: $\mathbf{A}_{\mathrm{HT}} = [\mathbf{Q}\mathbf{X}_1, \mathbf{Q}\mathbf{X}_2, \mathbf{P}\mathbf{X}_1, \mathbf{Z}_1]$. Zmiennymi instrumentalnymi dla $\mathbf{X}_1$ i $\mathbf{X}_2$ są więc odchylenia wartości tych zmiennych od średnich po czasie, zmiennymi instrumentalnymi dla $\mathbf{Z}_2$ – średnie po czasie, a zmienne $\mathbf{Z}_1$ nie wymagają instrumen-

[11] Zbiór instrumentów $\mathbf{A}_{HT}$ zaproponowany został przez B r e u s c h a, M i z o n a i S c h m i d t a [1989].

tów. Elementy macierzy $\mathbf{A}_{\mathrm{HT}}$ mogą być instrumentami, gdyż z przyjętych założeń wynikają następujące własności:

$\mathbf{QZ}_i = \mathbf{0}$ z konstrukcji $\mathbf{Q}$, gdyż $\mathbf{Z}_i$ jest stałe w czasie,

$\mathrm{E}[(\mathbf{QX})^{\mathrm{T}}(\boldsymbol{\alpha}+\boldsymbol{\varepsilon})] = \mathbf{0}$, gdyż zmienne x_{it} są nieskorelowane z ε_{it},

a $\mathbf{Q}(\alpha_i + \varepsilon_{it}) = \mathbf{Q}\varepsilon_{it}$, $\mathrm{E}[(\mathbf{Z}_1^{\mathrm{T}}(\boldsymbol{\alpha}+\boldsymbol{\varepsilon})] = \mathbf{0}$.

Uzyskany w ten sposób estymator 2MNK wyraża się wzorem:

$$\begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\varphi}} \end{pmatrix}_{\mathrm{HT}} = \left[\begin{pmatrix} \mathbf{X}_*^{\mathrm{T}} \\ \mathbf{Z}_*^{\mathrm{T}} \end{pmatrix} \mathbf{P}_{\mathbf{A}} \left(\mathbf{X}_*, \mathbf{Z}_* \right) \right]^{-1} \begin{pmatrix} \mathbf{X}_*^{\mathrm{T}} \\ \mathbf{Z}_*^{\mathrm{T}} \end{pmatrix} \mathbf{P}_{\mathbf{A}} \mathbf{y}_* \tag{2.33}$$

gdzie $\mathbf{P}_{\mathrm{AHT}}$ jest macierzą projekcji ortogonalnej na $\mathbf{A}_{\mathrm{HT}}$, określoną analogicznie do $\mathbf{P}_{\mathrm{A}}$.

Estymator (2.33), nazywany też estymatorem uogólnionej metody zmiennych instrumentalnych (GIV), jest też efektywnym estymatorem uogólnionej metody momentów (GMM). Warto podkreślić, że zmienne $\mathbf{X}_1$ są wykorzystywane jako instrumenty dwukrotnie: raz jako odchylenia od średnich grupowych $(\mathbf{QX}_1)$, i jako średnie grupowe $(\mathbf{PX}_1)$. Jest to możliwe dzięki wykorzystaniu danych panelowych.

Zgodnie z warunkiem rzędu, dla identyfikacji rozważanego modelu musi zachodzić warunek $k_1 \geq g_2$, a więc liczba zmiennych egzogenicznych zmieniających się w czasie musi być co najmniej równa liczbie stałych w czasie zmiennych skorelowanych z efektami grupowymi α_i. Dokładniej, możliwe są następujące przypadki:

– jeśli $k_1 < g_2$, to model jest nieidentyfikowalny. Wówczas $\hat{\boldsymbol{\beta}}_{\mathrm{HT}} = \hat{\boldsymbol{\beta}}_{\mathrm{W}}$, a oszacowanie wektora $\boldsymbol{\varphi}$ nie jest możliwe;

– jeśli $k_1 = g_2$, to model jest jednoznacznie identyfikowalny. Wówczas $\hat{\boldsymbol{\beta}}_{\mathrm{HT}} = \hat{\boldsymbol{\beta}}_{\mathrm{W}}$, a wektor $\boldsymbol{\varphi}$ szacuje się za pomocą estymatora $\hat{\boldsymbol{\varphi}}_{\mathrm{2MNK}}$, danego wzorem (2.31);

– jeśli $k_1 > g_2$, to model jest identyfikowalny z nadmiarem i estymator dany wzorem (2.33) jest efektywniejszy niż estymator wewnątrzgrupowy.

H a u s m a n i T a y l o r [1981] zaproponowali też test dla warunków identyfikujących. Badanie identyfikowalności modelu jest ważne, gdyż, jak wynika z powyższego opisu, istnienie estymatorów (2.33), a więc istnienie efektywniejszego niż wewnątrzgrupowy estymatora wektora $\boldsymbol{\beta}$ i zgodnego estymatora wektora $\boldsymbol{\varphi}$, zależy od tego, czy model jest identyfikowalny z nadmiarem. Przy założeniu prawdziwości hipotezy zerowej, postaci:

$$\mathrm{H}_0 : \underset{N\to\infty}{p\lim} \frac{1}{N} \mathbf{X}_1^{\mathrm{T}} \boldsymbol{\alpha}_i = \mathbf{0} \ \text{ i } \ \underset{N\to\infty}{p\lim} \frac{1}{N} \mathbf{Z}_1^{\mathrm{T}} \boldsymbol{\alpha}_i = \mathbf{0},$$

oba estymatory: $\hat{\boldsymbol{\beta}}_{\mathrm{HT}}$ i $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ są zgodne. Przy tym, jak wspomniano powyżej, $\hat{\boldsymbol{\beta}}_{\mathrm{HT}}$ jest efektywny. Jeśli hipoteza ta jest nieprawdziwa, to estymator $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ jest zgodny, a $\hat{\boldsymbol{\beta}}_{\mathrm{HT}}$ – niezgodny.

Statystyka empiryczna ma postać[12]:

$$\hat{m} = \left(\hat{\boldsymbol{\beta}}_{\mathrm{HT}} - \hat{\boldsymbol{\beta}}_{\mathrm{W}}\right)^{\mathrm{T}} \left[\mathrm{D}^2(\hat{\boldsymbol{\beta}}_{\mathrm{W}}) - \mathrm{D}^2(\hat{\boldsymbol{\beta}}_{\mathrm{HT}})\right]^{-1} \left(\hat{\boldsymbol{\beta}}_{\mathrm{HT}} - \hat{\boldsymbol{\beta}}_{\mathrm{W}}\right) \tag{2.34}$$

Statystyka $\hat{\sigma}_{\varepsilon}^2 \hat{m}$ ma rozkład χ^2 z l stopniami swobody, gdzie $l = \min\{k_1 - g_2, NT - K\}$.

Proponowany test jest typu Walda. Nazywany jest w literaturze testem Hausmana, gdyż J. A. Hausman [1978] pokazał, że przy założeniu prawdziwości H_0 dla dowolnych estymatorów $\tilde{\boldsymbol{\beta}}_1$ i $\tilde{\boldsymbol{\beta}}_2$ wektora $\boldsymbol{\beta}$ macierz $[\mathrm{D}^2(\hat{\boldsymbol{\beta}}_{\mathrm{W}}) - \mathrm{D}^2(\hat{\boldsymbol{\beta}}_{\mathrm{HT}})]$ jest zgodnym estymatorem asymptotycznej wariancji $\sqrt{N}(\tilde{\boldsymbol{\beta}}_1 - \tilde{\boldsymbol{\beta}}_2)$. Może się jednak zdarzyć, że macierz $[\mathrm{D}^2(\hat{\boldsymbol{\beta}}_{\mathrm{W}}) - \mathrm{D}^2(\hat{\boldsymbol{\beta}}_{\mathrm{HT}})]$ nie jest dodatnio określona. Wówczas nie można obliczyć $\hat{m}$ na podstawie wzoru (2.34) i należy w inny sposób oszacować macierz asymptotycznej wariancji (por. Lee [2002], s. 73).

2.5. Metody estymacji modeli wykorzystujących dane czasowo-przekrojowe

Przedmiotem rozważań niniejszego podrozdziału są modele szacowane na podstawie danych, charakteryzujących się tym, że liczba obserwacji w czasie T przekracza liczbę obiektów N. Ze względu na długość wymiaru czasowego próby, modele takie mogą być estymowane metodami zbliżonymi do metod stosowanych w analizach szeregów czasowych. Uwzględnić jednak trzeba możliwość wystąpienia zjawiska heteroskedastyczności grupowej, odzwierciedlającej heterogeniczność obiektów, korelacji przekrojowej, obrazującej zależności między obiektami w tym samym okresie i autokorelacji, wynikającej z długości szeregów czasowych.

---

[12] Superskrypt –1 oznaczać może macierz odwrotną lub uogólnioną macierz odwrotną Moore'ego–Penrose'a. Obliczanie macierzy Moore'ego–Penrose'a jest konieczne, jeśli rząd: $\mathrm{rz}\left[\mathrm{D}^2(\hat{\boldsymbol{\beta}}_{\mathrm{W}}) - \mathrm{D}^2(\hat{\boldsymbol{\beta}}_{\mathrm{HT}})\right] = r < k$, gdzie k jest liczbą szacowanych parametrów, $\mathrm{D}^2(\hat{\boldsymbol{\beta}}_{\mathrm{W}})$ jest asymptotyczną wariancją macierzy $\sqrt{N}(\hat{\boldsymbol{\beta}}_{\mathrm{W}} - \boldsymbol{\beta})$, zaś $\mathrm{D}^2(\hat{\boldsymbol{\beta}}_{\mathrm{HT}})$ określa się analogicznie (por. Lee [2002]; Andrews [1987]).

2.5.1. Modele z heteroskedastycznością grupową składnika losowego

Dane czasowo-przekrojowe (ang. *time-series – cross-section data*, TSCS) charakteryzują się tym, że liczba obserwacji w czasie jest większa niż liczba obiektów. Model konstruowany na podstawie takich danych zapisać można w postaci ogólnej:

$$y_{it} = \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + \varepsilon_{it} \quad \Leftrightarrow \quad \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_N \end{bmatrix}\boldsymbol{\beta} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix} \quad \Leftrightarrow \quad \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.35}$$

gdzie dla $i = 1, ..., N$ oraz $t = 1, ..., T$, $\mathbf{y}_i$ jest wektorem obserwacji na zmiennej objaśnianej, a $\mathbf{X}_i$ – macierzą obserwacji na zmiennych objaśniających, pochodzących z i-tego obiektu, a pozostałe oznaczenia są analogiczne jak we wzorze (2.3).

Specyfikacja (2.35) oznacza w szczególności, że podobnie jak w modelach omawianych w poprzednich podrozdziałach, przyjmuje się założenie, że parametry strukturalne są identyczne dla wszystkich obserwowanych obiektów i okresów. Naturalnie występujące pomiędzy obiektami różnice uwzględnia się poprzez przyjęcie odpowiednich założeń co do struktury macierzy wariancji-kowariancji składnika losowego. W analogiczny sposób uwzględnia się zależności pomiędzy nimi. Podkreślić tu warto, że już samo połączenie obserwacji z różnych obiektów w jeden szereg jest uznaniem istnienia powiązań między nimi. Macierz wariancji-kowariancji składnika losowego równania (2.35) ma w najbardziej ogólnym przypadku postać blokową:

$$\mathbf{V} = \begin{bmatrix} \sigma_{11}\boldsymbol{\Omega}_{11} & \sigma_{12}\boldsymbol{\Omega}_{12} & \cdots & \sigma_{1N}\boldsymbol{\Omega}_{1N} \\ \sigma_{21}\boldsymbol{\Omega}_{21} & \sigma_{22}\boldsymbol{\Omega}_{22} & \cdots & \sigma_{2N}\boldsymbol{\Omega}_{2N} \\ & \vdots & & \\ \sigma_{N1}\boldsymbol{\Omega}_{N1} & \sigma_{N2}\boldsymbol{\Omega}_{N2} & \cdots & \sigma_{NN}\boldsymbol{\Omega}_{NN} \end{bmatrix} \tag{2.36}$$

przy czym $\mathbf{V}_{ij} = \sigma_{ij}\boldsymbol{\Omega}_{ij}$ jest macierzą wariancji-kowariancji między składnikami losowymi pochodzącymi z i-tego i j-tego obiektu. Postacie tych macierzy mogą być różne, w zależności od założeń dotyczących składnika losowego.

W najprostszym przypadku zakłada się, że w modelu (2.35) występuje jedynie zjawisko heteroskedastyczności grupowej, która jest efektem tego, iż różne obiekty podlegają różnym procesom stochastycznym. Kowariancje składników losowych z różnych obiektów są natomiast zerowe oraz nie występuje zjawisko autokorelacji. Założenia te oznaczają, że $\mathbf{V}$ jest macierzą diagonalną, a bloki na głównej przekątnej są równe $\sigma_i^2\,\mathbf{I}$, gdzie $\sigma_i^2 = \sigma_{ii}$ jest wariancją składnika losowego specyficzną dla i-tego obiektu. Zatem:

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \sigma_2^2\mathbf{I} & \cdots & \mathbf{0} \\ & \vdots & & \\ \mathbf{0} & \mathbf{0} & \cdots & \sigma_N^2\mathbf{I} \end{bmatrix} \tag{2.37}$$

Wariancja składnika losowego jest zatem stała w czasie dla każdego z analizowanych obiektów, ale różna dla różnych obiektów. Do testowania hipotezy o występowaniu heteroskedastyczności grupowej stosować można wiele testów, m. in.: Goldfelda–Quandta, White'a, mnożnika Lagrange'a, ilorazu wiarygodności, Walda. Sposób wnioskowania na podstawie trzech ostatnich testów w szczególnym przypadku heteroskedastyczności grupowej zostanie przedstawiony poniżej.

Statystyka empiryczna testu mnożnika Lagrange'a ma postać:

$$LM = \frac{T}{2}\sum_{i=1}^{n}\left(\frac{s_i^2}{s^2} - 1\right)^2 \tag{2.38}$$

gdzie s_i^2 jest oceną σ_i^2, a s^2 – oceną wariancji modelu *pooled*, tzn. modelu oszacowanego KMNK na podstawie wszystkich *NT* obserwacji. Statystyka *LM* ma rozkład χ^2 z *N* – 1 stopniami swobody. Jeśli *LM* > χ^2_{kryt}, to odrzucamy hipotezę zerową H_0: $\sigma_i^2 = \sigma_j^2$ dla *i*, *j* = 1, ..., *N*, o równości wariancji dla wszystkich obiektów.

Alternatywą dla testu mnożnika Lagrange'a jest test ilorazu wiarygodności (ang. *likehood ratio test*). Wartość statystyki empirycznej tego testu wyznacza się ze związku:

$$-2\ln\lambda = NT\ln\hat{\sigma}^2 - \sum_{i=1}^{N} T\ln\hat{\sigma}_i^2 \tag{2.39}$$

gdzie: $\hat{\sigma}^2 = \frac{1}{NT}\mathbf{e}^{\mathrm{T}}\mathbf{e}$, $\hat{\sigma}_i^2 = \frac{1}{NT}\mathbf{e}_i^{\mathrm{T}}\mathbf{e}_i$.

Współrzędne wektora reszt $\mathbf{e}_i$ wyznacza się na podstawie reszt estymatorów metody największej wiarygodności. Można tu wykorzystać np. reszty iteracyjnej UMNK z estymacją. Statystyka empiryczna ma rozkład χ^2 z *N* – 1 stopniami swobody. Jeśli zatem wartość obliczona przewyższa wartość krytyczną χ^2_{kryt}, to odrzucamy hipotezę H_0: $\sigma_i^2 = \sigma_j^2$ dla *i*, *j* = 1, ..., *N*.

Test Walda różni się od wymienionych powyżej tym, że może być zastosowany po oszacowaniu modelu uwzględniającego heteroskedastyczność grupową, gdyż obliczenie statystyki empirycznej wymaga znajomości ocen s_i^2 wariancji σ_i^2 obliczonych na podstawie reszt UMNK z estymacją. Statystyka empiryczna ma postać:

$$W = \frac{T}{2}\sum_{i=1}^{N}\left(\frac{s^2}{s_i^2} - 1\right)^2 \tag{2.40}$$

gdzie s^2 jest oceną wariancji wyznaczoną z modelu *pooled*. Statystyka W ma rozkład χ^2 z N stopniami swobody. Jeśli $LM > \chi^2_{\text{kryt}}$, to odrzucamy hipotezę zerową H_0: $\sigma_i^2 = \sigma_j^2$ dla $i, j = 1, ..., N$.

Zjawisko heteroskedastyczności grupowej jest, jak już wspomniano, szczególnym przypadkiem heteroskedastyczności składnika losowego. Do estymacji modelu, w którym zakłada się jego występowanie, można zatem zastosować m. in. ważoną metodę najmniejszych kwadratów (WMNK). Jej stosowanie wydaje się o tyle uzasadnione, że założenie o grupowej naturze heteroskedastyczności umożliwia dokonanie naturalnego podziału na podpróby (tworzą je obserwacje w czasie pochodzące z poszczególnych obiektów) i oszacowanie wariancji dla każdej podpróby na podstawie reszt KMNK. Estymator WMNK wyraża się, jak wiadomo, wzorem:

$$\hat{\boldsymbol{\beta}}_{\text{WMNK}} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\left(\mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{y}\right) = \left[\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\mathbf{X}_i^{\mathrm{T}}\mathbf{X}_i\right]^{-1}\left[\sum_{i=1}^{N}\frac{1}{\sigma_i^2}\mathbf{X}_i^{\mathrm{T}}\mathbf{y}_i\right] \tag{2.41}$$

Nieznane, rzeczywiste wartości wariancji σ_i^2 zastępuje się ich ocenami, wyznaczonymi na podstawie reszt KMNK, a weryfikacja modelu przebiega analogicznie, jak w przypadku modeli szacowanych na podstawie danych przekrojowych.

2.5.2. Modele z heteroskedastycznością grupową i korelacją przekrojową składnika losowego

Oprócz zjawiska heteroskedastyczności, w modelach szacowanych na podstawie próby czasowo-przekrojowej, bardzo częstym zjawiskiem jest korelacja przekrojowa składnika losowego, występująca zazwyczaj jednocześnie z heteroskedastycznością grupową.

Zjawisko korelacji przekrojowej (ang. *cross-sectional correlation*, *contemporaneous correlation*) polega na korelacji składników losowych pochodzących

z różnych obiektów, ale z tego samego okresu. Jego występowanie w modelach szacowanych na podstawie danych czasowo-przekrojowych wydaje się naturalne, gdyż łączenie danych z różnych obiektów w jeden szereg jest zazwyczaj uwarunkowane tym, że analizowane obiekty są ze sobą powiązane, np. funkcjonują w obrębie jednego systemu gospodarczego (np. województwa) lub są elementami tej samej klasyfikacji (np. sekcje PKD). Wpływające na ich kondycję czynniki makroekonomiczne działają na nie w zróżnicowanym stopniu.

Występowanie zjawiska korelacji przekrojowej oznacza, że każdy blok macierzy (2.36) ma postać:

$$\mathbf{V} = [\mathbf{V}_{ij}] = [\sigma_{ij}\mathbf{I}], \qquad i,\ j = 1, ..., N \tag{2.42}$$

Dla przykładu, jeśli zebrane zostałyby po 3 obserwacje w czasie dla 2 obiektów, macierz **V** miałaby postać:

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \sigma_{12} & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 & \sigma_{12} & 0 \\ 0 & 0 & \sigma_1^2 & 0 & 0 & \sigma_{12} \\ \sigma_{21} & 0 & 0 & \sigma_2^2 & 0 & 0 \\ 0 & \sigma_{21} & 0 & 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_{21} & 0 & 0 & \sigma_2^2 \end{bmatrix}.$$

Oprócz tego, że macierz **V** nie jest macierzą diagonalną, uwagę zwraca dodatkowo fakt, że elementy na głównej przekątnej tej macierzy różnią się względem obiektów, natomiast dla każdego obiektu wariancja jest stała w czasie. Oznacza to, że dopuszcza się jednoczesne występowanie zjawisk heteroskedastyczności grupowej i korelacji przekrojowej składnika losowego.

Do testowania występowania zjawiska korelacji przekrojowej stosować można test mnożnika Lagrange'a, opracowany przez Breuscha i Pagana lub test ilorazu wiarygodności.

Statystyka empiryczna testu Breuscha–Pagana ma postać:

$$\lambda_{LM} = T\sum_{i=2}^{N}\sum_{j=1}^{i-1} r_{ij}^2 \tag{2.43}$$

gdzie r_{ij} są współczynnikami korelacji reszt między obiektami *i*-tym i *j*-tym. Współczynniki korelacji oblicza się na podstawie reszt UMNK, ze względu na założenie o stałości parametrów strukturalnych względem obiektów.

Statystyka λ_{LM} ma rozkład χ^2 z $N \cdot (N-1)/2$ stopniami swobody. Jeśli $\lambda_{LM} > \chi^2_{\text{kryt}}$, to hipoteza zerowa H_0: $\sigma_{ij} = 0$ dla $i \neq j$ o tym, że zjawisko

korelacji przekrojowej nie występuje, jest odrzucana na rzecz hipotezy alternatywnej.

Test ilorazu wiarygodności weryfikuje występowanie korelacji *a posteriori*, gdyż wymaga znajomości ocen wariancji składnika losowego wyznaczonych na podstawie dwóch modeli: modelu z heteroskedastycznością grupową oraz modelu z heteroskedastycznością grupową i korelacją przekrojową. Może on być stosowany, jeśli wyznaczono estymatory równoważne estymatorom największej wiarygodności. Hipotezy formułowane są identycznie, jak w teście Breuscha–Pagana, a do ich weryfikacji służy statystyka:

$$\lambda_{LR} = T\left(\sum_{i=1}^{N} \ln \hat{\sigma}_i^2 - \ln\left|\hat{\boldsymbol{\Sigma}}\right|\right) \tag{2.44}$$

gdzie: $\boldsymbol{\Sigma} = [\sigma_{ij}]_{\mathrm{T}}$, elementami macierzy $\hat{\boldsymbol{\Sigma}}$ są oceny kowariancji σ_{ij}, wyznaczone na podstawie reszt modelu oszacowanego z uwzględnieniem zjawisk heteroskedastyczności grupowej i korelacji przekrojowej, $|\cdot|$ oznacza wyznacznik macierzy, zaś $\hat{\sigma}_i^2$ są ocenami wariancji, wyznaczonymi na podstawie reszt modelu oszacowanego z uwzględnieniem zjawiska heteroskedastyczności grupowej.

Statystyka λ_{LR} ma rozkład χ^2 z $N \cdot (N - 1)/2$ stopniami swobody. Jeśli $\lambda_{\mathrm{LR}} > \chi^2_{\mathrm{kryt}}$, to hipoteza zerowa jest odrzucana na rzecz hipotezy alternatywnej.

Do estymacji modeli, w których stwierdzono występowanie korelacji przekrojowej i heteroskedastyczności grupowej, wykorzystuje się jedną z dwóch metod: UMNK z estymacją lub technikę PCSE (ang. *panel-corrected standard errors*), zaproponowane przez B e c k a i K a t z a [1995]. Pierwsza z nich, zwana też metodą Parksa, jest metodą tradycyjnie stosowaną. Druga, opracowana dość niedawno, jest prostsza, choć jej stosowanie budzi wiele kontrowersji. Obie metody zostaną omówione poniżej.

UMNK z estymacją jest metodą powszechnie znaną, jednak w przypadku wykorzystania danych czasowo-przekrojowych wymiary macierzy **V** mogą być na tyle duże, że jej bezpośrednie odwrócenie jest utrudnione. Dlatego stosuje się następujące postępowanie. Po wyznaczeniu zgodnych (choć obciążonych) estymatorów kowariancji między obiektem *i*-tym i *j*-tym na podstawie reszt KMNK, w drugim kroku wyznacza się estymatory parametrów strukturalnych według wzoru:

$$\hat{\boldsymbol{\beta}}_{\mathrm{UMNK}} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\left(\mathbf{X}^{\mathrm{T}}\mathbf{V}^{-1}\mathbf{y}\right) \tag{2.45}$$

Macierz **V** zastępuje się macierzą $\hat{\mathbf{V}}$, której elementami są zgodne oceny kowariancji wyznaczone w pierwszym kroku na podstawie reszt KMNK.

Macierz **V** jest stopnia *NT*, jednak jej odwrócenie jest w praktyce stosunkowo proste, jeśli nie dokonuje się go bezpośrednio. Macierz tę, określoną wzorem (2.42), zapisać można równoważnie w postaci:

$$\mathbf{V} = \mathbf{\Sigma} \otimes \mathbf{I} \tag{2.46}$$

gdzie symbol $\otimes$ oznacza iloczyn Kroneckera, zaś

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2N} \\ \dots & \dots & \dots & \dots \\ \sigma_{N1} & \sigma_{N2} & \dots & \sigma_{NN} \end{bmatrix}_T .$$

Wówczas:

$$\mathbf{V}^{-1} = \mathbf{\Sigma}^{-1} \otimes \mathbf{I} \tag{2.47}$$

co oznacza, że w praktyce odwraca się macierz stopnia *N*. Jeśli przez σ^{ij} oznaczymy *ij*-ty element macierzy $\mathbf{\Sigma}^{-1}$, to wzór (2.45) zapisać można w postaci:

$$\hat{\beta}_{\text{EUMNK}} = \left[\sum_{i=1}^{N} \sum_{j=1}^{N} \sigma^{ij} \mathbf{X}_i^{\mathrm{T}} \mathbf{X}_j \right]^{-1} \left[\sum_{i=1}^{N} \sum_{j=1}^{N} \sigma^{ij} \mathbf{X}_i^{\mathrm{T}} \mathbf{y}_j \right] \tag{2.48}$$

Macierz wariancji-kowariancji estymatora jest określona przez wyrażenie w pierwszym nawiasie kwadratowym wzoru (2.48).

Estymatory największej wiarygodności wyznacza się stosując procedurę iteracyjną, aż do osiągnięcia zbieżności między:

$$\hat{\boldsymbol{\beta}}_{\text{EUMNK}} = \left(\mathbf{X}^{\mathrm{T}} \hat{\mathbf{V}}^{-1} \mathbf{X} \right)^{-1} \left(\mathbf{X}^{\mathrm{T}} \hat{\mathbf{V}}^{-1} \mathbf{y} \right)$$

i

$$\hat{\sigma}_{ij} = \frac{\mathbf{y}_i^{\mathrm{T}} \mathbf{y}_j + \hat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{X}_i^{\mathrm{T}} \mathbf{X}_j \hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{X}_i^{\mathrm{T}} \mathbf{y}_j - \hat{\boldsymbol{\beta}}^{\mathrm{T}} \mathbf{X}_j^{\mathrm{T}} \mathbf{y}_i}{T} \tag{2.49}$$

B e c k i K a t z [1995] zaproponowali oszacowanie parametrów strukturalnych modelu przy zastosowaniu KMNK. Uzyskane w ten sposób estymatory są zgodne i nieobciążone. Problem ich nieefektywności jest rozwiązywany przez zastosowanie specyficznej formuły obliczania średnich błędów szacunku parametrów, tzw. „skorygowanych błędów standardowych dla danych panelo-

wych” (ang. *panel corrected standard errors*, PCSE). Efektywność PCSE rośnie wraz ze wzrostem liczby okresów w stosunku do liczby obiektów. Macierz wariancji-kowariancji estymatorów proponowana przez Becka i Katza przyjmuje postać:

$$\mathrm{D}^2\left(\hat{\boldsymbol{\beta}}\right)_{\mathrm{PCSE}} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1}\left(\mathbf{X}\hat{\mathbf{V}}\mathbf{X}\right)\left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{-1} \tag{2.50}$$

Autorzy proponują oszacowanie nieznanej macierzy **V** jako iloczynu Kroneckera:

$$\hat{\mathbf{V}} = \frac{\mathbf{E}^{\mathrm{T}}\mathbf{E}}{T} \otimes \mathbf{I} \tag{2.51}$$

gdzie $\mathbf{E} = [e_{it}]_{T\times N}$ jest macierzą reszt modelu (2.35) oszacowanego KMNK, w której kolumna o numerze *i* zawiera reszty pochodzące z *i*-tego obiektu.

Elementami macierzy $\hat{\mathbf{V}}$ są:

$$s_{ij} = \frac{1}{T}\sum_{t=1}^{T} e_{it}e_{jt} \tag{2.52}$$

W szczególności, wariancje są równe:

$$s_i^2 = \frac{1}{T}\sum_{t=1}^{T} e_{it}^2 \tag{2.53}$$

Oznacza to, że kowariancje składników losowych pochodzących z *i*-tego i *j*-tego obiektu są stałe w czasie; wariancje są szacowane jako średnie błędy kwadratowe reszt.

Zarówno metoda Parksa, jak i metoda PCSE mają swoich zwolenników i przeciwników. Autorzy metody PCSE zarzucają podejściu Parksa to, że estymatory UMNK wariancji parametrów strukturalnych są obciążone (niedoszacowane), ponieważ wyznaczane są na podstawie kolejnych oszacowań macierzy wariancji-kowariancji. Metoda PCSE krytykowana jest zaś za to, że nie podejmuje się w niej próby modelowania procesów generujących występowanie zjawisk, a jedynie koryguje błędy standardowe (M a d d a l a [1998]), oraz za dość skomplikowany sposób obliczania PCSE w przypadku modelu z opóźnioną zmienną objaśnianą (W o r r a l l, P r a t t [2004]).

2.5.3. Modele z heteroskedastycznością grupową, korelacją przekrojową i autokorelacją składnika losowego

Powyższe rozważania dotyczą przypadku, gdy w modelu nie występuje autokorelacja składnika losowego w czasie. Możliwe jest rozszerzenie analizy o uwzględnienie również tego zjawiska. Przyjmuje się zazwyczaj jedno z następujących założeń:

1) składniki losowe z różnych obiektów są nieskorelowane; autokorelacja występuje jedynie „wewnątrz" poszczególnych obiektów. Oznacza to w szczególności, że nie dopuszcza się korelacji przekrojowej,

2) oprócz autokorelacji dopuszcza się również korelację składników losowych z różnych obiektów – korelację przekrojową.

W każdym z powyższych dwóch przypadków oceny współczynnika autokorelacji ρ_i właściwego dla i-tego obiektu można uzyskać przy założeniu, że jest on: a) inny dla każdego obiektu: $\rho_i \neq \rho_j$ dla $i \neq j$, b) jednakowy dla wszystkich obiektów: $\rho_i = \rho_j = \rho$.

Poniżej omówione zostaną sposoby testowania, czy zjawisko autokorelacji występuje, oraz metody estymacji modeli, w których stwierdzono jej występowanie.

Weryfikacji hipotezy o występowaniu autokorelacji można dokonać za pomocą kilku testów statystycznych, m. in. testu Durbina–Watsona lub testu mnożnika Lagrange'a dla danych czasowo-przekrojowych. W przypadku testu Durbina–Watsona pochodzenie reszt, wykorzystywanych do konstrukcji statystyki empirycznej, zależy od założeń dotyczących współczynnika autokorelacji. Przy założeniu, że współczynnik ten jest jednakowy dla wszystkich obiektów, do obliczenia statystyki D–W wykorzystuje się reszty modelu oszacowanego KMNK na podstawie całej próby. Jeśli zakłada się, że wartość współczynnika autokorelacji może się różnić między obiektami, to do obliczenia statystyki D–W wykorzystuje się reszty modeli oszacowanych KMNK dla każdego obiektu osobno. Podobnie postępuje się stosując test mnożnika Lagrange'a.

Dobór odpowiedniej metody estymacji zależy od założeń przyjętych co do innych zjawisk występujących w modelu i zmienności współczynnika autokorelacji względem obiektów.

Założenie o zmienności (lub stałości) współczynnika autokorelacji względem obiektów jest niezależne od założeń przyjmowanych co do występowania korelacji przekrojowej. Jeśli zakłada się, że $\rho_i = \rho_j = \rho$, to narzuca się na model pewne ograniczenie. Może ono jednak być niezbędne, np. gdy szeregi czasowe są stosunkowo krótkie. Jego przyjęcie korzystnie wpływa wówczas na własności estymatora współczynnika ρ w małych próbach. B e c k, K a t z [1995] zalecają, aby zawsze zakładać jednakową wartość ρ. Zgodnie z tym punktem widzenia obliczanie N różnych ocen współczynnika autokorelacji skutkuje wzrostem (skumulowaniem) błędów szacunku, co negatywnie wpływa na jakość estymato-

rów parametrów strukturalnych szacowanego modelu. Przeprowadzone eksperymenty Monte Carlo (B e c k, K a t z [1996]) wykazały, że przyjęcie założenia $\rho_i = \rho_j = \rho$, nawet jeśli jest ono fałszywe, daje ostatecznie lepsze oszacowania niż wymagające obliczania n różnych ocen założenie $\rho_i \neq \rho_j$ dla $i \neq j$. Ci sami autorzy wskazują również, że przyjęcie założenia o jednakowym dla wszystkich obiektów współczynniku ρ jest naturalną konsekwencją założenia, że parametry strukturalne nie różnią się między obiektami.

Przy założeniu, że $\rho_i \neq \rho_j$ dla $i \neq j$, początkową wartość oceny r_i współczynnika ρ_i wyznacza się na podstawie regresji KMNK, dla każdego obiektu osobno.

Przy założeniu, że $\rho_i = \rho_j = \rho$, możliwe jest wyznaczenie trzech alternatywnych estymatorów r współczynnika ρ, z których każdy jest zgodny. Są one dane wzorami:

$$
\begin{aligned}
&\text{a)}\ r = \frac{1}{N}\sum_{i=1}^{N} r_i \quad \text{lub} \\
&\text{b)}\ r = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T} e_{it}e_{i,t-1}}{\sum_{i=1}^{N}\sum_{t=2}^{T} e_{it}^2} \quad \text{lub} \\
&\text{c)}\ r = \mathrm{correl}\left(e_{i,t-1}, e_{it}\right)
\end{aligned}
\tag{2.54}
$$

gdzie correl( ) oznacza współczynnik korelacji z próby. Niezależnie od wyboru estymatora (2.54), UMNK z estymacją daje estymatory asymptotycznie efektywne.

Jeśli zjawisko autokorelacji występuje, to po wyznaczeniu (w dowolny sposób) początkowej wartości r (lub r_i) stosuje się metody estymacji opisane poniżej, zależne od tego, czy w modelu występuje jednocześnie zjawisko korelacji przekrojowej.

W przypadku, gdy zakłada się, że występuje tylko autokorelacja „wewnątrz” poszczególnych obiektów, zaś nie dopuszcza się korelacji przekrojowej, tzn. correl(ε_{it}, ε_{js}) = 0 dla $i \neq j$, macierz $\mathbf{V}$ jest macierzą blokowo-diagonalną postaci:

$$
\mathbf{V} = \begin{bmatrix}
\sigma_1^2\mathbf{\Omega}_1 & \mathbf{0} & \cdots & \mathbf{0} \\
\mathbf{0} & \sigma_2^2\mathbf{\Omega}_2 & \cdots & \mathbf{0} \\
 & \vdots & & \\
\mathbf{0} & \mathbf{0} & \cdots & \sigma_N^2\mathbf{\Omega}_N
\end{bmatrix}
\tag{2.55}
$$

Zakłada się przy tym zazwyczaj, że składniki losowe podlegają procesowi AR(1): $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + u_{it}$. Wariancja składnika losowego jest wtedy równa:

$$\mathrm{D}^2(\varepsilon_{it}) = \sigma_i^2 = \frac{\sigma_{ui}^2}{1-\rho_i^2} \tag{2.56}$$

W celu usunięcia autokorelacji stosuje się zazwyczaj przekształcenie Praisa–Winstena, które polega na tym, że szeregi czasowe dotyczące kolejnych obiektów transformuje się w następujący sposób:

$$\mathbf{y}_i^* = \begin{bmatrix} \sqrt{1-r_i^2}\, y_{i1} \\ y_{i2} - r_i y_{i1} \\ \vdots \\ y_{iT} - r_i y_{i,T-1} \end{bmatrix}, \qquad \mathbf{X}_i^* = \begin{bmatrix} \sqrt{1-r_i^2}\, \mathbf{x}_{i1} \\ \mathbf{x}_{i2} - r_i \mathbf{x}_{i1} \\ \vdots \\ \mathbf{x}_{iT} - r_i \mathbf{x}_{i,T-1} \end{bmatrix} \tag{2.57}$$

gdzie r_i jest oceną współczynnika autokorelacji, $\mathbf{x}_{it}$ jest wektorem wartości wszystkich zmiennych objaśniających przyjmowanych dla *i*-tego obiektu w okresie *t*.

W modelu szacowanym na podstawie transformowanych danych nie występuje już autokorelacja, a jedynie heteroskedastyczność grupowa. Estymacji dokonuje się metodami omówionymi w poprzednich podrozdziałach. Wariancję resztową oblicza się na podstawie wzoru (2.56), zastępując wartości rzeczywiste ρ_i i σ_{ui}^2 ich oszacowaniami: r_i i $s_{ui}^2 = \frac{(\mathbf{e}_i^*)^{\mathrm{T}} \mathbf{e}_i^*}{T}$, gdzie $\mathbf{e}_i^*$ oznacza wektor reszt modelu oszacowanego na podstawie danych transformowanych za pomocą przekształcenia (2.57).

W przypadku, gdy dopuszcza się korelację składników losowych z różnych obiektów, zachodzi równość: cov(u_{it}, u_{jt}) = σ_{uij}, a macierz **V** ma postać:

$$\mathbf{V} = \begin{bmatrix} \sigma_1^2 \mathbf{\Omega}_1 & \sigma_{12} \mathbf{\Omega}_{12} & \cdots & \sigma_{1n} \mathbf{\Omega}_{1n} \\ \sigma_{21} \mathbf{\Omega}_{21} & \sigma_2^2 \mathbf{\Omega}_2 & \cdots & \sigma_{12n} \mathbf{\Omega}_{2n} \\ & \vdots & & \\ \sigma_{n1} \mathbf{\Omega}_{n1} & \sigma_{n2} \mathbf{\Omega}_{n2} & \cdots & \sigma_n^2 \mathbf{\Omega}_n \end{bmatrix} \tag{2.58}$$

Przekształcenie Praisa–Winstena, usuwając autokorelację składnika losowego w czasie, prowadzi w tym przypadku do modelu z heteroskedastycznością grupową i korelacją przekrojową. Oznacza to, że macierz **V** przyjmuje dla danych transformowanych postać analogiczną do (2.42). Do estymacji modelu

można zatem zastosować metody opisane w pkt 2.5.2, przy czym estymatory kowariancji σ_{ij} wyznacza się ze wzoru:

$$\sigma_{ij} = \frac{\sigma_{uij}}{1-\rho_i \rho_j},$$

zastępując wartości rzeczywiste ρ_i, ρ_j i σ_{uij} ich oszacowaniami: r_i, r_j i $s_{uij} = \frac{\left(\mathbf{e}_i^*\right)^{\mathrm{T}} \mathbf{e}_j^*}{T}$, gdzie $\mathbf{e}_i^*$ i $\mathbf{e}_j^*$ oznaczają wektory reszt modelu oszacowanego na podstawie danych transformowanych za pomocą przekształcenia (2.57).

2.6. Panelowe modele wielorównaniowe

W badaniach ekonomicznych niejednokrotnie zachodzi konieczność estymowania modeli wielorównaniowych. Przykładem może być układ równań popytu dla różnych sektorów, działów przemysłu lub regionów. Metody estymacji takich modeli bazują na metodach stosowanych dla modeli szeregów czasowych lub danych przekrojowych. Dobór odpowiedniej metody zależy od rodzaju modelu. Poniżej przedstawione zostaną pokrótce metody stosowane do estymacji panelowych modeli o równaniach pozornie niezależnych oraz modeli o równaniach łącznie współzależnych. Rozważania dotyczyć będą, podobnie jak w poprzednich rozdziałach, modeli, w których wyodrębnia się tylko efekt grupowy α_i. Metody postępowania w przypadku rozdziału składnika losowego na trzy składowe: $u_{it} = \alpha_i + \lambda_t + \varepsilon_{it}$ są analogiczne. W literaturze polskojęzycznej zostały one omówione w: D a ń s k a [2000] (rozdz. 2).

2.6.1. Metody estymacji modeli prostych o równaniach pozornie niezależnych (SUR)

Odpowiednim sposobem postępowania w przypadku modeli prostych jest często potraktowanie ich jako układu równań pozornie niezależnych (ang. *seemingly unrelated regressions*, SUR), zaproponowane przez Z e l l n e r a [1962]. Zastosowanie tej metody powoduje, jak wiadomo, wzrost efektywności estymatorów, uzyskany dzięki uwzględnieniu korelacji składników losowych między poszczególnymi równaniami. Do estymacji modelu panelowego z efektami losowymi została ona zastosowana po raz pierwszy przez A v e r y'e g o [1977].

Panelowy model SUR ma postać:

$$\mathbf{y}_j = \mathbf{Z}_j \boldsymbol{\delta}_j + \mathbf{u}_j \quad \text{dla } j = 1, ..., M \tag{2.59}$$

gdzie $\mathbf{y}_j$ jest wektorem o wymiarach ($NT \times 1$), $\mathbf{Z}_j$ – macierzą o wymiarach ($NT \times k_j^*$), $\boldsymbol{\delta}_j^{\mathrm{T}} = \left(\boldsymbol{\alpha}_j, \boldsymbol{\beta}_j^{\mathrm{T}}\right)$, $\boldsymbol{\beta}_j$ jest wektorem ($k_j \times 1$), $k_j^* = k_j + 1$. Składnik losowy $\mathbf{u}_j$ dowolnego j-tego równania rozdzielany jest na dwie składowe i zapisywany w postaci:

$$\mathbf{u}_j = \mathbf{Z}_\alpha \boldsymbol{\alpha}_j + \boldsymbol{\varepsilon}_j \quad \text{dla} \quad j = 1, ..., M \tag{2.60}$$

gdzie: $\mathbf{Z}_\alpha = \left(\mathbf{I}_N \otimes \mathbf{i}_T\right)$, zaś $\boldsymbol{\alpha}_j^{\mathrm{T}} = \left(\alpha_{1j}, \alpha_{2j}, ..., \alpha_{Nj}\right)$ i $\boldsymbol{\varepsilon}_j^{\mathrm{T}} = \left(\varepsilon_{11j}, \varepsilon_{12j}, ..., \varepsilon_{1Tj}, ..., \varepsilon_{N1j}, \varepsilon_{N2j}, ..., \varepsilon_{NTj}\right)$ są wektorami losowymi o wartości oczekiwanej równej zero. Zakłada się również, że w dowolnym j-tym równaniu efekty grupowe α_{ij} są niezależne od ε_{itj} $\left(\varepsilon_{itj}: \mathrm{IID}\left(0, \sigma_\varepsilon^2\right)\right.$ oraz $\left.\alpha_{ij}: \mathrm{IID}\left(0, \sigma_\alpha^2\right)\right)$. Przyjmuje się też założenie Zellnera, że składniki losowe poszczególnych równań są skorelowane w danym okresie (dla danego obiektu), a nieskorelowane w różnych okresach (dla różnych obiektów). Macierz wariancji-kowariancji składników losowych pochodzących z j-tego i l-tego równania postaci (2.59) ma postać:

$$\boldsymbol{\Omega}_{jl} = \mathrm{E}\left(\mathbf{u}_j \mathbf{u}_l^{\mathrm{T}}\right) = \sigma_{\alpha jl}^2 \left(\mathbf{I}_N \otimes \mathbf{J}_T\right) + \sigma_{\varepsilon jl}^2 \left(\mathbf{I}_N \otimes \mathbf{J}_T\right) \tag{2.61}$$

Macierz (2.61) ma więc formę analogiczną do macierzy $\boldsymbol{\Omega}$ wariancji-kowariancji składnika losowego modelu jednorówaniowego (wzór (2.14) w podrozdz. 2.2). Jednakże w przypadku rozważanym obecnie występują dodatkowo wariancje międzyrównaniowe $\sigma_{\alpha_{jl}}^2$ i $\sigma_{\varepsilon_{jl}}^2$, które również podlegają estymacji.

Łączna macierz wariancji-kowariancji dla całego układu M równań (2.59) ma postać:

$$\boldsymbol{\Omega} = \Sigma\left(\mathbf{u}\mathbf{u}^{\mathrm{T}}\right) = \boldsymbol{\Sigma}_\alpha \otimes \left(\mathbf{I}_N \otimes \mathbf{J}_T\right) + \boldsymbol{\Sigma}_\varepsilon \otimes \left(\mathbf{I}_N \otimes \mathbf{I}_T\right) \tag{2.62}$$

gdzie $\boldsymbol{\Sigma}_\alpha = [\sigma_{\alpha jl}^2]$ i $\boldsymbol{\Sigma}_\varepsilon = [\sigma_{\varepsilon jl}^2]$ są macierzami stopnia M. Ponadto składowe składnika losowego całego układu M równań $\boldsymbol{\alpha}^{\mathrm{T}} = \left(\boldsymbol{\alpha}_1^{\mathrm{T}}, ..., \boldsymbol{\alpha}_M^{\mathrm{T}}\right)$ i $\boldsymbol{\varepsilon}^{\mathrm{T}} = \left(\boldsymbol{\varepsilon}_1^{\mathrm{T}}, ..., \boldsymbol{\varepsilon}_M^{\mathrm{T}}\right)$ mają rozkłady $\boldsymbol{\alpha} \sim (\mathbf{0}, \boldsymbol{\Sigma}_\alpha \otimes \mathbf{I}_N)$, $\boldsymbol{\varepsilon} \sim (\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon \otimes \mathbf{I}_{NT})$, co oznacza, że spełniają założenia SUR.

Macierz (2.62) zapisać można w postaci alternatywnej (por. np. D a ń s k a [2000]):

$$\mathbf{\Omega} = \mathbf{\Sigma}_1 \otimes \mathbf{P} + \mathbf{\Sigma}_\alpha \otimes \mathbf{Q} \tag{2.63}$$

gdzie: $\mathbf{\Sigma}_1 = T\mathbf{\Sigma}_\alpha + \mathbf{\Sigma}_\varepsilon$, $\mathbf{P}$ jest operatorem międzygrupowym (wzór (2.15)), $\mathbf{Q}$ – operatorem wewnątrzgrupowym (wzór (2.8)). Wówczas $\mathbf{\Omega}^{-1} = \mathbf{\Sigma}_1^{-1} \otimes \mathbf{P} + \mathbf{\Sigma}_\alpha^{-1} \otimes \mathbf{Q}$, co wykazał B a l t a g i [1980], a nieobciążone estymatory składowych macierzy (2.63) wyznaczyć można jako:

$$\hat{\mathbf{\Sigma}}_\varepsilon = \mathbf{U}^{\mathrm{T}}\mathbf{Q}\mathbf{U}/N(T-1) \quad \text{oraz} \quad \hat{\mathbf{\Sigma}}_1 = \mathbf{U}^{\mathrm{T}}\mathbf{Q}\mathbf{U}/N \tag{2.64}$$

gdzie $\mathbf{U} = (\mathbf{u}_1, ..., \mathbf{u}_M)$ jest macierzą składników losowych wszystkich M równań modelu o wymiarach ($NT \times M$). Elementy macierzy $\mathbf{U}$ zastąpić można resztami MNK (por. Avery [1977]) albo resztami estymatora wewnątrzgrupowego (A m e m i y a [1971]; B a l t a g i [1980]).

Ponieważ $\mathbf{\Omega}$ jest macierzą blokowo-diagonalną, to zastosowanie UMNK do łącznej estymacji układu M równań (2.59) jest równoważne zastosowaniu jej do każdego równania osobno, o ile zbiory zmiennych objaśniających tych równań różnią się. Jeśli jednak macierz zmiennych objaśniających w każdym równaniu jest identyczna, to estymatory te nie są równoważne. Estymatory typu Zellnera są asymptotycznie efektywne, jeżeli $\hat{\mathbf{\Sigma}}_\varepsilon$ jest zgodny, a $\hat{\mathbf{\Sigma}}_\alpha$ ma skończoną dodatnio określoną granicę (por. P r u c h a [1984]).

2.6.2. Metody estymacji modeli wielorównaniowych o równaniach łącznie współzależnych

Powyższe rozważania dotyczyły modeli wielorównaniowych prostych, a przedstawiona metoda estymacji stanowiła rozszerzenie procedury Zellnera stosowanej dla modeli szeregów czasowych lub danych przekrojowych. Metody estymacji modeli wielorównaniowych o równaniach łącznie współzależnych również są adaptacją metod stosowanych do estymacji na podstawie danych „jednowymiarowych" i bazują na dwustopniowej MNK (2MNK) lub trójstopniowej MNK (3MNK).

W celu prezentacji metod estymacji pojedynczych równań (bazujących na 2MNK) załóżmy, że pierwsze równanie modelu o równaniach łącznie współzależnych ma postać:

$$\mathbf{y}_1 = \mathbf{Z}_1\boldsymbol{\delta}_1 + \mathbf{u}_1 \tag{2.65}$$

gdzie: $\mathbf{Z}_1 = [\mathbf{Y}_1, \mathbf{X}_1]$, $\boldsymbol{\delta}_1^{\mathrm{T}} = (\boldsymbol{\gamma}_1^{\mathrm{T}}, \boldsymbol{\beta}_1^{\mathrm{T}})$, przy czym $\mathbf{Y}_1$ jest g_1-elementowym zbiorem zmiennych endogenicznych, $\mathbf{X}_1$ – k_1-elementowym zbiorem zmiennych egzoge-

nicznych występujących w tym równaniu. Jak wiadomo, równanie (2.65) jest identyfikowalne, jeżeli liczba zmiennych egzogenicznych niewystępujących w pierwszym równaniu k_2 jest nie mniejsza niż liczba g_1. Niech ponadto $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2]$ będzie zbiorem wszystkich zmiennych egzogenicznych modelu. Składnik losowy modelu (2.65) ma postać:

$$\mathbf{u}_1 = \mathbf{Z}_\alpha \boldsymbol{\alpha}_1 + \boldsymbol{\varepsilon}_1 \tag{2.66}$$

gdzie: $\mathbf{Z}_\alpha$, $\boldsymbol{\alpha}_1$, $\boldsymbol{\varepsilon}_1$ określone są analogicznie jak w równaniu (2.60).

Macierz wariancji-kowariancji równania (2.66) ma postać:

$$\mathrm{E}\left(\mathbf{u}_1\mathbf{u}_1^{\mathrm{T}}\right) = \boldsymbol{\Omega}_{11} = \sigma_{\alpha_{11}}^2 \left(\mathbf{I}_N \otimes \mathbf{J}_T\right) + \sigma_{\varepsilon_{11}}^2 \mathbf{I}_{NT} \tag{2.67}$$

Podwójny subskrypt (11) przy składowych α i ε wskazuje, że jest to macierz właściwa dla pierwszego równania.

W wyniku pomnożenia równania (2.65) przez operator wewnątrzgrupowy $\mathbf{Q}$ otrzymuje się:

$$\tilde{\mathbf{y}}_1 = \tilde{\mathbf{Z}}_1 \boldsymbol{\delta}_1 + \tilde{\mathbf{u}}_1 \tag{2.68}$$

gdzie: $\tilde{\mathbf{y}}_1 = \mathbf{Q}\mathbf{y}_1$, $\tilde{\mathbf{Z}}_1 = \mathbf{Q}\mathbf{Z}_1$, $\tilde{\mathbf{u}}_1 = \mathbf{Q}\mathbf{u}_1$. Parametry tego równania estymuje się za pomocą 2MNK, zastępując zmienne endogeniczne skorelowane ze składnikiem losowym przez $\tilde{\mathbf{X}} = \mathbf{Q}\mathbf{X}$. Uzyskany w ten sposób wewnątrzgrupowy estymator 2MNK (W2MNK) wektora $\boldsymbol{\delta}_1$ ma formę:

$$\tilde{\boldsymbol{\delta}}_{1,\mathrm{W2MNK}} = \left(\tilde{\mathbf{Z}}_1^{\mathrm{T}} \mathbf{P}_{\tilde{\mathbf{X}}} \tilde{\mathbf{Z}}_1\right)^{-1} \tilde{\mathbf{Z}}_1^{\mathrm{T}} \mathbf{P}_{\tilde{\mathbf{X}}} \tilde{\mathbf{y}}_1 \tag{2.69}$$

a jego wariancja wynosi: $\mathrm{D}^2(\tilde{\boldsymbol{\delta}}_{1,\mathrm{W2MNK}}) = \delta_{\varepsilon_{11}}^2 \left(\tilde{\mathbf{Z}}_1^{\mathrm{T}} \mathbf{P}_{\tilde{\mathbf{X}}} \tilde{\mathbf{Z}}_1\right)^{-1}$, gdzie $\mathbf{P}_{\tilde{\mathbf{X}}} = \tilde{\mathbf{X}}\left(\tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{X}}\right)^{-1}\tilde{\mathbf{X}}^{\mathrm{T}}$.

Międzygrupowy estymator 2MNK (M2MNK) uzyskuje się stosując 2MNK do równania (2.65) pomnożonego stronami przez operator międzygrupowy $\mathbf{P}$, a więc równania:

$$\bar{\mathbf{y}}_1 = \bar{\mathbf{Z}}_1 \boldsymbol{\delta}_1 + \bar{\mathbf{u}}_1 \tag{2.70}$$

gdzie: $\overline{\mathbf{y}}_1 = \mathbf{P}\mathbf{y}_1$, $\overline{\mathbf{Z}}_1 = \mathbf{P}\mathbf{Z}_1$, $\overline{\mathbf{u}}_1 = \mathbf{P}\mathbf{u}_1$. Jako instrumenty dla zmiennych endogenicznych przyjmuje się $\overline{\mathbf{X}} = \mathbf{P}\mathbf{X}$. Uzyskany w ten sposób estymator ma postać:

$$\hat{\boldsymbol{\delta}}_{1,\mathrm{M2MNK}} = \left(\overline{\mathbf{Z}}_1^{\mathrm{T}}\mathbf{P}_{\overline{\mathbf{X}}}\overline{\mathbf{Z}}_1\right)^{-1}\overline{\mathbf{Z}}_1^{\mathrm{T}}\mathbf{P}_{\overline{\mathbf{X}}}\overline{\mathbf{y}}_1 \tag{2.71}$$

zaś jego wariancja: $\mathrm{D}^2(\hat{\boldsymbol{\delta}}_{1,\mathrm{M2MNK}}) = \delta_{1_{11}}^2\left(\overline{\mathbf{Z}}_1^{\mathrm{T}}\mathbf{P}_{\overline{\mathbf{X}}}\overline{\mathbf{Z}}_1\right)^{-1}$, gdzie $\sigma_{1_{11}}^2 = T\sigma_{\alpha_{11}}^2 + \sigma_{\varepsilon_{11}}^2$.

Estymator dwustopniowej metody najmniejszych kwadratów dla modeli panelowych (2MNK-MP) (ang. *error component two-stage least squares estimator*), będący macierzowo ważoną średnią estymatorów (2.69) i (2.71) (por. B a l t a g i [1981]) ma postać:

$$\hat{\boldsymbol{\delta}}_{1,\mathrm{PM2MNK}} = \left[\frac{\tilde{\mathbf{Z}}_1^{\mathrm{T}}\mathbf{P}_{\tilde{\mathbf{X}}}\tilde{\mathbf{Z}}_1}{\sigma_{\varepsilon_{11}}^2} + \frac{\overline{\mathbf{Z}}_1^{\mathrm{T}}\mathbf{P}_{\overline{\mathbf{X}}}\overline{\mathbf{Z}}_1}{\sigma_{1_{11}}^2}\right]^{-1}\left[\frac{\tilde{\mathbf{Z}}_1^{\mathrm{T}}\mathbf{P}_{\tilde{\mathbf{X}}}\tilde{\mathbf{y}}_1}{\sigma_{\varepsilon_{11}}^2} + \frac{\overline{\mathbf{Z}}_1^{\mathrm{T}}\mathbf{P}_{\overline{\mathbf{X}}}\overline{\mathbf{y}}_1}{\sigma_{1_{11}}^2}\right] \tag{2.72}$$

Uzyskuje się go w wyniku zastosowania UMNK do estymacji systemu równań, postaci:

$$\begin{pmatrix}\tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{y}}_1\\ \overline{\mathbf{X}}^{\mathrm{T}}\overline{\mathbf{y}}_1\end{pmatrix} = \begin{pmatrix}\tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{Z}}_1\\ \overline{\mathbf{X}}^{\mathrm{T}}\overline{\mathbf{Z}}_1\end{pmatrix}\delta_1 + \begin{pmatrix}\tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{u}}_1\\ \overline{\mathbf{X}}^{\mathrm{T}}\overline{\mathbf{u}}_1\end{pmatrix} \tag{2.73}$$

gdzie: $\mathrm{E}\begin{pmatrix}\tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{u}}_1\\ \overline{\mathbf{X}}^{\mathrm{T}}\overline{\mathbf{u}}_1\end{pmatrix} = \mathbf{0}$ oraz $\mathrm{D}^2\begin{pmatrix}\tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{u}}_1\\ \overline{\mathbf{X}}^{\mathrm{T}}\overline{\mathbf{u}}_1\end{pmatrix} = \begin{bmatrix}\sigma_{\varepsilon_{11}}^2\tilde{\mathbf{X}}^{\mathrm{T}}\tilde{\mathbf{X}} & \mathbf{0}\\ \mathbf{0} & \sigma_{1_{11}}^2\overline{\mathbf{X}}^{\mathrm{T}}\mathbf{X}\end{bmatrix}$.

Pierwszy czynnik wzoru (2.72) określa wariancję estymatora PM2MNK.

Jak wspomniano powyżej, estymator (2.72) może być przedstawiony jako macierzowo ważona średnia estymatorów W2MNK i M2MNK, co zostało pokazane w: B a l t a g i [2008] (s. 122–124). Nieznane składowe wariancji $\sigma_{\nu_{11}}^2$ i $\sigma_{1_{11}}^2$ występujące we wzorze (2.72) zastąpić można zgodnymi estymatorami uzyskanymi odpowiednio na podstawie reszt W2MNK i M2MNK. Dokładniej zostało to opisane w: B a l t a g i [2008] (s. 122–124), a w języku polskim w pracy B. D a ń s k i e j [2000]. Przedstawiona tam też została alternatywna metoda uzyskania estymatora wektora $\boldsymbol{\delta}_1$.

W celu prezentacji możliwości łącznej estymacji systemu równań łącznie współzależnych (bazującej na 3MNK) zapiszmy ten system w postaci:

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\delta} + \mathbf{u} \tag{2.74}$$

gdzie $\mathbf{y} = \left(\mathbf{y}_1^{\mathrm{T}}, ..., \mathbf{y}_M^{\mathrm{T}}\right)$, $\mathbf{Z} = \mathrm{diag}\left(\mathbf{Z}_j\right)$, $\boldsymbol{\delta}^{\mathrm{T}} = \left(\boldsymbol{\delta}_1^{\mathrm{T}}, ..., \boldsymbol{\delta}_M^{\mathrm{T}}\right)$, $\mathbf{u}' = \left(\mathbf{u}_1^{\mathrm{T}}, ..., \mathbf{u}_M^{\mathrm{T}}\right)$, zaś $\mathbf{Z}_j = [\mathbf{Y}_j, \mathbf{X}_j]$ jest macierzą o wymiarach $[NT \times (g_j + k_j)]$ dla j = 1, ..., M. Oznacza to, że j-te równanie modelu zawiera g_j zmiennych endogenicznych i k_j zmiennych egzogenicznych.

Składnik losowy j-tego równania układu (2.74) ma postać: $\mathbf{u}_j = \mathbf{Z}_\alpha \boldsymbol{\alpha}_j + \boldsymbol{\varepsilon}_j$, a macierze wariancji-kowariancji są dane wzorami: (2.61) dla j-tego i i-tego równania oraz (2.62) dla całego układu (2.74). Ich forma jest więc identyczna, jak w przypadku modelu SUR. Mnożąc równanie (2.74) stronami przez $\mathbf{I}_M \otimes \mathbf{Q}$ uzyskuje się:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{Z}}\boldsymbol{\delta} + \tilde{\mathbf{u}} \tag{2.75}$$

gdzie: $\tilde{\mathbf{y}} = \left(\mathbf{I}_M \otimes \mathbf{Q}\right)\mathbf{y}$, $\tilde{\mathbf{Z}} = \left(\mathbf{I}_M \otimes \mathbf{Q}\right)\mathbf{Z}$, $\tilde{\mathbf{u}} = \left(\mathbf{I}_M \otimes \mathbf{Q}\right)\mathbf{u}$. Stosując do modelu (2.75) 3MNK ze zbiorem instrumentów $\mathbf{I}_M \otimes \tilde{\mathbf{X}}$ (gdzie $\tilde{\mathbf{X}} = \mathbf{QX}$), otrzymuje się tzw. wewnątrzgrupowy estymator trójstopniowej MNK (W3MNK) dany wzorem:

$$\tilde{\boldsymbol{\delta}}_{\mathrm{W3MNK}} = \left[\tilde{\mathbf{Z}}^{\mathrm{T}}\left(\boldsymbol{\Sigma}_\varepsilon^{-1} \otimes \mathbf{P}_{\tilde{\mathbf{X}}}\right)\tilde{\mathbf{Z}}\right]^{-1}\left[\tilde{\mathbf{Z}}^{\mathrm{T}}\left(\boldsymbol{\Sigma}_\varepsilon^{-1} \otimes \mathbf{P}_{\tilde{\mathbf{X}}}\right)\tilde{\mathbf{y}}\right] \tag{2.76}$$

Analogicznie uzyskuje się tzw. międzygrupowy estymator trójstopniowej MNK (M3MNK). Należy zastosować 3MNK ze zbiorem instrumentów $\mathbf{I}_M \otimes \overline{\mathbf{X}}$ (gdzie $\overline{\mathbf{X}} = \mathbf{PX}$) do estymacji modelu, powstającego z pomnożenia (2.74) przez $\mathbf{I}_M \otimes \mathbf{P}$, i mającego postać:

$$\overline{\mathbf{y}} = \overline{\mathbf{Z}}\boldsymbol{\delta} + \overline{\mathbf{u}} \tag{2.77}$$

gdzie $\overline{\mathbf{y}} = \left(\mathbf{I}_M \otimes \mathbf{P}\right)\mathbf{y}$, $\overline{\mathbf{Z}} = \left(\mathbf{I}_M \otimes \mathbf{P}\right)\mathbf{Z}$, $\overline{\mathbf{u}} = \left(\mathbf{I}_M \otimes \mathbf{P}\right)\mathbf{u}$.

Estymator M3MNK dany jest wzorem:

$$\hat{\boldsymbol{\delta}}_{\mathrm{M3MNK}} = \left[\overline{\mathbf{Z}}^{\mathrm{T}}\left(\boldsymbol{\Sigma}_1^{-1} \otimes \mathbf{P}_{\overline{\mathbf{X}}}\right)\overline{\mathbf{Z}}\right]^{-1}\left[\overline{\mathbf{Z}}^{\mathrm{T}}\left(\boldsymbol{\Sigma}_1^{-1} \otimes \mathbf{P}_{\overline{\mathbf{X}}}\right)\overline{\mathbf{y}}\right] \tag{2.78}$$

Estymator trójstopniowej metody najmniejszych kwadratów dla modeli panelowych (PM3MNK) (ang. *error component three-stage least squares estimator*), ma postać:

$$\hat{\boldsymbol{\delta}}_{\mathrm{PM3MNK}} = \left[\tilde{\mathbf{Z}}^{\mathrm{T}}\left(\boldsymbol{\Sigma}_{\varepsilon}^{-1} \otimes \mathbf{P}_{\tilde{\mathbf{X}}}\right)\tilde{\mathbf{Z}} + \overline{\mathbf{Z}}^{\mathrm{T}}\left(\boldsymbol{\Sigma}_{1}^{-1} \otimes \mathbf{P}_{\overline{\mathbf{X}}}\right)\overline{\mathbf{Z}}\right]^{-1} \cdot \\ \cdot \left[\tilde{\mathbf{Z}}^{\mathrm{T}}\left(\boldsymbol{\Sigma}_{\varepsilon}^{-1} \otimes \mathbf{P}_{\tilde{\mathbf{X}}}\right)\tilde{\mathbf{y}} + \overline{\mathbf{Z}}^{\mathrm{T}}\left(\boldsymbol{\Sigma}_{1}^{-1} \otimes \mathbf{P}_{\overline{\mathbf{X}}}\right)\overline{\mathbf{y}}\right] \tag{2.79}$$

Otrzymuje się go w wyniku estymacji UMNK systemu równań utworzonego z (2.75) i (2.77) pomnożonych odpowiednio przez $\mathbf{I}_M \otimes \tilde{\mathbf{X}}^{\mathrm{T}}$ i $\mathbf{I}_M \otimes \overline{\mathbf{X}}^{\mathrm{T}}$. Estymator (2.79) może być przedstawiony jako macierzowo ważona średnia estymatorów W3MNK i M3MNK. Nieznane składowe wariancji $\boldsymbol{\Sigma}_{\varepsilon}$ i $\boldsymbol{\Sigma}_1$, występujące w (2.79), zastąpić można (podobnie jak w metodzie dwustopniowej) zgodnymi estymatorami uzyskanymi odpowiednio na podstawie reszt W2MNK i M2MNK. Dokładniej zostało to opisane w: Baltagi [2008] (s. 130–132), a w języku polskim w pracy Dańskiej [2000], w których przedstawiono też alternatywną metodę uzyskania estymatora wektora $\boldsymbol{\delta}_1$.

ROZDZIAŁ 3

PROBLEMY ESTYMACJI I TESTOWANIA DYNAMICZNYCH MODELI PANELOWYCH

W poprzednich rozdziałach podkreślano, że dynamiczne modele panelowe wymagają zastosowania specyficznych metod estymacji. Zaprezentowanie tych metod oraz odpowiednich testów statystycznych jest zasadniczym celem niniejszego rozdziału. Prezentacja metod estymacji poprzedzona jest szczegółowym uzasadnieniem przyczyn utraty zgodności i nieobciążoności estymatorów stosowanych dla modeli statycznych, która ma miejsce, jeśli postać modelu jest dynamiczna. Własności estymatorów: wewnątrzgrupowego i międzygrupowego, w przypadku modelu dynamicznego, oraz problemy związane z szacowaniem efektów długookresowych i krótkookresowych w modelach panelowych są omówione w podrozdz. 3.1. Podrozdział 3.2 poświęcony jest prezentacji metod estymacji dynamicznych modeli panelowych, bazujących na uogólnionej metodzie momentów i metodzie zmiennych instrumentalnych. Są to metody mające największe znaczenie w praktyce, zaś pojawiające się w literaturze przedmiotu propozycje alternatywne, o których wspomniano w podrozdz. 1.6, wydają się mieć znaczenie znacznie mniejsze. Wyjątkiem może być propozycja zastosowania estymatora wewnątrzgrupowego ze skorygowanym sposobem obliczania wariancji estymatorów parametrów strukturalnych. Ze względu na relatywnie dużą częstotliwość zastosowań tej metody została ona przedstawiona dokładniej w pkt 3.2.7. Testy statystyczne, stosowane do oceny jakości dynamicznych modeli panelowych opisane są w podrozdz. 3.3. Końcowy podrozdz. 3.4 poświęcony jest porównaniu własności estymatorów, przedstawionych w podrozdz. 3.2.

3.1. Panelowe modele statyczne a dynamiczne – własności estymatorów

Estymator wewnątrzgrupowy dla statycznych modeli panelowych jest zgodny, niezależnie od tego, czy efekty grupowe są losowe (model RE), czy nielosowe (model FE). W przypadku modelu dynamicznego estymator ten traci zgodność, przy czym dotyczy to zarówno modelu FE, jak i RE. Wynika to z faktu, że opóźniona zmienna objaśniana jest zależna od stałych w czasie

efektów grupowych α_i, a dokładniej, że transformowany za pomocą operatora wewnątrzgrupowego wektor obserwacji na opóźnionej zmiennej objaśnianej[1] jest skorelowany z transformowanym w ten sam sposób wektorem składników losowych $\boldsymbol{\varepsilon}$.

Poniżej pokazana zostanie przyczyna utraty zgodności estymatora wewnątrzgrupowego na przykładzie modelu autoregresyjnego. Model taki ma postać:

$$y_{it} = \gamma y_{i,t-1} + \alpha_i + \varepsilon_{it} \tag{3.1}$$

Postać ogólna estymatora wewnątrzgrupowego $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ dana jest wzorem (2.9) w rozdz. 2. Dla parametru γ modelu (3.1) miałby on postać:

$$\hat{\gamma}_{\mathrm{W}} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it} - \bar{y}_i)(y_{i,t-1} - \bar{y}_{i,-1})}{\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{i,t-1} - \bar{y}_{i,-1})^2} \tag{3.2}$$

gdzie $\bar{y}_i$ oraz $\bar{y}_{i,-1}$ oznaczają średnie po czasie odpowiednio z wartości y_{it}, $y_{i,t-1}$.

Podstawiając (3.1) do powyższego wzoru, otrzymujemy:

$$\hat{\gamma}_{\mathrm{W}} = \gamma + \frac{\frac{1}{NT}\sum_i\sum_t(\varepsilon_{it} - \bar{\varepsilon}_i)(y_{i,t-1} - \bar{y}_{i,-1})}{\frac{1}{NT}\sum_i\sum_t(y_{i,t-1} - \bar{y}_{i,-1})^2} \tag{3.3}$$

Jeśli $N \to \infty$, a T jest skończone, to estymator ten nie jest zgodny ani nieobciążony, ponieważ wyrażenie w liczniku wzoru (3.3) nie dąży wtedy do zera. N i c k e l l [1981] wykazał, że:

$$\operatorname*{p\,lim}_{N\to\infty}\frac{1}{NT}\sum_i\sum_t(\varepsilon_{it} - \bar{\varepsilon}_i)(y_{i,t-1} - \bar{y}_{i,-1}) = -\frac{\sigma_\varepsilon^2}{T^2}\cdot\frac{(T-1) - T\gamma + \gamma^T}{(1-\gamma)^2} \neq 0 \tag{3.4}$$

gdzie σ_ε^2 oznacza wariancję składnika losowego ε_{it}.

---

[1] Wynikiem takiej transformacji jest wektor, którego elementami są odchylenia od średnich grupowych; i-ta współrzędna tego wektora ma postać: $y_{i,t-1} - \bar{y}_{i.}$, gdzie $\bar{y}_{i.}$ oznacza średnią arytmetyczną zmiennej y po czasie.

Z zależności (3.4) wynika, że estymator (3.3) byłby zgodny i nieobciążony jedynie wtedy, gdy $T \to \infty$. Próby wykorzystywane w praktyce charakteryzują się jednak zazwyczaj niewielkimi wartościami T.

Wynik uzyskany przez S. Nickella zapisać można inaczej (por. np. H s i a o [2003], s. 72):

$$\underset{N\to\infty}{p\lim}(\hat{\gamma}_{\mathrm{W}} - \gamma) = \frac{1+\gamma}{T-1}\left[1 - \frac{1-\gamma^T}{T(1-\gamma)}\right]\left[1 - \frac{2\gamma}{(T-1)(1-\gamma)}\left(1 - \frac{1-\gamma^T}{T(1-\gamma)}\right)\right]^{-1} \tag{3.5}$$

Ze wzoru (3.5) widać, że estymator wewnątrzgrupowy jest niezgodny $\mathrm{O}(T^{-1})$ i asymptotyczne obciążenie jest ujemne dla dodatnich wartości γ. Ponadto wielkość tego obciążenia nie zależy od wartości σ_ε^2.

Dla modelu z niezależnymi zmiennymi objaśniającymi, postaci:

$$y_{it} = \gamma y_{i,t-1} + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_{it} + \alpha_i + \varepsilon_{it} \tag{3.6}$$

P. S e v e s t r e i A. T r o g n o n [1985] wykazali, że asymptotyczne obciążenie estymatora $\hat{\gamma}_{\mathrm{W}}$ zmienia się w ten sposób, że we wzorze (3.5) prawą stronę należy pomnożyć przez:

$$\left[\underset{N\to\infty}{p\lim}\frac{1}{NT}\mathbf{y}_{-1}^{\mathrm{T}}\mathbf{A}\mathbf{y}_{-1}\right]\left[\underset{N\to\infty}{p\lim}\frac{1}{NT}\mathbf{y}_{-1}^{\mathrm{T}}\mathbf{A}\mathbf{M}\mathbf{A}\mathbf{y}_{-1}\right]^{-1} \tag{3.7}$$

gdzie $\mathbf{M} = \mathbf{I} - \mathbf{A}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{A}$, $\mathbf{A} = \mathbf{I}_N \otimes \mathbf{A}_T$, $\mathbf{A}_T = \mathbf{I}_T - \frac{1}{T}\mathbf{i}_T\mathbf{i}_T^{\mathrm{T}}$, $\mathbf{i}_T$ oznacza wektor jedynek o T współrzędnych, $\mathbf{y}_{-1}$ jest wektorem opóźnionych zmiennych objaśnianych: $\mathbf{y}_{-1} = [y_{10}, ..., y_{1,T-1}, ..., y_{N0}, ..., y_{N,T-1}]^{\mathrm{T}}$. Obciążenie estymatora wektora $\boldsymbol{\beta}$ parametrów przy zmiennych egzogenicznych modelu (3.6) wynosi zaś:

$$\underset{N\to\infty}{p\lim}(\hat{\boldsymbol{\beta}}_{\mathrm{W}} - \boldsymbol{\beta}) = -\left[\underset{N\to\infty}{p\lim}(\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{A}\mathbf{y}_{-1}\right]\underset{N\to\infty}{p\lim}(\hat{\gamma}_{\mathrm{W}} - \gamma) \tag{3.8}$$

M. H. P e s a r a n i R. S m i t h [1995] pokazali, że jeżeli w rzeczywistości obiekty są heterogeniczne i parametry strukturalne modelu różnią się względem obiektów (wektorem parametrów strukturalnych modelu (3.6) jest $\boldsymbol{\beta}_i$, gdzie $i = 1, ..., N$), to estymator wewnątrzgrupowy jest niezgodny również w przypadku, gdy jednocześnie $T \to \infty$ i $N \to \infty$.

Własności estymatora KMNK parametru γ modelu autoregresyjnego postaci (3.1) zależą natomiast od założeń dotyczących efektów grupowych.

W przypadku statycznego modelu z efektami losowymi (RE), przy założeniu, że zmienne objaśniające są egzogeniczne i nieskorelowane z efektami grupowymi, estymator KMNK, choć nieefektywny, jest zgodny i nieobciążony. Natomiast w przypadku dynamicznym estymator KMNK parametru γ traci zgodność, co spowodowane jest korelacją opóźnionej zmiennej objaśnianej z efektami grupowymi. Estymator ten ma postać:

$$\hat{\gamma}_{\mathrm{KMNK}} = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it} \cdot y_{i,t-1}}{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,t-1}^2} = \gamma + \frac{\sum_{i=1}^{N}\sum_{t=1}^{T} (\alpha_i + \varepsilon_{it}) y_{i,t-1}}{\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,t-1}^2} \tag{3.9}$$

Asymptotyczne obciążenie estymatora $\hat{\gamma}_{\mathrm{KMNK}}$ dla $N \to \infty$ jest określone jako granica według prawdopodobieństwa drugiego składnika po prawej stronie wzoru (3.5). Można wykazać (por. np. H s i a o [2003], s. 74), że:

$$\underset{N\to\infty}{p\lim} \frac{1}{NT} \sum_{i=1}^{N}\sum_{t=1}^{T} (\alpha_i + \varepsilon_{it}) y_{i,t-1} = \frac{1-\gamma^T}{T(1-\gamma)} \operatorname{cov}(y_{i0}, \alpha_i) + \frac{\sigma_\alpha^2}{T(1-\gamma)^2} [(T-1) - T\gamma + \gamma^T] \tag{3.10}$$

oraz

$$\underset{N\to\infty}{p\lim} \frac{1}{NT} \sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,t-1}^2 = \frac{1-\gamma^{2T}}{T(1-\gamma^2)} \cdot \sum_{t=1}^{T} \frac{y_{i0}^2}{N} + \frac{\sigma_\alpha^2}{(1-\gamma)^2} \cdot \frac{1}{T}\left(T - 2\frac{1-\gamma^T}{1-\gamma} + \frac{1-\gamma^{2T}}{1-\gamma^2} \right) +$$
$$+ \frac{2}{T(1-\gamma)} \left(\frac{1-\gamma^T}{1-\gamma} - \frac{1-\gamma^{2T}}{1-\gamma^2} \right) \operatorname{cov}(y_{i0}, \alpha_i) + \frac{\sigma_\varepsilon^2}{T(1-\gamma^2)^2} [(T-1) - T\gamma + \gamma^T] \tag{3.10a}$$

gdzie σ_ε^2 oznacza wariancję składnika losowego ε_{it}, a σ_α^2 – wariancję efektów α_i. Zazwyczaj zakłada się, że wartości początkowe y_{i0} są arbitralnie określonymi stałymi lub są generowane przez ten sam proces, który generuje pozostałe wartości y_{it}, zatem kowariancja $\operatorname{cov}(y_{i0}, \alpha_i)$ jest nieujemna i estymator $\hat{\gamma}_{\mathrm{KMNK}}$ jest obciążony w górę. Jeśli założy się jedynie, że wartości początkowe są ograniczone, tzn. że $\underset{N\to\infty}{p\lim}\, y_{i0}^2 / N$ jest skończona, to $\hat{\gamma}_{\mathrm{KMNK}}$ jest obciążony

w górę, jeśli $N \to \infty$ lub $T \to \infty$. Obciążenie to rośnie ze wzrostem wariancji σ_α^2. Opisane wyżej rezultaty asymptotyczne są prawdziwe również w małych próbach, o czym świadczą wyniki eksperymentów Monte Carlo przeprowadzonych przez N e r l o v e'a [1967].

Dołączenie do modelu niezależnych zmiennych objaśniających (model (3.6)) powoduje nieznaczną redukcję obciążenia estymatora $\hat{\gamma}_{\mathrm{KMNK}}$. Oceny KMNK parametrów $\boldsymbol{\beta}$ są natomiast bliższe zeru niż rzeczywiste wartości tych parametrów.

Powyższe rozważania na temat estymatora KMNK dotyczyły modelu RE. W przypadku dynamicznego modelu FE postaci (3.6), jeśli $T \to \infty$, to opóźniona zmienna objaśniana y_{it-1} spełnia warunek słabej egzogeniczności: $\mathrm{E}(u_{it} \mid y_{i,t-1}, y_{i,t-2}, ...; \alpha_i) = 0$. Zatem regresja KMNK zmiennej y_{it} na $y_{i,t-1}$, $\mathbf{x}_{it}$ i specyficzne dla obiektów stałe α_i daje estymator zgodny. Dla skończonych wartości T argumentacja jest natomiast analogiczna, jak w przypadku modeli RE.

Formuły, określające asymptotyczne obciążenie estymatorów KMNK dla modelu autoregresyjnego rzędu p wyprowadzone są w pracy A. T r o g n o n a [1978]. Kierunek obciążenia jest jednak w przypadku takiego modelu niemożliwy do określenia *a priori*.

Problem estymacji efektów długookresowych na podstawie modeli panelowych rozważany jest w pracy A. P i r o t t e'a [1999]. Sytuacja w przypadku takich modeli różni się od przypadku modeli szeregów czasowych przede wszystkim ze względu na niewielki wymiar czasowy próby. Krótkość szeregów czasowych powoduje, że w przypadku modeli panelowych nie może być zastosowana procedura E n g l e'a i G r a n g e r a [1987]. Ta dwuetapowa procedura stosowana do estymacji modeli szeregów czasowych umożliwia oszacowanie efektów długookresowych na podstawie modelu statycznego, a efekty krótkookresowe są szacowane na podstawie modelu korekty błędem (ang. *error correction model*, ECM). Dla modeli panelowych B. H. B a l t a g i i J. M. G r i f f i n [1983, 1984] pokazali, że efekty długookresowe można oszacować na podstawie regresji międzygrupowej, tzn. szacując model statyczny na podstawie średnich grupowych (model (2.4), rozdz. 2). Estymatorem wektora parametrów strukturalnych $\boldsymbol{\beta}$ modelu (2.4) jest tzw. estymator międzygrupowy $\hat{\boldsymbol{\beta}}_{\mathrm{B}}$, dany wzorem (2.23). Jest to możliwe również w przypadku modelu heterogenicznego, z wektorem parametrów strukturalnych postaci $\boldsymbol{\beta}_i$, gdzie $i = 1, ..., N$, dla $N \to \infty$ i $T \to \infty$, co pokazali P e s a r a n i S m i t h [1995]. Pirotte [1999] rozważa przypadek gdy $N \to \infty$, T jest skończone, a parametry strukturalne są jednakowe względem obiektów. Wykazał on, że jeśli poprawną specyfikacją modelu jest model dynamiczny z efektami losowymi, to granicą stochastyczną estymatora międzygrupowego $\hat{\boldsymbol{\beta}}_{\mathrm{B}}$ modelu statycznego jest ocena efektu długookresowego. Granicą stochastyczną estymatora wewnątrzgrupowego $\hat{\boldsymbol{\beta}}_{\mathrm{W}}$ modelu statycznego jest zaś ocena efektów krótkookresowych. Zatem w modelu

dynamicznym efekty długookresowe można szacować na podstawie statycznej regresji międzygrupowej, a efekty krótkookresowe – na podstawie statycznej regresji wewnątrzgrupowej. Wyniki uzyskane przez A. P i r o t t e'a [1999] nie zależą od założeń dotyczących macierzy wariancji-kowariancji składnika losowego w modelu statycznym, zatem są one prawdziwe również dla modeli z autokorelacją i z efektami losowymi. P. E g g e r i M. P f a f f e r m a y r [2004] pokazali, że asymptotyczne obciążenie estymatorów: międzygrupowego i wewnątrzgrupowego jako estymatorów efektów długookresowych i krótkookresowych zależy od własności procesu generującego dane, długości szeregu czasowego i przekrojowego zróżnicowania wartości zmiennych objaśniających.

3.2. Metody estymacji dynamicznych modeli panelowych

Dynamiczny model panelowy ma postać:

$$y_{it} = \gamma y_{i,t-1} + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + u_{it} = \gamma y_{i,t-1} + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + \alpha_i + \varepsilon_{it}, \quad i = 1, ..., N, t = 1, ..., T \quad (3.11)$$

gdzie $\varepsilon_{it} \sim \mathrm{N}(0, \sigma_\varepsilon^2)$ dla każdego i, t, α_i – efekty grupowe; jeśli α_i są losowe, to $\alpha_i \sim \mathrm{N}(0, \sigma_\alpha^2)$, $\mathbf{x}_{it} = [x_{kit}]_{K\times 1}$ – wektor zmiennych objaśniających o K współrzędnych, $\boldsymbol{\beta}$ – wektor parametrów ($K \times 1$), jednakowych dla wszystkich i oraz t.

Jak wspomniano w podrozdz. 3.1, do estymacji panelowych modeli dynamicznych należy stosować odrębne metody, inne niż stosuje się dla modeli statycznych. Najważniejsze propozycje takich metod, przedstawiane we współczesnej literaturze przedmiotu bazują na uogólnionej metodzie momentów (GMM). Pojawiają się też propozycje stosowania metod opartych na metodzie zmiennych instrumentalnych lub metodzie największej wiarygodności oraz korekt niwelujących obciążenie estymatorów stosowanych dla modeli statycznych. Wydaje się jednak, że praktyczne znaczenie tych metod jest stosunkowo niewielkie. GMM jest metodą najbardziej uniwersalną, niewymagającą przyjmowania zbyt wielu założeń i relatywnie łatwą do praktycznego zastosowania.

Metody bazujące na GMM mogą być wykorzystywane m. in. do estymacji modeli autoregresyjnych z rozkładem opóźnień (ang. *autoregressive-distribured lag model*, ADL), Jako szczególny przypadek takiego modelu potraktować można model dynamiczny, estymowany na podstawie danych panelowych. S. B o n d [2002] (s. 26) stwierdza, że GMM jest szczególnie użyteczna do estymacji modeli, które zawierają endogeniczne lub z góry ustalone zmienne objaśniające, a proces generujący szeregi czasowe nie jest w pełni specyfiko-

walny. Wydaje się, że powyższe stwierdzenie, autorstwa jednego z najbardziej uznanych autorytetów w dziedzinie metod estymacji panelowych modeli dynamicznych, wyjaśnia popularność, jaką zyskały metody bazujące na GMM w rozważanym tu kontekście[2].

Postać i liczba warunków momentów, możliwych do wykorzystania w procesie estymacji modelu (3.11) za pomocą GMM, zależy od założeń, dotyczących korelacji między zmiennymi $\mathbf{x}_{it}$ a składowymi α_i i ε_{it}. Przy założeniu, że ε_{it} nie wykazuje korelacji w czasie oraz $\mathbf{x}_{it}$ są skorelowane z α_i, możliwe jest przyjęcie trzech alternatywnych założeń na temat korelacji $\mathbf{x}_{it}$ z ε_{it}. Zmienne $\mathbf{x}_{it}$ mogą być:

– endogeniczne, tzn. $\mathbf{x}_{it}$ jest skorelowany z wartością bieżącą ε_{it} i wartościami opóźnionymi $\varepsilon_{i,t-s}$, ale nieskorelowana z wartościami przyszłymi $\varepsilon_{i,t+s}$,

– z góry ustalone, tzn. $\mathbf{x}_{it}$ jest nieskorelowany z wartością bieżącą ε_{it}, ale skorelowany z wartościami opóźnionymi $\varepsilon_{i,t-s}$,

– ściśle egzogeniczne, tzn. $\mathbf{x}_{it}$ jest nieskorelowany z wartością bieżącą ε_{it}, wartościami opóźnionymi $\varepsilon_{i,t-s}$ i z wartościami przyszłymi $\varepsilon_{i,t+s}$.

Możliwe jest też przyjęcie dodatkowych, silniejszych założeń o niewystępowaniu korelacji jednoczesnej: $\mathrm{E}(\mathbf{x}_{it}u_{it}) = \mathbf{0}$ dla t = 1, ..., T. Wówczas w zależności od tego, czy $\mathbf{x}_{it}$ jest wektorem zmiennych z góry ustalonych, czy ściśle egzogenicznych wskazać można dodatkowe zmienne instrumentalne, a więc dodatkowe warunki momentów.

Kolejne dodatkowe warunki momentów są spełnione, jeśli założy się, że $\mathbf{x}_{it}$ są nieskorelowane z α_i albo, jak w pracy M. A r e l l a n o i O. B o v e r [1995], że poziomy zmiennych $\mathbf{x}_{it}$ są skorelowane z α_i, ale pierwsze różnice $\Delta\mathbf{x}_{it}$ i efekty grupowe α_i nie są skorelowane.

Wybór między powyższymi alternatywnymi założeniami wydaje się arbitralny. Ponieważ jednak w większości przypadków dodatkowe warunki momentów są tzw. warunkami ponadidentyfikującymi (ang. *overidentyfying restrictions*), to ich właściwość można badać za pomocą różnicowego testu Sargana. Jest on omówiony w podrozdz. 3.3.

Do wymienionych powyżej założeń, różni autorzy dołączają jeszcze dodatkowe założenia, np. założenia dotyczące warunków początkowych lub słabej stacjonarności. W ten sposób uzyskuje się alternatywne estymatory dla panelowych modeli dynamicznych, spośród których te o największym znaczeniu praktycznym omówione zostaną w kolejnych podrozdziałach.

W podrozdz. 3.2.1 omówiona jest metoda A n d e r s o n a i H s i a o [1981, 1982], oparta na metodzie zmiennych instrumentalnych. Pozostałe propozycje bazują na GMM.

---

[2] Uogólniona metoda momentów w wersji dla szeregów czasowych została zaproponowana przez L. P. H a n s e n a [1982]. Jej idea została przedstawiona w zarysie w dodatku II niniejszej publikacji.

GMM pierwszych różnic (ang. *first-differenced GMM*, FDGMM), zaproponowana przez M. A r e l l a n o i O. B o n d a [1991] omówiona jest w pkt 3.2.2. Zasadniczą ideę FDGMM przedstawić można następująco: oblicza się pierwsze różnice modelu w celu usunięcia stałych w czasie efektów grupowych α_i, a następnie zmienne objaśniające w modelu pierwszych różnic zastępuje instrumentami, którymi są poziomy zmiennych, opóźnione o dwa lub więcej okresów. Estymatory parametrów strukturalnych uzyskuje się stosując GMM do modelu pierwszych różnic. Możliwość zastosowania tej metody może być problematyczna w przypadku, kiedy opóźnione poziomy zmiennych są słabymi instrumentami dla zmiennych zróżnicowanych (tzn. zmienne instrumentalne są zbyt słabo skorelowane ze zmienną objaśniającą). Estymatory FDGMM mogą być wówczas silnie obciążone[3]. Sytuacja taka ma miejsce m. in. wtedy, gdy liczba obserwacji w czasie jest mała i wykorzystuje się tzw. szeregi czasowe o wysokim stopniu trwałości (ang. *persistent time series*). W procesie konstrukcji i estymacji dynamicznych modeli panelowych wymieniona cecha szeregów czasowych jest wyjątkowo istotna, ponadto wydaje się, że pojęcie to jest rzadko spotykane w innych kontekstach. Dlatego poniżej podana jest definicja takiego szeregu.

Definicja. Szeregiem czasowym o wysokim stopniu trwałości nazywa się szereg, którego przyszłe wartości są silnie skorelowane z wartościami bieżącymi (ang. *A highly persistent time series is a time series process, where outcomes in the distant future are highly correlated with current outcomes*).

R. B l u n d e l l i S. B o n d [1998] rozważali problem słabych instrumentów w kontekście panelowych modeli dynamicznych i zaproponowali tzw. systemowy estymator GMM (ang. *system GMM*, SGMM). Praca ta jest omówiona w pkt 3.2.6. Estymator systemowy wykorzystuje założenia dotyczące wartości początkowych, umożliwiające uzyskanie warunków momentów, które są właściwe również w przypadku szeregów o wysokim stopniu trwałości. Zasadnicza idea SGMM polega na oszacowaniu systemu równań zarówno na przyrostach, jak i na poziomach. Instrumentami dla zmiennych objaśniających w równaniach na poziomach są opóźnione pierwsze różnice zmiennych. Instrumenty te są właściwe przy założeniu, że spełnione jest założenie słabej stacjonarności dla warunków początkowych. Idea SGMM jest rozwinięciem sugestii przedstawionych przez M. A r e l l a n o i O. B o v e r [1995], dotyczących wykorzystania dodatkowych warunków momentów dla równań na poziomach i wskazania właściwych instrumentów. Praca M. A r e l l a n o, O. B o v e r [1995] omówiona jest w pkt 3.2.4. Dodatkowe warunki momentów, które mogą być wykorzystywane do estymacji dynamicznych modeli panelowych, jeśli liczba

[3] Uwagi na temat obciążenia estymatorów, spowodowanego zastosowaniem słabych instrumentów, znaleźć można m. in. w: S t a i g e r, S t o c k [1997]; B o b b a, C o v i e l l o [2007].

obiektów $N \to \infty$, a wymiar czasowy próby T jest skończony, wskazali również S. C. Ahn i P. Schmidt [1995]. Ich wykorzystanie poprawia efektywność estymatora FDGMM. Warunki te są nieliniowe, ale autorzy wskazują możliwość przekształcenia ich do warunków liniowych. Omówienie pracy S. C. Ahna i P. Schmidta [1995] zawarte jest w pkt 3.2.5.

Punkt 3.2.3 zawiera opis propozycji F. Windmeijera [2005], dotyczącej sposobu korygowania obciążenia estymatorów wariancji dwustopniowych estymatorów GMM, które jest szczególnie znaczne w małych próbach.

Punkt 3.2.7 zawiera natomiast opis propozycji J. F. Kivieta [1995]. Polega ona na zastosowaniu do estymacji dynamicznego modelu panelowego estymatora wewnątrzgrupowego, z uwzględnieniem wyznaczonej przez autora poprawki, korygującej jego obciążenie.

3.2.1. Estymator metody zmiennych instrumentalnych Andersona–Hsiao

T. W. Anderson i C. Hsiao w dwóch kolejnych artykułach z lat 1981 i 1982 przedstawili ideę wykorzystania do estymacji dynamicznych modeli panelowych metody zmiennych instrumentalnych (IV). Artykuł z roku 1981 dotyczy modelu autoregresyjnego. Rozszerzenie rozważań na model zawierający dodatkowe zmienne objaśniające przedstawione zostało w pracy: Anderson, Hsiao [1982]. Metoda zmiennych instrumentalnych, znana obecnie jako metoda Andersona–Hsiao, została przez autorów zaproponowana jako praktyczna alternatywa dla stanowiącej również przedmiot ich rozważań metody największej wiarygodności (MNW). Dwie, jak się wydaje najważniejsze w kontekście późniejszych badań sugestie, zawarte w pracach T. W. Andersona i C. Hsiao [1981], [1982], to podkreślenie wagi warunków początkowych w procesie estymacji dynamicznych modeli panelowych, oraz wskazanie możliwości wyeliminowania z modelu efektów grupowych α_i poprzez obliczenie pierwszych różnic.

Anderson i Hsiao [1981] (w dalszym ciągu AH81) rozważali model autoregresyjny postaci:

$$y_{it} = \gamma y_{i,t-1} + u_{it} = \gamma y_{i,t-1} + \alpha_i + \varepsilon_{it}, \quad \text{gdzie } |\gamma| < 1 \tag{3.12}$$

Przyjmując różne założenia dotyczące wartości początkowych y_{i0}, badali oni własności estymatora MNW i estymatora wewnątrzgrupowego (WG). AH81 wyróżnili cztery przypadki:

1) wartości początkowe y_{i0} są nielosowe, obserwowalne i spełniają warunek:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}y_{i0}^2<\infty,$$

2) wartości początkowe są nielosowe, niezależne od α_i i ε_{it}, oraz mają jednakową średnią c i wariancję σ_0^2,

3) wartości początkowe są nielosowe, niezależne od α_i i ε_{it}, oraz mają różne średnie i jednakową wariancję σ_0^2.

4) wartości początkowe są nielosowe, stacjonarne,

Wyznaczenie łącznej funkcji wiarygodności dla i-tego obiektu (w celu zastosowania estymatora MNW) jest najprostsze w przypadku 2). Wyznacza się ją w zależności od funkcji gęstości dla próby i od wartości początkowych:

$$f_i(y_{iT},...,y_{i1},y_{i0})=f_i(y_{iT},...,y_{i1}|y_{i0})\,f_i(y_{i0}).$$

Wartości c i σ_0^2 estymuje się jednocześnie z wartościami γ, σ_α^2 (wariancją efektów grupowych) i σ_ε^2 (wariancją składnika losowego). W przypadkach 3) i 4) wyznaczanie funkcji wiarygodności komplikuje się. Dla przykładu, w przypadku 4) jej postać jest następująca:

$$f_i(y_{iT},...,y_{i1},y_{i0})=\left(\frac{1}{2\pi}\right)^{(T+1)/2}\left(\frac{1}{\rho\sigma^2}\right)^{(T+1)/2}|\mathbf{\Omega}|^{1/2}$$

$$\exp\left\{-\frac{1}{\rho\sigma^2}(y_{i0,}y_{i1}-\gamma y_{i0},...,y_{iT}-\gamma y_{i,T-1})^{\mathrm{T}}\mathbf{\Omega}^{-1}(y_{i0,}y_{i1}-\gamma y_{i0},...,y_{iT}-\gamma y_{i,T-1})\right\},$$

gdzie $\Omega=\frac{1-\rho}{\rho}\begin{bmatrix}\frac{1}{1-\gamma^2} & \mathbf{0}^{\mathrm{T}}\\ \mathbf{0} & \mathbf{I}_T\end{bmatrix}+\begin{bmatrix}\frac{1}{1-\gamma}\\ \mathbf{1}_T\end{bmatrix}\begin{bmatrix}\frac{1}{1-\gamma} & \mathbf{1}_T^{\mathrm{T}}\end{bmatrix}$, $\mathbf{0}$ i $\mathbf{1}$ oznaczają wektory składające się odpowiednio z zer lub jedynek, ρ oznacza współczynnik korelacji międzygrupowej, $\mathrm{E}(\mathbf{uu}^{\mathrm{T}})=\sigma^2\mathbf{\Omega}$, $\mathbf{u}=[u_{11},...,u_{1T},...,u_{N1},...,u_{NT}]^{\mathrm{T}}$.

W przypadku estymatora WG, wartości początkowe y_{i0} traktuje się jako nieznane stałe. Własności różnych estymatorów, w zależności od założeń dotyczących y_{i0}, zestawione są w tabl. 3.1.

T a b l i c a 3.1

Zgodność estymatorów MNW i FE w zależności od założeń, dotyczących y_{i0}

Założenie o y_{i0}	Estymator MNW		Estymator FE	
	$T \to \infty$, N skończone	$N \to \infty$, T skończone	$T \to \infty$, N skończone	$N \to \infty$, T skończone
1)	zgodny	zgodny	zgodny	niezgodny
2)	zgodny	zgodny	zgodny	niezgodny
3)	zgodny	niezgodny	zgodny	niezgodny
4)	zgodny	zgodny	zgodny	niezgodny

Ź r ó d ł o: AH81, za: N e r l o v e [2000].

Z tab. 3.1 wynika w szczególności, że jeśli wymiar czasowy próby T jest skończony, to estymator WG nie powinien być stosowany. Estymator ten jest zaś zgodny i nieobciążony, jeśli $T \to \infty$, co ma jednak małe znaczenie praktyczne, ponieważ zazwyczaj dane panelowe charakteryzują się niewielką liczbą obserwacji w czasie. Z tab. 3.1 wynika również, że dla skończonych wartości T i $N \to \infty$ estymatory MNW są zgodne, za wyjątkiem przypadku 3). Wynik ten przemawiałby zatem za estymowaniem modeli dynamicznych za pomocą MNW.

Jest ona jednak żmudna i niedogodna w zastosowaniach praktycznych. Ponadto założenia dotyczące wartości początkowych są trudne lub niemożliwe do zweryfikowania w praktyce. Jako praktyczną alternatywę AH81 proponują oszacowania parametru γ modelu (3.12) za pomocą metody zmiennych instrumentalnych (IV). Zastosowanie tej metody wymaga przyjęcia następujących założeń: $\mathrm{E}(\varepsilon_{it}) = 0$, $\mathrm{E}(\varepsilon_{it}\varepsilon_{js}) = \sigma_{\varepsilon}^2$ dla $i = j$, $t = s$, $\mathrm{E}(\varepsilon_{it}\varepsilon_{js}) = 0$ w pozostałych przypadkach, efekty grupowe są nielosowe.

Na wstępie, w celu wyeliminowania z modelu efektów grupowych, AH81 obliczają pierwsze różnice modelu (3.12). Otrzymuje się wówczas model postaci:

$$\left(y_{it} - y_{i,t-1}\right) = \gamma\left(y_{i,t-1} - y_{i,t-2}\right) + \left(\varepsilon_{it} - \varepsilon_{i,t-1}\right) \tag{3.13}$$

Ponieważ w modelu (3.13) zmienna objaśniająca $\Delta y_{i,t-1} = (y_{i,t-1} - y_{i,t-2})$ jest skorelowana ze składnikiem losowym $(\varepsilon_{it} - \varepsilon_{i,t-1})$, AH81 proponują zastosowanie metody IV. Jako możliwe instrumenty wskazują oni dwie zmienne: $y_{i,t-2}$ oraz $\Delta y_{i,t-2} = (y_{i,t-2} - y_{i,t-3})$. Obie one są skorelowane z $y_{i,t-1} - y_{i,t-2}$, ale nieskorelowane z $\Delta\varepsilon_{it} = (\varepsilon_{it} - \varepsilon_{i,t-1})$, ponieważ z założeń modelu (3.12) wynika, że ε_{it} nie wykazują autokorelacji.

Estymatory parametru γ mają, w zależności od doboru zmiennej instrumentalnej, postać:

$$\hat{\gamma}_{\mathrm{AH}} = \begin{cases} \dfrac{\sum_{i=1}^{N}\sum_{t=3}^{T}\left(y_{it}-y_{i,t-1}\right)\left(y_{i,t-2}-y_{i,t-3}\right)}{\sum_{i=1}^{N}\sum_{t=3}^{T}\left(y_{i,t-1}-y_{i,t-2}\right)\left(y_{i,t-2}-y_{i,t-3}\right)} & \text{dla zmiennej instrumentalnej } \Delta y_{i,t-2} \\ \dfrac{\sum_{i=1}^{N}\sum_{t=2}^{T}\left(y_{it}-y_{i,t-1}\right)y_{i,t-2}}{\sum_{i=1}^{N}\sum_{t=2}^{T}\left(y_{i,t-1}-y_{i,t-2}\right)y_{i,t-2}} & \text{dla zmiennej instrumentalnej } y_{i,t-2} \end{cases} \tag{3.14}$$

Estymatory (3.14), nazywane estymatorami Andersona–Hsiao, są zgodne, jeśli $N \to \infty$ lub $T \to \infty$. Ich własności są ponadto niezależne od założeń o wartościach początkowych.

W późniejszej chronologicznie literaturze (por. A r e l l a n o [1989]) zaleca się stosowanie jako instrumentu dla $\Delta y_{i,t-1}$ opóźnionego o dwa okresy poziomu zmiennej, tzn. $y_{i,t-2}$. Przyczyną jest to, że zastosowanie $\Delta y_{i,t-2}$ jako zmiennej instrumentalnej może wpłynąć na bardzo duży wzrost wariancji ocen parametrów. Ponadto obliczenie różnicy $(y_{i,t-2} - y_{i,t-3})$ skraca wymiar czasowy próby o 3 okresy, a zastosowanie opóźnionego poziomu $y_{i,t-2}$ – tylko o 2 okresy. Jest to istotne szczególnie w przypadku, gdy wymiar czasowy próby jest niewielki. Jeśli $T \geq 3$, to wyboru instrumentu można również dokonać badając korelację między $(y_{i,t-1} - y_{i,t-2})$ a $y_{i,t-2}$ oraz między $(y_{i,t-1} - y_{i,t-2})$ a $(y_{i,t-2} - y_{i,t-3})$. W pracy AH81 znaleźć można ponadto porównanie asymptotycznej efektywności obu estymatorów, danych wzorem (3.14).

W artykule z 1982 r. T. W. A n d e r s o n i C. H s i a o (dalej AH82) rozszerzają omówione powyżej rozważania na przypadek, gdy w modelu (3.12) występują dodatkowe zmienne objaśniające $\mathbf{x}_{it}$[4]. Model taki ma postać:

$$y_{it} = \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_{it} + \gamma y_{i,t-1} + \alpha_i + \varepsilon_{it} \tag{3.15}$$

[4] G. R i d d e r i T. W a n s b e e k [1990] zauważyli możliwość dokonania pewnych uproszczeń w stosunku do propozycji AH82, wynikających z założeń o stacjonarności zmiennych (objaśnianej i egzogenicznych).

gdzie $|\gamma| < 1$; $\mathrm{E}(\alpha_i) = \mathrm{E}(\varepsilon_{it}) = 0$, $\mathrm{E}(\alpha_i \varepsilon_{it}) = 0$, $\mathrm{E}(\varepsilon_{it}\varepsilon_{js}) = \sigma_\varepsilon^2$ dla $i = j$, $t = s$, $\mathrm{E}(\varepsilon_{it}\varepsilon_{js}) = 0$ w pozostałych przypadkach.

Podobnie jak w artykule z 1981 r., również w pracy AH82 rozważono kilka różnych przypadków w zależności od założeń o wartościach początkowych. Dodatkowo wyszczególniono przypadki, kiedy wszystkie zmienne egzogeniczne są stałe w czasie lub wszystkie zmieniają się w czasie, a także przypadek mieszany.

W przypadku, gdy zmienne objaśniające $\mathbf{x}_{it}$ są stałe w czasie, autorzy AH82 rozważają cztery różne zestawy założeń o wartościach początkowych, analogiczne jak w pracy AH81. Do estymacji parametrów modelu proponują oni zastosowanie estymatorów MNW i/lub wewnątrzgrupowych (WG)[5]. W każdym przypadku badane są własności asymptotyczne proponowanych estymatorów, w zależności od T i N.

Jak stwierdzają sami autorzy (AH82, s. 58), iteracyjna procedura wyznaczania estymatora MNW jest bardzo skomplikowana. Jako alternatywę proponują zastosowanie metody zmiennych instrumentalnych. Ocenę parametru γ uzyskać można na podstawie wzoru (3.14). Wstawiając następnie $\hat{\gamma}_{\mathrm{AH}}$ do równania:

$$\bar{y}_i - \gamma \bar{y}_{i,t-1} = \mathbf{x}_i^{\mathrm{T}} \boldsymbol{\beta} + \alpha_i + \bar{\varepsilon}_i$$

gdzie $\bar{y}_i$, $\bar{y}_{i,-1}$, $\bar{\varepsilon}_i$ oznaczają średnie po czasie zmiennych y_{it}, $y_{i,t-1}$ i ε_{it}, wyznaczyć można ocenę wektora parametrów $\boldsymbol{\beta}$ za pomocą KMNK.

Własności asymptotyczne estymatorów MNW, WG i IV są porównane w pracy AH82 (tab. 1, s. 60). Ze względu na dużą liczbę zestawionych w niej przypadków nie jest ona prezentowana w niniejszej pracy.

W przypadku, gdy zmienne objaśniające $\mathbf{x}_{it}$ zmieniają się w czasie, AH82 wyróżniają model autokorelacyjny (ang. *serial correlation model*) i model zależności stanowej (ang. *state dependence model*). Powołują się przy tym na pracę H e c k m a n a [1978], który pierwszy wprowadził takie rozróżnienie.

Model autokorelacyjny ma postać:

$$\begin{cases} y_{it} = w_{it} + \boldsymbol{\beta}^{\mathrm{T}} \mathbf{x}_{it} + \alpha_i, & i = 0, 1, \ldots, N, \quad t = 1, \ldots, T \\ w_{it} = \gamma w_{i,t-1} + \varepsilon_{it}, & i = 1, \ldots, N, \quad t = 1, \ldots, T \end{cases} \tag{3.16}$$

[5] Za pomocą estymatora WG można oszacować tylko parametr γ. Estymator parametru β nie istnieje, gdyż przekształcenie wewnątrzgrupowe usuwa z modelu zmienne stałe w czasie, a zatem w omawianym przypadku również zmienną egzogeniczną.

lub równoważnie:

$$y_{it} = \gamma y_{i,t-1} + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_{it} - \gamma\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_{i,t-1} + \alpha_i + \varepsilon_{it} \tag{3.17}$$

Model zależności stanowej ma natomiast postać:

$$\begin{cases} y_{it} = w_{it} + \alpha_i, & i = 0, 1, ..., N, \quad t = 1, ..., T \\ w_{it} = \gamma w_{i,t-1} + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_{it} + \varepsilon_{it}, & i = 1, ..., N, \quad t = 1, ..., T \end{cases} \tag{3.18}$$

lub równoważnie:

$$y_{it} = \gamma y_{i,t-1} + \boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_{it} + \alpha_i + \varepsilon_{it} \tag{3.19}$$

W obu powyższych przypadkach nieobserwowalne zmienne w_{it} wykazują autokorelację. W modelu autokorelacyjnym (3.16) na kształtowanie się zmiennej y_{it} wpływa jedynie bieżąca wartość zmiennych egzogenicznych. Obecność w modelu opóźnionej zmiennej endogenicznej wynika jedynie z korelacji w czasie składnika losowego. Natomiast w modelu zależności stanowej (3.18) y_{it} kształtuje się również pod wpływem opóźnionych wartości y_{it}. Jak piszą autorzy AH82 (s. 61), nieobserwowalne wartości w_{it} „przenoszą" działanie czynników $\mathbf{x}_{it}$ i można je nazwać stanem. Zasadnicze rozróżnienie między dwoma powyższymi modelami związane jest z występowaniem dynamicznej reakcji na zmiany wartości zmiennych egzogenicznych. W modelu autokorelacyjnym, jeżeli wartość zmiennej x wzrasta w okresie t, a później wraca do poprzedniego poziomu, to nie ma to wpływu na rozkład zmiennej y w okresie $t + 1$. W modelu zależności stanowej natomiast, zmiana wartości x w okresie t wpływa na rozkład zmiennej y w okresie $t + 1$.

Autorzy wprowadzają też dalsze rozróżnienie modeli: autokorelacyjnego i zależności stanowej, w zależności od wartości początkowych. Wyróżniane są dwa przypadki dla modelu autokorelacyjnego (wartości w_{i0} są nielosowe lub losowe i stacjonarne) i aż osiem dla modelu zależności stanowej.

Dla wszystkich tych przypadków proponowane są estymatory MNW. Ich własności asymptotyczne są zestawione w tab. 2 (AH82, s. 67) i tab. 3 (AH82, s. 78). Dodatkowo proponowane są również, jako mniej skomplikowane, metody alternatywne. Dla modelu autokorelacyjnego jest to wielostopniowa UMNK z estymacją, a dla modelu zależności stanowej – metoda IV. Konieczność stosowania w praktyce metod alternatywnych w stosunku do MNW jest zauważona przez T. W. Andersona i C. Hsiao, którzy stwierdzają, że niekiedy badacz dysponuje zbyt małą liczbą informacji, aby poprawnie wskazać warunki początkowe; jeśli zaś są one wskazane błędnie, to estymator MNW jest niezgodny.

Zastosowanie metody IV do estymacji modelu zależności stanowej wymaga, podobnie jak w przypadku modelu autoregresyjnego, obliczenia pierwszych różnic modelu (3.19). Model pierwszych różnic ma postać:

$$\left(y_{it} - y_{i,t-1}\right) = \gamma\left(y_{i,t-1} - y_{i,t-2}\right) + \boldsymbol{\beta}^{\mathrm{T}}\left(\mathbf{x}_{it} - \mathbf{x}_{i,t-1}\right) + \left(\varepsilon_{it} - \varepsilon_{i,t-1}\right) \tag{3.20}$$

Zmiennymi instrumentalnymi dla $\left(y_{i,t-1} - y_{i,t-2}\right)$, tak samo jak dla modelu (3.13), mogą być $y_{i,t-2}$ lub $\Delta y_{i,t-2}$. Uzyskane w ten sposób estymatory $\hat{\gamma}_{\mathrm{AH}}$ i $\hat{\boldsymbol{\beta}}_{\mathrm{AH}}$ są zgodne dla $N \to \infty$ lub $T \to \infty$. Nie zależą one ponadto od wartości początkowych. Oceny składowych wariancji wyznacza się według wzorów:

$$\hat{\sigma}^2_{\varepsilon,\mathrm{AH}} = \frac{1}{2N(T-1)} \sum_{i=1}^{N} \sum_{t=2}^{T} \left[\left(y_{it} - y_{i,t-1}\right) - \hat{\boldsymbol{\beta}}^{\mathrm{T}}_{\mathrm{AH}}\left(\mathbf{x}_{it} - \mathbf{x}_{i,t-1}\right) - \hat{\gamma}_{\mathrm{AH}}\left(y_{i,t-1} - y_{i,t-2}\right)\right]^2 \tag{3.21}$$

$$\hat{\sigma}^2_{\alpha,\mathrm{AH}} = \frac{1}{N} \sum_{i=1}^{N} \left(\overline{\mathbf{y}}_i - \overline{\mathbf{y}}_{i,t-1}\hat{\gamma}_{\mathrm{AH}} - \hat{\mathbf{x}}_i \hat{\boldsymbol{\beta}}_{\mathrm{AH}}\right)^2 - \frac{1}{T}\hat{\sigma}^2_{\varepsilon,\mathrm{AH}} \tag{3.22}$$

Estymator (3.20) jest zgodny, jeśli $N \to \infty$ lub $T \to \infty$, a estymator (3.21) tylko w przypadku, gdy $N \to \infty$.

W ostatnim rozważanym przez autorów AH82 przypadku, gdy w modelu występują jednocześnie zmienne egzogeniczne stałe i zmieniające się w czasie, sugerują oni zastosowanie metod, będących połączeniem omówionych powyżej.

Metoda Andersona–Hsiao daje zgodne, ale nie zawsze efektywne estymatory parametrów. Utrata efektywności wynika, co zauważyli S. C. Ahn i P. Schmidt [1995], z niewykorzystania wszystkich dostępnych warunków momentów oraz nieuwzględnienia struktury zróżnicowanego składnika losowego $\Delta\varepsilon_{it}$. Metody uzyskania estymatorów efektywnych zostały wskazane w późniejszych pracach, np.: Arellano, Bond [1991]; Ahn, Schmidt [1995]; Blundell, Bond [1998]. Są one omówione w kolejnych punktach niniejszego podrozdziału.

3.2.2. Estymator GMM pierwszych różnic Arellano i Bonda (FDGMM)

Estymator GMM pierwszych różnic (FDGMM, ang. *first–differenced GMM*) został zaproponowany przez M. Arellano i S. Bonda [1991]. Praca ta jest rozwinięciem idei przedstawionych przez T. W. Andersona

i C. Hsiao [1981, 1982] oraz D. Holza-Eakina, W. Neweya i H. S. Rosena [1988]. Jako metodę estymacji dynamicznych modeli panelowych M. Arellano i S. Bond [1991] zaproponowali uogólnioną metodę momentów (GMM), opracowaną dla modeli szeregów czasowych przez L. P. Hansena [1982]. Wskazali oni też zmienne instrumentalne dla opóźnionej zmiennej objaśnianej oraz z góry ustalonych i endogenicznych zmiennych objaśniających. Te instrumenty (opóźnione poziomy poszczególnych zmiennych) oraz pierwsze różnice zmiennych egzogenicznych tworzą potencjalnie dużą macierz zmiennych instrumentalnych. Wykorzystując tę macierz, Arellano i Bond zaproponowali jedno- i dwustopniowy estymator GMM oraz właściwe dla nich odporne estymatory składowych wariancji (ang. *variance components estimator*, VCE). Wykazali oni też, że VCE dla metody dwustopniowej jest silnie obciążony w dół (VCE dla metody jednostopniowej jest nieobciążony).

M. Arellano i S. Bond [1991] rozważają najpierw przypadek modelu autoregresyjnego postaci:

$$y_{it} = \gamma y_{i,t-1} + u_{it} = \gamma y_{i,t-1} + (\alpha_i + \varepsilon_{it}) \tag{3.23}$$

gdzie $u_{it} = \alpha_i + \varepsilon_{it}$, oraz spełnione są założenia: $|\gamma| < 1$, $\mathrm{E}(\alpha_i) = \mathrm{E}(\varepsilon_{it}) = 0$, $\mathrm{E}(\alpha_i \varepsilon_{it}) = 0$, $\mathrm{E}(y_{i1}\varepsilon_{it}) = 0$ dla i = 1, ..., N, t = 2, ..., T, składnik losowy ε_{it} nie wykazuje autokorelacji: $\mathrm{E}(\varepsilon_{it}\varepsilon_{is}) = 0$ dla i = 1, ..., N, $t \neq s$.

W celu wyeliminowania z równania (3.23) stałych w czasie efektów grupowych, oblicza się pierwsze różnice[6], co daje:

$$y_{it} - y_{i,t-1} = \gamma(y_{i,t-1} - y_{i,t-2}) + (\varepsilon_{it} - \varepsilon_{i,t-1}) \quad \text{dla } t = 3, ..., T \tag{3.24}$$

Zgodny i efektywny estymator modelu (3.24) uzyskać można stosując GMM, gdyż jeśli $N \to \infty$, a T jest skończone, to uwzględnione założenia implikują $m = 0{,}5(T-1)(T-2)$ warunków momentów postaci:

$$\mathrm{E}(\Delta\varepsilon_{it} y_{i,t-s}) = 0 \quad \text{dla } t = 3, ..., T, \quad s \geq 2 \tag{3.25}$$

Zauważmy bowiem, że warunek ten dla t = 3, pierwszego okresu, w którym jest spełniony, ma postać:

$$y_{i3} - y_{i2} = \gamma(y_{i2} - y_{i1}) + (\varepsilon_{i3} - \varepsilon_{i2}).$$

[6] Idea obliczenia pierwszych różnic modelu została zaproponowana przez T. W. Andersona i C. Hsiao [1981].

Zmienną instrumentalną może być y_{i1}, ponieważ jest ona silnie skorelowana z $(y_{i2} - y_{i1})$, a nieskorelowana z $(\varepsilon_{i3} - \varepsilon_{i2})$ (ε_{it} są z założenia nieskorelowane w czasie). Dla $t = 4$ mamy:

$$y_{i4} - y_{i3} = \gamma(y_{i3} - y_{i2}) + (\varepsilon_{i4} - \varepsilon_{i3}),$$

zatem zmiennymi instrumentalnymi dla $(y_{i3} - y_{i2})$ mogą być y_{i1}, i y_{i2}, ponieważ obie są nieskorelowane z $(\varepsilon_{i4} - \varepsilon_{i3})$.

Kontynuując to postępowanie, w każdym kolejnym okresie próby wskazać można jeden dodatkowy instrument. Ostatecznie dla okresu T właściwymi zmiennymi instrumentalnymi są: $y_{i1}, y_{i2}, ..., y_{i,T-2}$.

Warunki (3.25) zapisać można w postaci wektorowej:

$$\mathrm{E}\left(\mathbf{Z}_i^{\mathrm{T}} \Delta\boldsymbol{\varepsilon}_i\right) = \mathbf{0} \tag{3.26}$$

gdzie:

$$\mathbf{Z}_i = \begin{bmatrix} [y_{i1}] & \mathbf{0} & \cdots & \mathbf{0} \\ 0 & [y_{i1}, y_{i2}] & & \mathbf{0} \\ \vdots & & \ddots & \mathbf{0} \\ 0 & \cdots & \mathbf{0} & [y_{i1}, ..., y_{i,T-2}] \end{bmatrix}_{(T-2)\times m}, \qquad \Delta\boldsymbol{\varepsilon}_i = \begin{bmatrix} \varepsilon_{i2} - \varepsilon_{i1} \\ \cdots \\ \varepsilon_{iT} - \varepsilon_{i,T-1} \end{bmatrix}_{(\mathrm{T}-2)\times 1} \tag{3.27}$$

Każdy wiersz macierzy $\mathbf{Z}_i$ zawiera zmienne instrumentalne właściwe dla danego okresu.

Dalsze postępowanie polega na zastosowaniu uogólnionej metody momentów (GMM). Założenia i zarys idei GMM dla modeli szeregów czasowych są omówione w niniejszej publikacji w dodatku II. M. A r e l l a n o i S. B o n d [1991] zastosowali tę metodę do estymacji modelu (3.24). Model ten zapisać można macierzowo w postaci:

$$\Delta\mathbf{y} = \Delta\mathbf{y}_{-1}\gamma + \Delta\boldsymbol{\varepsilon} \tag{3.28}$$

gdzie: $\Delta\mathbf{y} = [\Delta\mathbf{y}_1^{\mathrm{T}}, ..., \Delta\mathbf{y}_N^{\mathrm{T}}]^{\mathrm{T}}$, $\Delta\mathbf{y}_i$ (i = 1, ..., N) określone jest analogicznie jak $\Delta\boldsymbol{\varepsilon}_i$ we wzorze (3.27), $\Delta\mathbf{y}_{-1} = [\Delta\mathbf{y}_{1,-1}^{\mathrm{T}}, ..., \Delta\mathbf{y}_{N,-1}^{\mathrm{T}}]^{\mathrm{T}}$, $\Delta\mathbf{y}_{i,-1} = [y_{i3} - y_{i2}, ..., y_{iT} - y_{i,T-2}]^{\mathrm{T}}$, $\Delta\boldsymbol{\varepsilon} = [\Delta\boldsymbol{\varepsilon}_1, ..., \Delta\boldsymbol{\varepsilon}_N]^{\mathrm{T}}$. Mnożąc (3.28) przez macierz zmiennych instrumentalnych $\mathbf{Z}$, gdzie $\mathbf{Z} = [\mathbf{Z}_1^{\mathrm{T}}, ..., \mathbf{Z}_N^{\mathrm{T}}]^{\mathrm{T}}$ otrzymuje się:

$$\mathbf{Z}^{\mathrm{T}}\Delta\mathbf{y} = \mathbf{Z}^{\mathrm{T}}\Delta\mathbf{y}_{-1}\gamma + \mathbf{Z}^{\mathrm{T}}\Delta\boldsymbol{\varepsilon} \tag{3.29}$$

Ponieważ liczba warunków momentów zawyczaj jest większa niż liczba nieznanych parametrów, estymator parametru γ wyznacza się minimalizując formę kwadratową:

$$Q_N(\gamma) = \left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i^{\mathrm{T}}\left(\Delta\mathbf{y}_i - \gamma\Delta\mathbf{y}_{i,-1}\right)\right]^{\mathrm{T}}\mathbf{W}_N\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i^{\mathrm{T}}\left(\Delta\mathbf{y}_i - \gamma\Delta\mathbf{y}_{i,-1}\right)\right] \tag{3.30}$$

gdzie $\mathbf{W}_N$ jest symetryczną, dodatnio określoną macierzą wag. Estymator ten ma postać:

$$\hat{\gamma}_{\mathrm{GMM}} = \left[\left(\sum_{i=1}^{N}\Delta\mathbf{y}_{i,-1}^{\mathrm{T}}\mathbf{Z}_i\right)\mathbf{W}_N\left(\sum_{i=1}^{N}\mathbf{Z}_i^{\mathrm{T}}\Delta\mathbf{y}_{i,-1}\right)\right]^{-1}\left[\left(\sum_{i=1}^{N}\Delta\mathbf{y}_{i,-1}^{\mathrm{T}}\mathbf{Z}_i\right)\mathbf{W}_N\left(\sum_{i=1}^{N}\mathbf{Z}_i^{\mathrm{T}}\Delta\mathbf{y}_{i,-1}\right)\right] \tag{3.31}$$

i nazywany jest estymatorem GMM pierwszych różnic (FDGMM) lub estymatorem Arellano i Bonda. Zastosowanie (3.31) w praktyce wymaga wcześniejszej estymacji macierzy wag. Jako początkową macierz wag przyjąć można $\mathbf{W}_N = \mathbf{I}$ lub, zgodnie z propozycją Arellano i Bonda,

$$\mathbf{W}_N = \frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i^{\mathrm{T}}\mathbf{G}\mathbf{Z}_i \tag{3.32}$$

gdzie:

$$\mathbf{G} = \begin{bmatrix} 2 & -1 & 0 & \cdots \\ -1 & 2 & \ddots & 0 \\ 0 & \ddots & \ddots & -1 \\ \vdots & 0 & -1 & 2 \end{bmatrix}.$$

Początkowy estymator $\hat{\gamma}_{\mathrm{GMM1}}$ parametru γ uzyskany na podstawie wzoru (3.31) z macierzą wag (3.32) nazywa się jednostopniowym estymatorem GMM pierwszych różnic (FDGMM1). Estymator dwustopniowy $\hat{\gamma}_{\mathrm{GMM2}}$ uzyskuje się zastępując we wzorze (3.31) macierz $\mathbf{W}_N$ jej zgodnym i asymptotycznie efektywnym estymatorem postaci ogólnej:

$$\hat{\mathbf{W}}_N^{\mathrm{opt}} = \left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i^{\mathrm{T}}\Delta\hat{\boldsymbol{\varepsilon}}_i\Delta\hat{\boldsymbol{\varepsilon}}_i^{\mathrm{T}}\mathbf{Z}_i\right)^{-1} \tag{3.33}$$

Estymator (3.33) wyznacza się stosując procedurą iteracyjną opisaną w dodatku II, prowadzącą do wzoru (D.II.18). $\Delta\hat{\boldsymbol{\varepsilon}}_i$ oznacza wektor reszt początkowego estymatora $\hat{\gamma}_{\mathrm{GMM1}}$.

Estymatory $\hat{\gamma}_{\mathrm{GMM}}$ i $\hat{\mathbf{W}}_N^{\mathrm{opt}}$ wyprowadza się nie czyniąc żadnych założeń o rozkładzie α_i i ε_{it}. Dla zapewnienia właściwości instrumentów niezbędne jest jednak założenie o braku autokorelacji ε_{it}[7]. Jeśli spełnione są założenia modelu (3.23) i dodatkowo ε_{it} jest homoskedastyczny, to $\mathrm{E}\left(\Delta\boldsymbol{\varepsilon}_i\Delta\boldsymbol{\varepsilon}_i^{\mathrm{T}}\right)=\sigma_\varepsilon^2\mathbf{G}$ i optymalną macierz wag można wyznaczyć jako:

$$\tilde{\mathbf{W}}_N^{\mathrm{opt}}=\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i^{\mathrm{T}}\mathbf{G}\mathbf{Z}_i\right)^{-1} \tag{3.34}$$

Macierz (3.34) nie zawiera nieznanych parametrów, zatem przy założeniu homoskedastyczności ε_{it} jednostopniowy estymator FDGMM jest równoważny estymatorowi dwustopniowemu. Jeśli zaś ε_{it} jest heteroskedastyczny, to estymator dwustopniowy jest bardziej efektywny. B o n d, H o e f f l e r i T e m p l e [2001] (s. 18) stwierdzają jednak, powołując się na eksperymenty Monte Carlo, że wzrost efektywności jest z reguły niewielki, oraz że estymator dwustopniowy zbiega do swojego rozkładu asymptotycznego dość powoli. W skończonych próbach błędy szacunku tego estymatora mogą ponadto być silnie obciążone w dół[8].

Macierz wariancji-kowariancji estymatora (3.31) jest w ogólnym przypadku dana wzorem:

$$\operatorname*{p\,lim}_{N\to\infty}\left[\left(\frac{1}{N}\sum_{i=1}^{N}\Delta\mathbf{y}_{i,-1}^{\mathrm{T}}\mathbf{Z}_i\right)\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i^{\mathrm{T}}\Delta\boldsymbol{\varepsilon}_i\Delta\boldsymbol{\varepsilon}_i^{\mathrm{T}}\mathbf{Z}_i\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i^{\mathrm{T}}\Delta\mathbf{y}_{i,-1}\right)\right]^{-1} \tag{3.35}$$

a przy dodatkowych założeniach dotyczących normalności rozkładu i sferyczności ε_{it} jej środkowy czynnik redukuje się do $\sigma_\varepsilon^2\tilde{\mathbf{W}}_N^{\mathrm{opt}}$.

Omówione powyżej postępowanie można zmodyfikować na przypadek modelu, w którym występują niezależne zmienne objaśniające: $y_{it}=\gamma y_{i,t-1}+\boldsymbol{\beta}^{\mathrm{T}}\mathbf{x}_{it}+\alpha_i+\varepsilon_{it}$, a więc modelu postaci (3.6).

---

[7] W pracy A r e l l a n o i B o n d a [1991] zaproponowany jest odpowiedni test statystyczny, badający występowanie autokorelacji. Jest on omówiony w podrozdz. 3.3.

[8] Stwierdzenia zawarte w tym akapicie dotyczą również estymatora systemowego, który jest omawiany w pkt 3.2.6.

Postać dodatkowych warunków momentów i sposób wyznaczenia macierzy wag $\mathbf{W}_N$ zależy od charakteru zmiennych $\mathbf{x}_{it}$. Znaczenie mają założenia dotyczące korelacji tych zmiennych zarówno ze składnikiem losowym ε_{it}, jak i z efektami grupowymi α_i.

Jeśli zmienne $\mathbf{x}_{it}$ są skorelowane z efektami grupowymi α_i, to postać optymalnej macierzy instrumentów zależy od tego, czy są one ściśle egzogeniczne, czy z góry ustalone.

Jeśli zmienne $\mathbf{x}_{it}$ są ściśle egzogeniczne, tzn. $\mathrm{E}(\mathbf{x}_{it}\varepsilon_{is}) = \mathbf{0}$ dla $t, s = 1, ..., T$, to dla $i = 1, ..., N$ zachodzi warunek:

$$\mathrm{E}(\mathbf{x}_{it}\Delta\varepsilon_{is}) = \mathbf{0} \quad \text{dla} \quad t, s = 1, ..., T \tag{3.36}$$

Wówczas właściwymi instrumentami we wszystkich równaniach modelu pierwszych różnic uzyskanego z (3.11) są $\mathbf{x}_{i1}, ..., \mathbf{x}_{iT}$. Macierz instrumentów ma wtedy postać:

$$\mathbf{Z}_i = \mathrm{diag}(y_{i1}, ..., y_{it}, \mathbf{x}_{i1}^{\mathrm{T}}, ..., \mathbf{x}_{i,t+1}^{\mathrm{T}}) \tag{3.37}$$

Oznacza to, że w macierzy zmiennych instrumentalnych $\mathbf{Z}_i$ danej wzorem (3.27) do każdego elementu głównej przekątnej dodać trzeba zmienne $\mathbf{x}_{i1}, ..., \mathbf{x}_{iT}$. Taki sposób postępowania powoduje zwiększenie wymiarów macierzy $\mathbf{Z}_i$ i utrudnia obliczenia. Dlatego zazwyczaj wykorzystuje się pierwsze różnice wektora $\mathbf{x}_{it}$ jako instrumenty dla samych siebie. Stopień utraty efektywności jest w takim przypadku zazwyczaj niewielki. Warunki momentów przyjmują wówczas postać:

$$\mathrm{E}(\Delta\mathbf{x}_{it}\Delta\varepsilon_{is}) = \mathbf{0} \quad \text{dla} \quad t = 1, ..., T, \; i = 1, ..., N \tag{3.38}$$

a macierz instrumentów:

$$\mathbf{Z}_i = \begin{bmatrix} \left[y_{i1}, \Delta\mathbf{x}_{i2}^{\mathrm{T}}\right] & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \left[y_{i1}, y_{i2}, \Delta\mathbf{x}_{i3}^{\mathrm{T}}\right] & & \mathbf{0} \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \left[y_{i1}, ..., y_{i,T-2}, \Delta\mathbf{x}_{iT}^{\mathrm{T}}\right] \end{bmatrix} \tag{3.39}$$

Jeśli zmienne $\mathbf{x}_{it}$ są z góry ustalone, tzn. $\mathrm{E}(\mathbf{x}_{it}\varepsilon_{is}) \neq \mathbf{0}$ dla $t > s$, oraz $\mathrm{E}(\mathbf{x}_{it}\varepsilon_{is}) = \mathbf{0}$ dla $s \geq t$, oznacza to, że bieżące i opóźnione wartości zmiennych $\mathbf{x}_{it}$ są nieskorelowane z bieżącymi wartościami składnika losowego. Właściwymi

instrumentami dla modelu pierwszych różnic uzyskanego z (3.11) w okresie s są zatem $\mathbf{x}_{i,s-1}$, ..., $\mathbf{x}_{i1}$. Warunki momentów mają wówczas postać:

$$\mathrm{E}\left(\mathbf{x}_{i,t-s}\Delta\varepsilon_{it}\right)=\mathbf{0} \quad \text{dla} \quad t = 3, ..., T, \; s \geq 2 \tag{3.40}$$

Model pierwszych różnic powstały z równania (3.11) dla $t = 3$ ma postać:

$$y_{i3}-y_{i2}=\gamma\left(y_{i2}-y_{i1}\right)+\left(\mathbf{x}_{i3}^{\mathrm{T}}-\mathbf{x}_{i2}^{\mathrm{T}}\right)\boldsymbol{\beta}+\left(\varepsilon_{i3}-\varepsilon_{i2}\right).$$

Właściwymi instrumentami dla tego okresu są $\mathbf{x}_{i1}^{\mathrm{T}}$ i $\mathbf{x}_{i2}^{\mathrm{T}}$, ponieważ są one nieskorelowane ze składnikiem losowym $\left(\varepsilon_{i3}-\varepsilon_{i2}\right)$. Analogicznie model pierwszych różnic dla $t = 4$ ma postać:

$$y_{i4}-y_{i3}=\gamma\left(y_{i3}-y_{i2}\right)+\left(\mathbf{x}_{i4}^{\mathrm{T}}-\mathbf{x}_{i3}^{\mathrm{T}}\right)\boldsymbol{\beta}+\left(\varepsilon_{i4}-\varepsilon_{i3}\right),$$

a liczba właściwych instrumentów wzrasta do trzech, ponieważ $\mathbf{x}_{i1}^{\mathrm{T}}$, $\mathbf{x}_{i2}^{\mathrm{T}}$ i $\mathbf{x}_{i3}^{\mathrm{T}}$ są nieskorelowane z $\left(\varepsilon_{i4}-\varepsilon_{i3}\right)$. Kontynuując powyższe postępowanie uzyskuje się macierz instrumentów postaci:

$$\mathbf{Z}_i=\begin{bmatrix} \left[y_{i1},\mathbf{x}_{i1}^{\mathrm{T}},\mathbf{x}_{i2}^{\mathrm{T}}\right] & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \left[y_{i1},y_{i2},\mathbf{x}_{i1}^{\mathrm{T}},\mathbf{x}_{i2}^{\mathrm{T}},\mathbf{x}_{i3}^{\mathrm{T}}\right] & & \mathbf{0} \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \left[y_{i1},...,y_{i,T-2},\mathbf{x}_{i1}^{\mathrm{T}},...,\mathbf{x}_{i,T-1}^{\mathrm{T}}\right] \end{bmatrix} \tag{3.41}$$

W przypadku, gdy wśród elementów wektora $\mathbf{x}_{it}$ znajdują się zarówno zmienne ściśle egzogeniczne, jak i z góry ustalone, macierz instrumentów jest odpowiednią kombinacją macierzy (3.39) i (3.41).

Jeśli nie wszystkie zmienne $\mathbf{x}_{it}$ są skorelowane z efektami grupowymi α_i M. Arellano i B. Bond [1991] proponują wykorzystanie idei J. A. Hausmana i W. E. Taylora [1981][9] i dekompozycję wektora $\mathbf{x}_{it}$ na dwa podzbiory: $\mathbf{x}_{it}=\left[\mathbf{x}_{1it} \quad \mathbf{x}_{2it}\right]$ takie, że składowe wektora $\mathbf{x}_{1it}$ są nieskorelowane z α_i, a składowe $\mathbf{x}_{2it}$ – skorelowane z α_i. Odpowiednio zmodyfikowana macierz instrumentów $\mathbf{Z}_i^*$ przyjmuje postać:

[9] Metoda Hausmana i Taylora omówiona jest w podrozdz. 2.4.

$$\mathbf{Z}_i^* = \begin{bmatrix} \mathbf{Z}_i & \mathbf{0} & \dots & & \mathbf{0} \\ \mathbf{0} & \left[\mathbf{x}_{1i1}^{\mathrm{T}}, \mathbf{x}_{1i2}^{\mathrm{T}}\right] & & & \mathbf{0} \\ \vdots & & \mathbf{x}_{1i3}^{\mathrm{T}} & & \mathbf{0} \\ & & & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & & \mathbf{0} & \mathbf{x}_{1iT}^{\mathrm{T}} \end{bmatrix} \tag{3.42}$$

gdzie $\mathbf{Z}_i$ jest określona wzorem (3.41).

Estymatory GMM jedno- i dwustopniowy wektora parametrów strukturalnych $[\gamma\ \boldsymbol{\beta}^{\mathrm{T}}]^{\mathrm{T}}$ modelu (3.11) uzyskuje się analogicznie, jak w przypadku modelu autoregresyjnego, opisanego powyżej. Macierz zmiennych instrumentalnych $\mathbf{Z}_i$ zastępuje się macierzą określoną w zależności od charakteru zmiennych $\mathbf{x}_{it}$ wzorami (3.39), (3.41) lub (3.42). Wzór (3.31), określający estymator GMM pierwszych różnic (FDGMM), po uwzględnieniu występowania w modelu wektora parametrów $\boldsymbol{\beta}$, przyjmuje postać:

$$\begin{aligned} \begin{bmatrix} \hat{\gamma} \\ \hat{\boldsymbol{\beta}} \end{bmatrix}_{\mathrm{GMM}} &= \left[\left(\sum_{i=1}^{N} (\Delta\mathbf{y}_{i,-1}^{\mathrm{T}})^* \mathbf{Z}_i\right) \mathbf{W}_N \left(\sum_{i=1}^{N} \mathbf{Z}_i^{\mathrm{T}} (\Delta\mathbf{y}_{i,-1})^*\right)\right]^{-1} \times \\ &\times \left[\left(\sum_{i=1}^{N} (\Delta\mathbf{y}_{i,-1}^{\mathrm{T}})^* \mathbf{Z}_i\right) \mathbf{W}_N \left(\sum_{i=1}^{N} \mathbf{Z}_i^{\mathrm{T}} (\Delta\mathbf{y}_{i,-1})^*\right)\right] \end{aligned} \tag{3.43}$$

Wzór (3.43) jest analogiczny do (3.31); różnica polega na odmiennym określeniu macierzy $\mathbf{Z}_i$ oraz na uwzględnieniu tego, że w modelu występują zmienne $\mathbf{x}_{it}$: wektor $\Delta\mathbf{y}_{i,-1}$ ze wzoru (3.31) jest zastąpiony przez wektor $(\Delta\mathbf{y}_{i,-1})^* = \Delta\mathbf{y}_{i,-1}\left[\Delta\mathbf{y}_{i,-1} \quad \Delta\mathbf{x}_{it}\right]^{\mathrm{T}}$.

Jak już wspomniano, jeżeli opóźnione poziomy zmiennych są słabymi instrumentami dla zmiennych zróżnicowanych, to estymatory FDGMM są obciążone i nieefektywne. W modelu autoregresyjnym (3.23) sytuacja taka ma miejsce, jeśli wartość parametru γ jest bliska 1, lub jeśli wariancja efektów grupowych α_i rośnie wraz z wariancją ε_{it} (stosunek wariancji efektów grupowych do wariancji właściwego składnika losowego jest zbyt duży)[10]. Rezultaty symulacji przedstawione w pracy R. B l u n d e l l a i S. B o n d a [1998] pokazują, że estymatory FDGMM mogą być, szczególnie w przypadku małej liczby obserwacji w czasie, obciążone w dół. Na przykład (B l u n d e l l, B o n d [1998],

[10] Bardziej ogólnie, zmienne stosowane jako instrumenty w równaniach na przyrostach mają tendencję do bycia słabymi, jeśli szeregi obserwacji dla poszczególnych obiektów mają własności zbliżone do szeregów posiadających pierwiastki jednostkowe.

tab. 2a), dla $T = 4$, $N = 100$ i rzeczywistej wartości parametru $\gamma = 0{,}9$, estymator FDGMM ma średnią 0,23, a odchylenie standardowe 0,83. Może się zdarzyć, że w pewnych, szczególnych sytuacjach, wprowadzenie do modelu zmiennych niezależnych $\mathbf{x}_{it}$ poprawi własności estymatora FDGMM (zob. Bond i in. [2001]), jednak na ogół w takim przypadku należy stosować inne metody estymacji.

Sprawdzenia, czy w danym modelu estymator FDGMM jest obciążony, można dokonać poprzez porównanie uzyskanego oszacowania parametru γ z ocenami uzyskanymi innymi metodami. Jak uzasadniono w podrozdz. 3.1 estymator KMNK parametru γ dla modelu (3.23) jest obciążony w górę, a estymator wewnątrzgrupowy tego parametru jest, przy ustalonym T, obciążony w dół. Zatem wartość zgodnego estymatora parametru γ powinna zawierać się w przedziale pomiędzy wartościami tych dwóch estymatorów. Jeśli ocena FDGMM parametru γ jest bliska ocenie uzyskanej na podstawie estymatora wewnątrzgrupowego, może to oznaczać że instrumenty wykorzystane w procesie estymacji są słabe, co powoduje obciążenie estymatorów FDGMM. Przedstawione powyżej rezultaty, dotyczące obciążenia estymatorów w modelu autoregresyjnym, są prawdziwe dla modelu, w którym występują niezależne zmienne objaśniające tylko w przypadku, gdy zmienne te są nieskorelowane z α_i oraz ściśle egzogeniczne ze względu na ε_{it}, co zostało szerzej omówione w: Sevestre, Trognon [1996]. Mimo to, jak stwierdzają S. Bond, A. Hoeffler, J. Temple [2001] (s. 7), porównanie estymatora FDGMM parametru γ modelu (3.11) z estymatorem wewnątrzgrupowym i estymatorem KMNK może być wskazane bez względu na charakter zmiennych objaśniających. Zbliżone wartości ocen parametru γ, uzyskane za pomocą FDGMM i estymatora wewnątrzgrupowego, mogą w każdym przypadku sygnalizować występowanie problemu słabych instrumentów. Można wówczas przeprowadzić badanie jakości instrumentów na podstawie postaci zredukowanej modelu lub rozważyć zastosowanie alternatywnych estymatorów, o których wiadomo, że mają lepsze własności w małych próbach w przypadku szeregów o wysokim stopniu trwałości.

3.2.3. Skorygowany estymator wariancji Windmeijera

Estymator wariancji estymatora dwustopniowej FDGMM wyznaczony na podstawie macierzy (3.35) jest, jak stwierdzili M. Arellano i S. Bond [1991], obciążony, przy czym obciążenie to jest szczególnie znaczące w małych próbach. Fakt ten został potwierdzony w późniejszej literaturze na podstawie eksperymentów Monte Carlo. Estymator wariancji jednostopniowego estymatora Arellano–Bonda jest zaś nieobciążony.

Do konstrukcji estymatora jednostopniowego wykorzystuje się macierz wag daną wzorem (3.32), która jest niezależna od estymowanych parametrów.

Konstrukcja estymatora dwustopniowego wymaga natomiast zastosowania macierzy wag danej wzorem (3.33), która zależy od reszt jednostopniowego estymatora $[\hat{\gamma} \quad \hat{\boldsymbol{\beta}}^{\mathrm{T}}]^{\mathrm{T}}_{\mathrm{FDGMM1}}$. Ta macierz wag jest budowana na podstawie początkowego, zgodnego estymatora wektora parametrów strukturalnych modelu.

F. Windmeijer [2005] wykazał, że właśnie obecność początkowych estymatorów parametrów strukturalnych $[\hat{\gamma} \quad \hat{\boldsymbol{\beta}}^{\mathrm{T}}]^{\mathrm{T}}_{\mathrm{FDGMM1}}$ w macierzy wag jest zasadniczą przyczyną obciążenia estymatora dwustopniowego w małych próbach, o ile warunki momentów są liniowe względem parametrów. Windmeijer wykazał też, że wielkość tego obciążenia można estymować, a zatem możliwe jest wyznaczenie wartości, o jaką powinny zostać skorygowane estymatory wariancji parametrów.

Asymptotyczne wariancje jedno- i dwustopniowego estymatora GMM dane są wzorami:

$$\hat{\mathrm{D}}^2(\hat{\boldsymbol{\delta}}_1) = \sigma^2_{\mathrm{GMM1}} = \frac{1}{N}(\mathbf{C}^{\mathrm{T}}\mathbf{W}_N\mathbf{C})^{-1}\mathbf{C}^{\mathrm{T}}\mathbf{W}_N\mathbf{W}_N^{-1}(\boldsymbol{\delta}_1)\mathbf{W}_N\mathbf{C}(\mathbf{C}^{\mathrm{T}}\mathbf{W}_N\mathbf{C})^{-1} \tag{3.44}$$

$$\hat{\mathrm{D}}^2(\hat{\boldsymbol{\delta}}_2) = \sigma^2_{\mathrm{GMM2}} = \frac{1}{N}(\mathbf{C}^{\mathrm{T}}\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)\mathbf{C})^{-1} \tag{3.45}$$

gdzie $\mathbf{C} = \dfrac{\partial \overline{\mathbf{g}}(\boldsymbol{\delta})}{\partial \boldsymbol{\delta}^T}$, $\overline{\mathbf{g}}(\boldsymbol{\delta}) = \dfrac{1}{N}\sum_{i=1}^{N}\mathbf{g}_i(\boldsymbol{\delta})$, $\mathbf{g}_i(\boldsymbol{\delta})$ jest wektorem, występującym w warunkach momentów: $\mathrm{E}\big(\mathbf{g}(\mathbf{x}_i\boldsymbol{\delta})\big) = \mathrm{E}\big(\mathbf{g}_i\boldsymbol{\delta})\big) = \mathbf{0}$, $\hat{\boldsymbol{\delta}}_1$ jest zgodnym estymatorem parametrów modelu jednostopniowej GMM, $\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)$ oznacza macierz wag wykorzystywaną do konstrukcji estymatora jednostopniowego, daną wzorem (3.32).

Wykorzystując rozwiniecie w szereg Taylora wyrażenia $(\hat{\boldsymbol{\delta}}_2 - \boldsymbol{\delta})$, gdzie $\hat{\boldsymbol{\delta}}_2$ jest dwustopniowym estymatorem wektora $\boldsymbol{\delta}$ parametrów strukturalnych modelu (3.11) ($\boldsymbol{\delta} = [\gamma\boldsymbol{\beta}^{\mathrm{T}}]^{\mathrm{T}}$), F. Windmeijer [2005] pokazał, że w małych próbach estymatorem wariancji wektora $\hat{\boldsymbol{\delta}}_2$, skorygowanym ze względu na obciążenie, jest:

$$\begin{aligned}\hat{\mathrm{D}}^2_{\mathrm{C}}(\hat{\boldsymbol{\delta}}_2) = {} & \frac{1}{N}\Big(\mathbf{C}^{\mathrm{T}}\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)\mathbf{C}\Big)^{-1} + \mathbf{G}_{\hat{\boldsymbol{\delta}}_2,\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)}\hat{\mathrm{D}}^2(\hat{\boldsymbol{\delta}}_1)\mathbf{G}^{\mathrm{T}}_{\hat{\boldsymbol{\delta}}_2,\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)} - \\ & + \frac{1}{N}\Big(\mathbf{G}_{\hat{\boldsymbol{\delta}}_2,\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)}\Big(\mathbf{C}^{\mathrm{T}}\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)\mathbf{C}\Big)^{-1} + \Big(\mathbf{C}^{\mathrm{T}}\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)\mathbf{C}\Big)^{-1}\mathbf{G}^{\mathrm{T}}_{\hat{\boldsymbol{\delta}}_2,\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)}\Big)\end{aligned} \tag{3.46}$$

gdzie $\hat{\mathrm{D}}^2(\hat{\boldsymbol{\delta}}_1)$ oznacza estymator wariancji estymatora jednostopniowej GMM, zaś $\mathbf{G}_{\hat{\boldsymbol{\delta}}_2,\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)}$ oznacza macierz stopnia *k*, której *s*-ta kolumna $\mathbf{G}_{\hat{\boldsymbol{\delta}}_2,\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)}[\cdot,s]$ określona jest wzorem:

$$\mathbf{G}_{\hat{\boldsymbol{\delta}}_2,\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)}[\cdot,s] = -\left(\mathbf{C}^{\mathrm{T}}\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)\mathbf{C}\right)^{-1}\mathbf{C}^{\mathrm{T}}\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)\frac{\partial\mathbf{W}_N^{-1}(\boldsymbol{\delta})}{\partial\boldsymbol{\delta}_s}\bigg|_{\hat{\boldsymbol{\delta}}_1}\mathbf{W}_N(\hat{\boldsymbol{\delta}}_1)\overline{\mathbf{g}}(\hat{\boldsymbol{\delta}}_2),$$

gdzie: $\dfrac{\partial\mathbf{W}_N^{-1}(\boldsymbol{\delta})}{\partial\boldsymbol{\delta}_s} = \dfrac{1}{N}\sum_{i=1}^{N}\left(\dfrac{\partial\mathbf{g}_i(\boldsymbol{\delta})}{\partial\boldsymbol{\delta}_s}\mathbf{g}_i(\boldsymbol{\delta})^{\mathrm{T}} + \mathbf{g}_i(\boldsymbol{\delta})\dfrac{\partial\mathbf{g}_i(\boldsymbol{\delta})^{\mathrm{T}}}{\partial\boldsymbol{\delta}_s}\right).$

Skuteczność zastosowania skorygowanego sposobu obliczania wariancji estymatora $\hat{\hat{\boldsymbol{\delta}}}_2$ danego wzorem (3.46), zwanego korektą Windmeijera, jest pokazana przez autora na przykładzie modeli szacowanych w pracach Arellano i Bonda [1991] oraz Blundella i Bonda [2000]. Konstrukcja i wyniki estymacji modelu zatrudnienia Arellano i Bonda [1991] są omówione w rozdz. 4. Blundell i Bond [2000] estymowali funkcję produkcji na podstawie danych pochodzących z 8 lat z 509 firm zajmujących się działalnością badawczo-rozwojową w USA. Windmeijer [2005] dokonał reestymacji tego modelu, stosując jedno- i dwustopniowe FDGMM oraz systemową GMM (SGMM) Blundella i Bonda [1998], która jest omówiona w pkt 3.2.6. Tabela 3 w: Windmeijer [2005] (s. 43) zawiera wyniki estymacji jedno- i dwustopniowej za pomocą FDGMM i SGMM. Dla metod dwustopniowych błędy szacunku parametrów obliczone są dwoma metodami: na podstawie wzoru (3.35) oraz na podstawie korekty Windmeijera (3.46), a dla metody jednostopniowej – tylko błędy nieskorygowane. Mimo że liczba firm jest bardzo duża, skorygowane błędy szacunku (3.46) dwustopniowych estymatorów FDGMM są znacznie większe niż błędy nieskorygowane (3.35) i zbliżone do błędów jednostopniowej FDGMM. Analogiczna zależność zachodzi dla SGMM, z tym że skorygowane błędy estymatorów metody dwustopniowej są mniejsze niż nieskorygowane błędy metody jednostopniowej. Są one również mniejsze niż skorygowane błędy dwustopniowej FDGMM.

3.2.4. Estymator Arellano i Bover

Propozycja Arellano i Bover [1995] jest kontynuacją i rozwinięciem idei Andersona i Hsiao [1982], oraz Arellano i Bonda [1991]. Jako metodę estymacji dynamicznego modelu panelowego autorzy sugerują GMM, wskazując możliwość wykorzystania w procesie estymacji równań na poziomach. Określają oni też warunki momentów dla tych równań oraz właściwe

instrumenty. Bazując na ideach wymienionych wyżej autorów, A r e l l a n o i B o v e r [1995] wykorzystują też metodę H a u s m a n a i T a y l o r a [1981].

Przedstawiony w podrozdz. 2.4 model H a u s m a n a i T a y l o r a [1981] jest modelem statycznym, który zapisać można w postaci:

$$y_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\varphi} + \alpha_i + \varepsilon_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\varphi} + u_{it} \tag{3.47}$$

gdzie $\mathbf{X}_{it}$ i $\mathbf{Z}_i$ są wektorami zmiennych objaśniających, takich że elementy $\mathbf{X}_{it}$ zmieniają się po obiektach i czasie, a elementy $\mathbf{Z}_i$ są niezmienne w czasie, a zmieniają się względem obiektów. A r e l l a n o i B o v e r [1995] zastosowali do estymacji parametrów modelu (3.47) uogólnioną metodę najmniejszych kwadratów (UMNK) i wykazali, że uzyskany estymator jest równoważny optymalnemu estymatorowi GMM. Następnie rozważania zostały uogólnione na przypadek modelu dynamicznego. Metodę A r e l l a n o i B o v e r [1995] wygodniej będzie przedstawić, jeśli model (3.47) zapisany zostanie w formie:

$$\mathbf{y}_i = \mathbf{W}_i\boldsymbol{\eta} + \mathbf{u}_i \tag{3.48}$$

gdzie $\mathbf{y}_i = [y_{i1}, ..., y_{iT}]^{\mathrm{T}}$, $\mathbf{u}_i = [u_{i1}, ..., u_{iT}]^{\mathrm{T}}$, $\boldsymbol{\eta}^{\mathrm{T}} = [\boldsymbol{\beta}^{\mathrm{T}}, \boldsymbol{\varphi}^{\mathrm{T}}]$, $\mathbf{W}_i = [\mathbf{X}_i, \mathbf{i}_T\mathbf{Z}_i]$, $\mathbf{X}_i = [\mathbf{X}_{i1}, ..., \mathbf{X}_{iT}]^{\mathrm{T}}$, zaś $\mathbf{i}_T$ jest T-elementowym wektorem jedynkowym (T – liczba obserwacji w czasie), lub w formie jeszcze bardziej zwartej:

$$\mathbf{y} = \mathbf{W}\boldsymbol{\eta} + \mathbf{u} \tag{3.49}$$

gdzie $\mathbf{y} = [\mathbf{y}_1^{\mathrm{T}}, ..., \mathbf{y}_N^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{u} = [\mathbf{u}_1^{\mathrm{T}}, ..., \mathbf{u}_N^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{W} = [\mathbf{W}_1^{\mathrm{T}}, ..., \mathbf{W}_N^{\mathrm{T}}]^{\mathrm{T}}$. Zakłada się, że efekty grupowe modelu (3.48) są losowe, a więc, że macierz wariancji--kowariancji składnika losowego $\mathbf{u}_i$ ma strukturę taką, jak w modelu RE, oraz spełnione są założenia tego modelu (por. pkt 2.2.1).

Model (3.48) pomnożono następnie stronami przez operator nieosobliwy postaci:

$$\mathbf{H} = \begin{bmatrix} \mathbf{C} \\ \mathbf{i}_T^{\mathrm{T}} / T \end{bmatrix} \tag{3.50}$$

gdzie $\mathbf{C}$ jest dowolną macierzą o wymiarach $((T - 1) \times T)$, rzędu $T - 1$, taką że $\mathbf{C}\mathbf{i}_T = \mathbf{0}$. Podkreślając, że macierz $\mathbf{C}$ można zdefiniować całkowicie dowolnie, A r e l l a n o i B o v e r [1995] wskazali możliwość dwojakiego określenia jej elementów. Mogą one być równe pierwszym $T - 1$ wierszom operatora wewnątrzgrupowego $\mathbf{Q} = \mathbf{I}_N \otimes \left(\mathbf{I}_T - \frac{\mathbf{J}_T}{T}\right)$ albo alternatywnie mogą nimi być

operatory pierwszych różnic. Składnik losowy modelu (3.48), przekształcony za pomocą operatora **H**, ma postać:

$$\mathbf{u}_i^+ = \mathbf{H}\mathbf{u}_i = \begin{bmatrix} \mathbf{C}\mathbf{u}_i \\ \overline{\mathbf{u}}_i \end{bmatrix} \tag{3.51}$$

Pierwsze $T - 1$ elementów składnika losowego (3.51) nie zawiera efektów grupowych α_i, zatem, na podstawie przyjętych założeń, wszystkie zmienne objaśniające mogą być instrumentami w równaniach, odpowiadających pierwszym $T - 1$ obserwacjom. Niech ponadto $\mathbf{m}_i$ oznacza podzbiór zmiennych objaśniających, których poziomy są nieskorelowane z efektami grupowymi, którego liczebność jest co najwyżej równa liczbie składowych wektora $\boldsymbol{\eta}$. Elementy $\mathbf{m}_i$ mogą więc pochodzić z podzbiorów $\mathbf{X}_1$ i $\mathbf{Z}_1$, wyodrębnionych przez Hausmana i Taylora [1981][11]. Konstrukcja podzbioru $\mathbf{m}_i$ opiera się zatem na zmiennych $[\mathbf{Z}_{1,i}, \mathbf{X}_{1,i1}, ..., \mathbf{X}_{1,iT}]^{\mathrm{T}}$, przy czym $\mathbf{Z}_{1,i}$ i $\mathbf{X}_{1,it}$ oznaczają składowe wektorów $\mathbf{Z}_i$ i $\mathbf{X}_{it}$, należące jednocześnie do podzbiorów $\mathbf{Z}_1$ i $\mathbf{X}_1$. Zatem pełna macierz zmiennych instrumentalnych dla przekształconego za pomocą operatora (3.50) modelu ma postać:

$$\mathbf{M}_i = \begin{bmatrix} \mathbf{w}_i^{\mathrm{T}} & \cdots & \cdots & \mathbf{0} \\ \vdots & \ddots & & \vdots \\ \mathbf{0} & & \mathbf{w}_i^{\mathrm{T}} & \mathbf{0} \\ \mathbf{0} & \cdots & \cdots & \mathbf{m}_i^{\mathrm{T}} \end{bmatrix} \tag{3.52}$$

gdzie $\mathbf{w}_i^{\mathrm{T}} = [\mathbf{X}_i, \mathbf{Z}_i]$.Warunki momentów przyjmują postać:

$$\mathrm{E}(\mathbf{M}_i^{\mathrm{T}}\mathbf{H}\mathbf{u}_i) = \mathbf{0} \tag{3.53}$$

Model (3.49) mnoży się następnie stronami przez $\mathbf{M}^{\mathrm{T}}\overline{\mathbf{H}}$, gdzie $\mathbf{M} = [\mathbf{M}_1^{\mathrm{T}}, ..., \mathbf{M}_N^{\mathrm{T}}]^{\mathrm{T}}$, $\overline{\mathbf{H}} = \mathbf{I}_N \otimes \mathbf{H}$, a następnie szacuje jego parametry za pomocą UMNK. Uzyskany w ten sposób estymator Arellano i Bover ma formę:

$$\hat{\boldsymbol{\eta}} = \left[\mathbf{W}^{\mathrm{T}}\overline{\mathbf{H}}^{\mathrm{T}}\mathbf{M}(\mathbf{M}^{\mathrm{T}}\overline{\mathbf{H}}\overline{\boldsymbol{\Omega}}\overline{\mathbf{H}}^{\mathrm{T}}\mathbf{M})^{-1}\mathbf{M}^{\mathrm{T}}\overline{\mathbf{H}}\mathbf{W}\right]^{-1}\mathbf{W}^{\mathrm{T}}\overline{\mathbf{H}}^{\mathrm{T}}\mathbf{M}(\mathbf{M}^{\mathrm{T}}\overline{\mathbf{H}}\overline{\boldsymbol{\Omega}}\overline{\mathbf{H}}^{\mathrm{T}}\mathbf{M})^{-1}\mathbf{M}^{\mathrm{T}}\overline{\mathbf{H}}\mathbf{y} \tag{3.54}$$

---

[11] Dekompozycja $\mathbf{X}_{it}$ i $\mathbf{Z}_i$ na zbiory zmiennych egzogenicznych i endogenicznych jest opisana w podrozdz. 2.4. $\mathbf{X}_1$ i $\mathbf{Z}_1$ oznaczają podzbiory zmiennych $\mathbf{X}$ i $\mathbf{Z}$, które nie są skorelowane z ε_{it} ani z α_i.

Macierz $\overline{\mathbf{\Omega}}$ we wzorze (3.54) oznacza macierz wariancji-kowariancji składnika losowego $\mathbf{u}$ i ma postać $\overline{\mathbf{\Omega}} = \mathbf{I}_N \otimes \mathbf{\Omega}$, gdzie $\mathbf{\Omega} = \mathrm{E}(\mathbf{u}_i \mathbf{u}_i^{\mathrm{T}})$.

Zgodnym estymatorem macierzy $\mathbf{\Omega}^+ = \mathbf{H}\mathbf{\Omega}\,\mathbf{H}^{\mathrm{T}}$, będącej macierzą wariancji-kowariancji składnika losowego $\mathbf{u}_i^+$ danego wzorem (3.51) jest:

$$\hat{\mathbf{\Omega}}^+ = \sum_{i=1}^{N} \hat{\mathbf{u}}_i^+ \hat{\mathbf{u}}_i^{+\mathrm{T}} / N \tag{3.55}$$

gdzie $\hat{\mathbf{u}}_i^+$ są resztami, obliczonymi na podstawie dowolnego, początkowego estymatora zgodnego. Estymator (3.54) z macierzą wariancji-kowariancji (3.55) jest optymalnym estymatorem GMM. Wzrost efektywności estymacji uzyskuje się, zastępując wyrażenie $\sum_i \mathbf{M}_i^{\mathrm{T}} \mathbf{\Omega}^+ \mathbf{M}_i$ we wzorze (3.54) przez $\sum_i \mathbf{M}_i^{\mathrm{T}} \hat{\mathbf{u}}_i^+ \hat{\mathbf{u}}_i^{+\mathrm{T}} \mathbf{M}_i$. Jest to zgodne z sugestią Hansena [1982]. Estymator Hausmana i Taylora [1981], dany wzorem (2.33), jest równoważny estymatorowi (3.54), jeśli we wzorze (3.54) jako estymator macierzy $\mathbf{\Omega}^+$ przyjmie się $\tilde{\mathbf{\Omega}}^+ = \mathbf{H}\tilde{\mathbf{\Omega}}\,\mathbf{H}^{\mathrm{T}}$, z macierzą $\tilde{\mathbf{\Omega}} = \tilde{\sigma}_\varepsilon^2 \mathbf{I}_T + \tilde{\sigma}_\alpha^2 \mathbf{i}_T \mathbf{i}_T^{\mathrm{T}}$, gdzie $\tilde{\sigma}_\varepsilon^2$ i $\tilde{\sigma}_\alpha^2$ są zgodnymi estymatorami wariancji składowych ε i α; jako $\mathbf{m}_i$ należy zaś przyjąć $\mathbf{m}_i = [\mathbf{Z}_{1,i}, \overline{\mathbf{X}}_{1,i}]^{\mathrm{T}}$, gdzie $\overline{\mathbf{X}}_{1,i}$ jest średnią po czasie obliczoną z obserwacji $\mathbf{X}_{1,it}$ dla $t = 1, ..., T$.

Arellano i Bover [1995] wykazali, że ponieważ macierz zmiennych instrumentalnych $\mathbf{M}_i$ jest blokowo-diagonalna, to estymator (3.54) nie zależy od wyboru macierzy $\mathbf{C}$ we wzorze (3.50). Ponadto metoda Arellano i Bover nie wymaga, w przeciwieństwie do metody Hausmana i Taylora [1981], znajomości postaci macierzy $\mathbf{\Omega}^{-1/2}$. Zatem przedstawione powyżej rozważania stanowią uogólnienie propozycji Hausmana i Taylora, w tym sensie, że metoda Arellano i Bover może być stosowana nie tylko do estymacji modeli RE. Następnie przedstawione powyżej rozważania rozszerzone zostały na przypadek dynamicznego modelu panelowego.

Model (3.47) po uwzględnieniu opóźnionej zmiennej objaśniającej przyjmuje postać:

$$y_{it} = \gamma y_{i,t-1} + \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\varphi} + \alpha_i + \varepsilon_{it} = \gamma y_{i,t-1} + \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\varphi} + u_{it} \tag{3.56}$$

Model (3.56) można oczywiście zapisać w postaci wektorowej:

$$\mathbf{y}_i = \mathbf{W}_i^* \boldsymbol{\eta}^* + \mathbf{u}_i \tag{3.57}$$

Przy założeniu, że wartość y_{i0} zmiennej y z okresu $t = 0$ jest znana, macierze $\boldsymbol{\eta}^*$ i $\mathbf{W}^*$ definiuje się wówczas jako: $\boldsymbol{\eta}^{*\mathrm{T}} = [\gamma, \boldsymbol{\beta}^{\mathrm{T}}, \boldsymbol{\varphi}^{\mathrm{T}}]$, $\mathbf{W}_i^* = \left[\mathbf{y}_{i(-1)}, \mathbf{X}_i, \mathbf{i}_T \mathbf{Z}_i\right]$, gdzie $\mathbf{y}_{i(-1)} = [y_{i0}, ..., y_{i,T-1}]^{\mathrm{T}}$. Pozostałe oznaczenia są takie, jak we wzorze (3.48).

Następnie stosuje się postępowanie opisane powyżej dla modelu statycznego. Estymator dany wzorem (3.54) jest zgodnym estymatorem wektora $\boldsymbol{\eta}^*$ dla przypadku dynamicznego, pod warunkiem że liczba właściwych instrumentów jest wystarczająca, aby zapewnić identyfikowalność. Definiując macierz zmiennych instrumentalnych, należy oczywiście uwzględnić fakt, że w rozważanym obecnie przypadku pierwszy okres próby ma numer $t = 0$. Zatem w przypadku modelu dynamicznego:

$$\mathbf{w}_i = [\mathbf{X}_{i0}, ..., \mathbf{X}_{iT}, \mathbf{Z}_i]^{\mathrm{T}} \tag{3.58}$$

Konstrukcja podzbioru $\mathbf{m}_i$ jest analogiczna, jak w przypadku statycznym. Szczególnie dogodny jest jednak w tym przypadku wybór instrumentów, zgodny z propozycją Breuscha, Mizona i Schmidta [1989]. Podzbiór $\mathbf{m}_i$ ma w przypadku modelu dynamicznego postać:

$$\mathbf{m}_i = [\mathbf{Z}_{1,i}, \mathbf{X}_{1,i}, \tilde{\mathbf{X}}_{2,i1}, ..., \tilde{\mathbf{X}}_{2,iT}]^{\mathrm{T}} \tag{3.59}$$

gdzie $\tilde{\mathbf{X}}_{2,it} = \mathbf{X}_{2,it} - \frac{1}{T}\sum_{t=1}^{T} \mathbf{X}_{2,it}$. Zmienne $\tilde{\mathbf{X}}_{2,it}$ mogą w tym przypadku być instrumentami, mimo że dopuszcza się założenie o występowaniu korelacji między $\mathbf{X}_{2,it}$ a efektami grupowymi α_i. Jest to możliwe, ponieważ, jak zauważyli Breusch, Mizon i Schmidt [1989], wykorzystać można dodatkowe warunki momentów o tym, że korelacja między zmiennymi $\mathbf{X}_{2,it}$ a efektami α_i jest stała w czasie. Jeśli $\mathbf{m}_i$ jest określone wzorem (3.59), to wiersze macierzy $\mathbf{CX}_i$ są liniowymi kombinacjami elementów $\mathbf{m}_i$. Oznacza to, że dla równań odpowiadających wszystkim okresom próby można wykorzystać ten sam zbiór instrumentów, a zatem macierz $\mathbf{M}_i$ określona wzorem (3.52) redukuje się do $\mathbf{M}_i = \mathbf{I}_T \otimes \mathbf{m}_i^{\mathrm{T}}$. W konsekwencji transformacja modelu za pomocą operatora $\mathbf{H}$ staje się zbędna, a model może być estymowany bezpośrednio za pomocą trójstopniowej klasycznej metody najmniejszych kwadratów. Estymator GMM wektora parametrów $\boldsymbol{\eta}^*$ dany wzorem (3.54) przyjmuje wówczas postać:

$$\boldsymbol{\eta}^* = \left[\sum_i (\mathbf{W}_i^* \otimes \mathbf{m}_i)^{\mathrm{T}} \left(\hat{\Omega} \otimes \sum_i \mathbf{m}_i \mathbf{m}_i^{\mathrm{T}} \right)^{-1} \sum_i (\mathbf{W}_i^* \otimes \mathbf{m}_i) \right]^{-1} \sum_i (\mathbf{W}_i^* \otimes \mathbf{m}_i)^{\mathrm{T}} \cdot$$

$$\cdot \left(\hat{\Omega} \otimes \sum_i \mathbf{m}_i \mathbf{m}_i^{\mathrm{T}} \right)^{-1} \sum_i (\mathbf{y}_i^* \otimes \mathbf{m}_i) \tag{3.60}$$

gdzie $\mathbf{m}_i$ jest określone wzorem (3.59).

W powyższych rozważaniach dopuszcza się, że składnik losowy ε_{it} wykazuje korelację w czasie, a zatem zmienna objaśniająca $y_{i,t-1}$ jest zmienną endogeniczną. Przy założeniu, że ε_{it} nie wykazuje korelacji w czasie, prawdziwe są dodatkowe warunki momentów, co pozwala w inny sposób skonstruować macierz $\mathbf{M}_i$. Przy założeniu, że macierz $\mathbf{C}$, oprócz poprzednio przyjętych warunków, jest macierzą trójkątną górną, transformowany składnik losowy z dowolnego okresu t (dowolny element wektora $\mathbf{u}_i^+$) jest niezależny od α_i oraz ε_{i1}, ..., ε_{it-1}. W tym przypadku y_{i0}, y_{i1}, ..., $y_{i,t-1}$ są dodatkowymi instrumentami dla równania, odpowiadającego okresowi t, i macierz instrumentów $\mathbf{M}_i$ przyjmuje postać:

$$\mathbf{M}_i = \begin{bmatrix} (\mathbf{w}_i^{\mathrm{T}}, y_{i0}) & \cdots & \cdots & & \mathbf{0} \\ \mathbf{0} & (\mathbf{w}_i^{\mathrm{T}}, y_{i0}, y_{i1}) & & \vdots & \mathbf{0} \\ \vdots & & \ddots & & \\ \mathbf{0} & \cdots & \cdots & (\mathbf{w}_i^{\mathrm{T}}, y_{i0}, ..., y_{i,T-2}) & \mathbf{0} \\ \mathbf{0} & & & \mathbf{0} & \mathbf{m}_i^{\mathrm{T}} \end{bmatrix} \tag{3.61}$$

Estymator GMM, określony wzorem (3.54), w którym macierz $\mathbf{M}_i$ dana jest wzorem (3.61), $\mathbf{M} = [\mathbf{M}_1^{\mathrm{T}}, ..., \mathbf{M}_N^{\mathrm{T}}]^{\mathrm{T}}$ jest, jak wykazali Arellano i Bover [1995], niezależny od wyboru macierzy $\mathbf{C}$, o ile spełnia ona wymienione powyżej założenia.

3.2.5. Estymator Ahna i Schmidta

Ahn i Schmidt [1995] wskazali dodatkowe warunki momentów, które mogą być wykorzystywane do estymacji panelowych modeli dynamicznych, jeśli liczba obiektów $N \to \infty$, a wymiar czasowy próby T jest skończony. Wykorzystanie tych warunków poprawia efektywność estymatora Arellano i Bonda [1991]. Warunki te są nieliniowe, ale autorzy wskazują możliwość przekształcenia ich do warunków liniowych.

Rozważania A h n a i S c h m i d t a [1995] koncentrują się początkowo na modelu autoregresyjnym, a następnie rozszerzone są na model ze zmiennymi egzogenicznymi. Punktem wyjścia jest więc model (3.23), postaci:

$$y_{it} = \gamma y_{i,t-1} + u_{it} = \gamma y_{i,t-1} + \alpha_i + \varepsilon_{it} \tag{3.62}$$

gdzie $u_{it} = \alpha_i + \varepsilon_{it}$, $|\gamma| < 1$, oraz spełnione są założenia nazywane w literaturze „założeniami standardowymi": $E(\alpha_i) = E(\varepsilon_{it}) = 0$, $E(\alpha_i\varepsilon_{it}) = 0$, $E(\varepsilon_{is}\varepsilon_{it}) = 0$ dla $i = 1, ..., N$, $t \neq s$, $E(y_{i0}\varepsilon_{it}) = 0$ dla $i = 1, ..., N$, $t = 1, ..., T$.

Jak zauważyli A h n i S c h m i d t [1995], dla modelu (3.62), spełnionych jest $T(T-1)/2 + (T-2)$ warunków momentów postaci:

$$E(y_{is}\Delta u_{it}) = 0 \quad \text{dla } t = 2, ..., T,\ s = 0, ..., t-2 \tag{3.63}$$

oraz

$$E(u_{iT}\Delta u_{it}) = 0 \quad \text{dla } t = 2, ..., T-1 \tag{3.64}$$

Układ $T(T-1)/2$ warunków (3.63) został wykorzystany przez A r e l l a n o i B o n d a [1991], zaś układ $(T-2)$ warunków (3.64) jest dołączony przez S. C. Ahna i P. Schmidta. Układ (3.63) i (3.64) zawiera wszystkie możliwe warunki momentów, wynikające z założeń modelu (3.62). Formalny dowód tego twierdzenia, wymagający rozbudowanego zapisu macierzowego, oraz nieformalne jego uzasadnienie znaleźć można w artykule: A h n, S c h m i d t [1995] i w: B a l t a g i [2008] (podrozdz. 8.4). Wykorzystanie dodatkowych warunków (3.64) poprawia efektywność estymacji, co zauważyć można, jeśli model (3.62) zapisany zostanie w postaci:

$$\begin{cases} \Delta y_{it} = \gamma \Delta y_{i,t-1} + \Delta u_{it} \\ y_{iT} = \gamma \Delta y_{i,T-1} + u_{iT} \end{cases} \quad \text{dla } t = 2, 3, ..., T \tag{3.65}$$

Zastosowanie FDGMM Arellano–Bonda sprowadza się do oszacowania pierwszych $T-1$ równań układu (3.65), ze zbiorem instrumentów wskazanych przez A r e l l a n o i B o n d a [1991], opisanych w pkt 3.2.2. Ostatnie równanie zostało wykorzystane dopiero przez S. C. A h n a i P. S c h m i d t a [1995]. Mimo że nie można wskazać odpowiednich zmiennych instrumentalnych, równanie to wnosi dodatkowe informacje, ze względu na założenie, że u_{iT} jest nieskorelowane z Δu_{it} dla $t = 2, ..., T-1$.

A h n, S c h m i d t [1995] rozważyli też skutki przyjęcia dwóch dodatkowych, alternatywnych założeń: homoskedastyczności składnika losowego ε_{it} oraz „stacjonarności".

Założenie, że wariancja składnika losowego ε_{it} jest stała w czasie:

$$\mathrm{E}(\varepsilon_{it}^2) = \sigma_i^2 \quad \text{dla } t = 1, ..., T \tag{3.66}$$

implikuje $T-1$, dodatkowych w stosunku do (3.63) i (3.64), nieliniowych warunków momentów. Tak więc przy założeniu homoskedastyczności ε_{it} łączna liczba warunków jest równa $T(T-1)/2 + (2T-3)$. Warunki te można przekształcić do postaci liniowej, która jest określona przez (3.63), łącznie z:

$$\mathrm{E}(y_{it}\Delta u_{i,t+1} - y_{i,t+1}\Delta u_{i,t+2}) = 0 \quad \text{dla } t = 1, ..., T-2 \tag{3.67}$$

$$\mathrm{E}\left(\Delta u_{i,t+1}\frac{1}{T}\sum_{t=1}^{T} u_{it}\right) = 0\,, \quad \text{dla } t = 1, ..., T-1 \tag{3.68}$$

Natomiast przy założeniu „stacjonarności", a więc że:

$$\mathrm{E}(\alpha_i y_{it}) \text{ jest jednakowa dla } t = 1, ..., T \tag{3.69}$$

i utrzymaniu założeń modelu (3.62) oraz założenia (3.66), liczba warunków momentów wzrasta o jeden. Zbiór wszystkich $T(T-1)/2 + (2T-2)$ warunków można w sposób liniowy zapisać jako (3.63) oraz:

$$\mathrm{E}(u_{iT}\Delta y_{it}) = 0 \quad \text{dla } t = 1, ..., T-1 \tag{3.70}$$

$$\mathrm{E}(u_{it}y_{it} - u_{i,t-1}y_{i,t-1}) = 0 \quad \text{dla } t = 2, ..., T \tag{3.71}$$

Ahn i Schmidt [1995] udowadniają, że ponieważ wykorzystanie w GMM warunków momentów (3.63) i (3.64) oznacza efektywne wykorzystanie informacji dostarczanych przez drugie momenty, to uzyskany w ten sposób estymator GMM jest asymptotycznie równoważny optymalnemu estymatorowi minimalnej odległości (ang. *optimal minimum distance*) Chamberlaina [1982, 1984]. Estymator GMM jest jednak preferowany, gdyż jest łatwiejszy obliczeniowo, a ponadto jego zgodność wynika ze słabszych założeń niż zgodność estymatora Chamberlaina.

S. C. Ahn i P. Schmidt dokonali też porównania asymptotycznej efektywności estymatora GMM, wykorzystującego zaproponowane przez nich warunki momentów, z estymatorem FDGMM. Porównane zostały trzy estymatory GMM: i) estymator $\hat{\gamma}_{\mathrm{FDGMM}}$, wykorzystujący warunki momentów (3.63), ii) estymator $\hat{\gamma}_{\mathrm{AS1}}$, wykorzystujący układ warunków (3.63) i (3.64), oraz

iii) estymator $\hat{\gamma}_{\mathrm{AS2}}$, który wykorzystuje warunki (3.63), (3.67) i (3.68), a więc wymaga założenia o homoskedastyczności ε_{it}. Obliczone zostały stosunki wariancji poszczególnych par estymatorów dla różnych wartości T, γ i $\sigma_\alpha^2/\sigma_\varepsilon^2$. W tekście zamieszczono wyniki uzyskane dla dwóch przypadków: przy założeniu, że zmienna y jest stacjonarna i bez tego założenia. Przy założeniu stacjonarności zmiennej y wzrost efektywności estymacji, uzyskany dzięki wykorzystaniu dodatkowych warunków, jest znaczący. Efektywność estymacji, mierzona stosunkiem wariancji $\mathrm{D}^2(\hat{\gamma}_{\mathrm{FDGMM}})/\mathrm{D}^2(\hat{\gamma}_{\mathrm{AS1}})$ rośnie ze wzrostem T, co autorzy uznają za wynik nieoczekiwany, gdyż udział liczby dodatkowych warunków (3.64) w łącznej ich liczbie maleje ze wzrostem T. Efektywność estymacji rośnie też ze wzrostem stosunku wariancji $\sigma_\alpha^2/\sigma_\varepsilon^2$ oraz jeśli $\gamma \to 1$. Oznacza to, że wskazane przez AS warunki momentów (3.64) dostarczają właściwych informacji również w tych przypadkach, w których estymator $\hat{\gamma}_{\mathrm{FDGMM}}$ traci efektywność. Efektywność estymacji, uzyskana dzięki wykorzystaniu dodatkowych warunków i założeniu o homoskedastyczności ε_{it}, mierzona stosunkiem $\mathrm{D}^2(\hat{\gamma}_{\mathrm{FDGMM}})/\mathrm{D}^2(\hat{\gamma}_{\mathrm{AS2}})$ nie jest już na ogół znacząco lepsza. Rośnie ona szczególnie, jeśli $\gamma \to 1$, a $T = 3$ (por. A h n, S c h m i d t [1995], tab. 1, s. 15). Podobne wyniki uzyskano w przypadku, gdy nie zakładano stacjonarności zmiennej y. Zatem wzrost efektywności, niezależnie od tego, czy y jest zmienną stacjonarną czy niestacjonarną, wynika przede wszystkim z wykorzystania dodatkowych warunków momentów, a w niewielkim tylko stopniu z założenia homoskedastyczności.

S. C. Ahn i P. Schmidt rozważali też model, w którym oprócz opóźnionej zmiennej endogenicznej, występują niezależne zmienne objaśniające. Wśród zmiennych objaśniających wyróżniono, podobnie jak w H a u s m a n i T a y l o r [1981] i w A r e l l a n o i B o v e r [1995], zmienne stałe w czasie ($\mathbf{Z}_i$) i takie, których wartości zmieniają się w czasie ($\mathbf{X}_{it}$). Wyróżniono też podzbiory $\mathbf{X}_{1,it}$ i $\mathbf{Z}_{1i}$ zmiennych nieskorelowanych z efektami grupowymi α_i (oznacza to innymi słowy: dopuszczenie korelacji części zmiennych objaśniających α_i). Założono jednak ścisłą egzogeniczność wszystkich zmiennych w sensie braku ich korelacji ze składnikiem losowym ε_{it}. Model Ahna i Schmidta postaci:

$$y_{it} = \gamma y_{i,t-1} + \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\varphi} + \alpha_i + \varepsilon_{it} = \gamma y_{i,t-1} + \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\varphi} + u_{it} \qquad (3.72)$$

jest więc równoważny modelowi Arellano–Bover postaci (3.56).

Założenia modelu (3.72), dotyczące ścisłej egzogeniczności zmiennych objaśniających oraz ich korelacji z efektami α_i, wraz z „założeniami standardowymi” implikują, oprócz warunków momentów (3.63) i (3.64), spełnienie układu warunków postaci:

$$\begin{cases} \mathrm{E}(\mathbf{X}_{1,it}u_{is}) = \mathrm{E}(\mathbf{Z}_{1,i}u_{it}) = \mathbf{0} \\ \text{dla dowolnego } t \ \mathrm{E}(\mathbf{X}_{2,it}u_{is}) \text{ jest jednakowe dla wszystkich } s \\ \mathrm{E}(\mathbf{Z}_{2,i}u_{it}) \text{ jest jednakowe dla wszystkich } t \end{cases} \tag{3.73}$$

Przy dodatkowym założeniu homoskedastyczności ε_{it} układ warunków momentów składa się z równań (3.63), (3.67), (3.68) i (3.73). Dołączenie kolejnego założenia:

$$\mathrm{E}(\mathbf{X}_{2,it}\alpha_i) \text{ jest dla dowolnego } i \text{ jednakowa dla wszystkich } t \tag{3.74}$$

pozwala zastąpić układ warunków (3.73) analogicznym układem, w którym zmienne **X** i **Z** zastąpione są odchyleniami od średnich grupowych tych zmiennych.

Ahn, Schmidt [1995] wykazali, że dzięki założeniu o ścisłej egzogeniczności wszystkich zmiennych **X** i **Z** dostępne są do wykorzystania warunki momentów, w których jako instrumenty wykorzystuje się odchylenia od średnich grupowych tych zmiennych.

3.2.6. Systemowy estymator GMM Blundella i Bonda

Jak wspomniano w pkt 3.2.2, liczba instrumentów możliwych do wykorzystania dla modelu pierwszych różnic maleje, jeśli stopień trwałości szeregów czasowych stają się wyższy. Inaczej, ze wzrostem stopnia trwałości zmiennych, stosowanych jako instrumenty, słabnie korelacja między ich poziomami a odpowiednimi przyrostami. Zrodziło to konieczność poszukiwania alternatywnych w stosunku do propozycji Arellano i Bonda [1991] metod estymacji panelowych modeli dynamicznych. Blundell i Bond [1998] zaproponowali estymator GMM, który wykorzystuje zarówno równania na przyrostach, jak i na poziomach. Bazuje on na propozycjach Arellano, Bonda [1991], oraz Arellano, Bover [1995]. Blundell i Bond wykazali, że warunki momentów dla równań na poziomach, wskazane przez Arellano, Bover [1995], postaci (3.53), pozostają właściwe również w tych przypadkach, w których instrumenty wykorzystywane w równaniach na przyrostach stają się słabe. Powiązali oni ponadto warunki Arellano i Bover z procesem generującym wartości początkowe, oraz pokazali, że warunki te dostarczają informacji, które mogą być wykorzystane w celu skonstruowania dodatkowych warunków momentów. W tym sensie praca Blundella i Bonda jest nawiązaniem do idei Andersona i Hsiao [1981, 1982], którzy pierwsi zwrócili uwagę na znaczenie warunków początkowych w procesie estymacji dynamicznych modeli panelowych.

W części teoretycznej artykułu B l u n d e l l i B o n d [1998] skoncentrowali się na modelu autoregresyjnym postaci (3.23), podkreślając, że wszystkie wnioski są prawdziwe również dla modelu ze zmiennymi egzogenicznymi. Do założeń modelu postaci (3.23) dodano, powołując się na pracę A h n a i S c h m i d t a [1995], tzw. standardowy warunek początkowy dla wartości y_{i1} postaci:

$$\mathrm{E}(y_{i1}\varepsilon_{it}) = 0 \quad \text{dla } i = 1, ..., N, t = 2, ..., T \tag{3.75}$$

Na wstępie skoncentrowano się na przypadku, gdy $T = 3$, a zatem istnieje tylko jeden warunek momentu postaci: $\mathrm{E}(y_{i1}\ \Delta\varepsilon_{i3}) = 0$, skąd wynika, że γ jest jednoznacznie identyfikowalne. W takim przypadku jednostopniowy estymator FDGMM redukuje się do estymatora metody zmiennych instrumentalnych, wyznaczonego z równania regresji:

$$\Delta y_{i2} = (\gamma - 1) y_{i1} + \alpha_i + \varepsilon_{i2} \tag{3.76}$$

B l u n d e l l i B o n d [1998] wykazali, że estymator parametru (γ – 1) jest obciążony w górę:

$$p\lim(\hat{\gamma} - 1) = (\gamma - 1)\frac{k}{k + (\sigma_\alpha^2 / \sigma_\varepsilon^2)} \tag{3.77}$$

gdzie $k = (1-\gamma)/(1+\gamma)$. Wykazali oni też, że wartość statystyki F dla równania (3.76) jest zbliżona do rozkładu χ^2 z 1 stopniem swobody i parametrem τ takim, że:

$$\tau = \frac{(\sigma_\varepsilon^2 k)^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2 k} \to 0 \qquad \text{dla } \gamma \to 0 \tag{3.78}$$

Zatem jeśli $\gamma \to 1$ lub stosunek wariancji efektów grupowych do wariancji właściwego składnika losowego jest zbyt duży, to estymator FDGMM staje się obciążony i nieefektywny. Jako przyczynę obciążenia i braku efektywności estymatora FDGMM wskazano problem słabych instrumentów.

Następnie B l u n d e l l i B o n d [1998] pokazali, że dołączenie do założeń modelu (3.23) dodatkowego założenia dla warunków początkowych powoduje, iż możliwe są do wykorzystania dodatkowe warunki momentów. Założenie to, nazwane warunkiem miękkiej stacjonarności (ang. *mild stationarity restriction*) ma postać:

$$\mathrm{E}\left[\left(y_{i1}-\frac{\alpha_i}{1-\gamma})\alpha_i\right)\right]=0 \tag{3.79}$$

Oznacza ono, że wartość początkowa y_{i1} nie różni się systematycznie od wartości $\frac{\alpha_i}{1-\gamma}$, która jest średnią wartością szeregu y_{it}. Szereg y_{it} jest przy tym zbieżny do $\frac{\alpha_i}{1-\gamma}$ dla każdego i i t = 2, ..., T. Założenie (3.79) jest założeniem dotyczącym stacjonarności średniej[12] i wynika z niego warunek:

$$\mathrm{E}(\alpha_i \Delta y_{i2}) = 0 \text{ dla } i = 1, ..., N \tag{3.80}$$

Równoważność warunku (3.80) i warunku (3.79), będącego ograniczeniem nałożonym na warunki początkowe, generujące y_{i1}, można wykazać w następujący sposób.

Bez straty ogólności zapisać można, że[13]:

$$y_{i1}=\frac{\alpha_i}{1-\gamma}+u_{i1} \tag{3.81}$$

Zapisując postać równania (3.23) dla pierwszego okresu, w którym zależność jest obserwowalna, otrzymuje się:

$$y_{i2}=\gamma y_{i,1}+\alpha_i+\varepsilon_{i2}.$$

Odejmując od obu stron tego równania wartość y_{i1}, otrzymuje się równanie (3.76), a następnie, wstawiając (3.81):

$$\Delta y_{i2}=(\gamma-1)\frac{\alpha_i}{1-\gamma}+(\gamma-1)u_{i1}+\alpha_i+\varepsilon_{i2}=(\gamma-1)u_{i1}+\varepsilon_{i2}.$$

Stąd, po uwzględnieniu założeń modelu (3.23), wynika, że założenie (3.80) jest równoważne ograniczeniu $\mathrm{E}(u_{i1}\alpha_i) = 0$ dla i = 1, ..., N, a zatem warunkowi (3.79). Warunkiem wystarczającym jest zatem, aby początkowe wartości y_{i1} dla wszystkich obiektów i spełniały założenie słabej stacjonarności:

---

[12] Wyróżniamy dwa rodzaje niestacjonarności: w średniej (trend deterministyczny) i wariancji (trend stochastyczny) procesu.

[13] Wymaga to przyjęcia proponowanego szeroko w literaturze założenia (por. np. W o o l d r i d g e [2002], s. 412), że $y_{i0} = y_{i1}$.

$\mathrm{E}(y_{i1}|\alpha_i) = \alpha_i/(1-\gamma)$. Spełnienie tego warunku wymaga jedynie, aby pierwsze momenty y_{it} były stałe, stałość drugich momentów nie jest zaś wymagana.

Warunek (3.80) w połączeniu z założeniami modelu (3.23) implikuje $T - 2$ dodatkowych warunków momentów postaci[14]:

$$\mathrm{E}(u_{it}\ \Delta y_{i,t-1}) = 0 \quad \text{dla } i = 1, ..., N,\ t = 3, 4, ..., T \tag{3.82}$$

gdzie $u_{it} = \alpha_i + \varepsilon_{it}$. Koncentrując się ponownie na przypadku $T = 3$, Blundell i Bond [1998] wykazali, że warunki (3.82) dostarczają właściwych informacji również w przypadku, gdy $\gamma \to 1$ lub stosunek wariancji $\sigma_\alpha^2/\sigma_\varepsilon^2$ jest duży. Z założeń (3.82) wynika, że możliwe jest wykorzystanie opóźnionych pierwszych różnic zmiennej y ($\Delta y_{i,t-1}$) jako instrumentów dla równań na poziomach. Jest to zgodne z propozycją Arellano i Bover [1995]. Warto ponadto zauważyć, że możliwymi instrumentami, oprócz $\Delta y_{i,t-1}$, są przyrosty opóźnione o więcej niż jeden okres (analogicznie, jak wynika to z warunku (3.25) dla estymatora FDGMM), ale ich wykorzystanie przy konstrukcji estymatora Blundella i Bonda jest zbędne właśnie ze względu na warunki (3.25).

Systemowy estymator GMM Blundella i Bonda (SGMM) wykorzystuje oba układy warunków momentów: (3.25) i (3.82). W procesie estymacji wykorzystuje się zatem system ($T - 2$) równań na przyrostach i ($T - 2$) równań na poziomach, odpowiadających okresom $t = 3, ..., T$, dla których przyrosty $\Delta y_{i,t-1}$ są określone. Innymi słowy, do warunków i równań wykorzystywanych przez FDGMM dołączony jest dodatkowy układ równań (na poziomach), w których instrumentami są opóźnione pierwsze różnice. Instrumenty te mogą być wykorzystane, ponieważ mimo że efekty grupowe α_i są skorelowane z poziomami y_{it}, to założenie (3.80) gwarantuje, że nie są one skorelowane z przyrostami Δy_{it}.

Warunki (3.82) zapisać można w postaci macierzowej:

$$\mathrm{E}\left(\mathbf{Z}_i^{+\mathrm{T}}\mathbf{u}_i^{+}\right) = \mathbf{0} \tag{3.83}$$

gdzie:

$$\mathbf{Z}_i^{+} = \begin{bmatrix} \mathbf{Z}_i & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Delta y_{i2} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Delta y_{i3} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \ddots & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \Delta y_{i,T-1} \end{bmatrix} \tag{3.84}$$

[14] Warunki te są właściwe, podobnie jak warunki wykorzystywane przez FDGMM, niezależnie od tego, czy składnik losowy ε_{it} jest homoskedastyczny czy nie.

natomiast macierz $\mathbf{Z}_i$ jest określona, jak we wzorze (3.27), a $\mathbf{u}_i^+ = [\Delta\varepsilon_{i3}, ..., \Delta\varepsilon_{iT}, u_{i3}, ..., u_{iT}]^{\mathrm{T}}$.

Właściwość dodatkowych instrumentów $\Delta y_{i,t-1}$ można testować stosując test Sargana warunków ponadidentyfikujących lub porównując estymatory FDGMM i SGMM za pomocą różnicowego testu Sargana lub testu Hausmana. Konstrukcja tych testów jest omówiona w podrozdz. 3.3.

Podobnie jak w przypadku estymatora FDGMM, opisaną powyżej metodę można zmodyfikować i zastosować do estymacji modelu, w którym występują niezależne zmienne objaśniające $\mathbf{x}_{it}$. Model (3.23) przyjmuje wtedy postać:

$$y_{it} = \gamma y_{i,t-1} + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + \alpha_i + \varepsilon_{it} \tag{3.85}$$

a więc identyczną z modelem (3.11), ze wszystkimi jego założeniami. Jeśli zmienne $\mathbf{x}_{it}$ są skorelowane z efektami grupowymi α_i, postać warunków momentów dla zmiennych $\mathbf{x}_{it}$ zależy od tego, czy zmienne te są ściśle egzogeniczne czy nie. W przypadku, gdy są one ściśle egzogeniczne, to postępowanie jest takie samo, jak w przypadku estymatora FDGMM. Jeśli zaś zmienne $\mathbf{x}_{it}$ nie są ściśle egzogeniczne, oznacza to, że bieżące i opóźnione wartości tych zmiennych są skorelowane z bieżącymi wartościami składnika losowego. Wówczas, oprócz warunków określonych wzorem (3.40), można uzyskać dodatkowe warunki momentów przyjmując założenie początkowe typu (3.80) dla zmiennych $\mathbf{x}_{it}$. Ma ono postać:

$$\mathrm{E}(\alpha_i \Delta\mathbf{x}_{it}) = 0 \quad \text{dla } i = 1, ..., N,\ t = 2, ..., T \tag{3.86}$$

a więc oznacza brak korelacji między pierwszymi różnicami zmiennych $\mathbf{x}_{it}$ a efektami grupowymi. Stąd wynikają warunki momentów postaci:

$$\mathrm{E}(\Delta\mathbf{x}_{it-1} u_{it}) = 0 \quad \text{dla } i = 1, ..., N \text{ oraz } t = 3, ..., T \tag{3.87}$$

Zatem opóźnione pierwsze różnice $\Delta\mathbf{x}_{it-1}$ mogą być zmiennymi instrumentalnymi w równaniach na poziomach i zespół wszystkich możliwych warunków momentów tworzą warunki (3.40) i (3.87).

Postać ogólna estymatora SGMM jest dana wzorem (3.31) dla modelu autoregresyjnego i (3.43) – dla modelu z niezależnymi zmiennymi objaśniającymi. Należy jedynie uwzględnić konieczność odpowiedniej modyfikacji macierzy zmiennych instrumentalnych $\mathbf{Z}_i$ występującej w tych wzorach. Asymptotyczne wariancje jedno- i dwustopniowego estymatora SGMM określone są odpowiednio wzorami analogicznymi do (3.44) i (3.45). Skorygowany estymator wariancji Windmeijera dla estymatora dwustopniowego wyznaczyć można ze wzoru (3.46).

B l u n d e l l i B o n d [1998] wykazali, że zastosowanie systemowego estymatora GMM przynosi znaczący wzrost efektywności w stosunku do FDGMM szczególnie, gdy wartość parametru autoregresyjnego $\gamma \to 1$ lub gdy stosunek wariancji $\sigma_\alpha^2/\sigma_\varepsilon^2$ rośnie.

W symulacji Monte Carlo wykazano, że dla $T = 4$ i $\sigma_\alpha^2/\sigma_\varepsilon^2 = 1$ stosunek asymptotycznej wariancji estymatora FDGMM do wariancji estymatora SGMM równa się 1,75 dla $\gamma = 0$ i wzrasta do 3,26 dla $\gamma = 0{,}5$ oraz aż do 55,4 dla $\gamma = 0{,}9$. Efektywność SGMM rośnie więc niezwykle szybko ze wzrostem wartości parametru γ. W tym samym eksperymencie wykazano, że wzrost czasowej liczebności próby poprawia nieco własności estymatora FDGMM, ale nie jest to poprawa systematyczna. Wyniki eksperymentu B l u n d e l l a i B o n d a [1998] zostały potwierdzone w pracach B l u n d e l l a i B o n d a [2000], oraz B l u n d e l l a, B o n d a i W i n d m e i j e r a [2000]. H a h n [1999] badał implikacje kluczowego założenia dla warunków początkowych postaci (3.79). Dokonał on numerycznego porównania semiparametrycznych granic informacyjnych Rao–Cramera dla przypadku, kiedy założenie (3.79) jest uwzględnione oraz bez jego uwzględniania. W ten sposób stwierdzony został bardzo istotny wzrost efektywności estymacji.

3.2.7. Skorygowany estymator wewnątrzgrupowy Kivieta

Opisany w podrozdz. 3.1 estymator wewnątrzgrupowy jest w przypadku dynamicznych modeli panelowych obciążony w dół. Propozycja K i v i e t a [1995] polega na wyznaczeniu wartości poprawki, o jaką należy skorygować ten estymator, aby zlikwidować jego obciążenie dla małych prób, tzn. skończonych wartości T i N. W prezentowanych, w artykule, eksperymentach Monte Carlo Kiviet wykazał, że w skończonych próbach estymator wewnątrzgrupowy z korektą ma lepsze własności niż inne estymatory, w szczególności estymator A n d e r s o n a i H s i a o [1981, 1982] i estymator FDGMM A r e l l a n o i B o n d a [1991].

Model rozważany przez K i v i e t a [1995], zawierający ściśle egzogeniczne zmienne objaśniające x_{kit}, dla $k = 1, ..., K$, tworzące wektor $\mathbf{x}_{it}$, ma postać:

$$y_{it} = \gamma y_{i,t-1} + \mathbf{x}_{it}^{\mathrm{T}}\boldsymbol{\beta} + \alpha_i + \varepsilon_{it}, \; i = 1, ..., N, \; t = 1, ..., T \tag{3.88}$$

Zbiór założeń przyjętych przez Kivieta jest nieco zmodyfikowany w stosunku do zbioru standardowego, opisanego dla modelu (3.11), dlatego zostanie tutaj przedstawiony w całości. Założenia modelu (3.88) są następujące: składnik losowy ε_{it} jest homoskedastyczny i nie wykazuje korelacji w czasie ani względem obiektów $\mathrm{E}(\varepsilon_{js}\varepsilon_{it}) = 0$ dla $i \neq j$, $t \neq s$; nieobserwowalna heterogeniczność

obiektów, wyrażana poprzez efekty grupowe, ma charakter losowy i efekty dla różnych obiektów nie są skorelowane: $E(\alpha_i \alpha_j) = 0$ dla $i \neq j$; efekty grupowe są nieskorelowane z właściwym składnikiem losowym: $E(\varepsilon_{jt}\, \alpha_i) = 0$ dla dowolnych i, j, s; zmienne objaśniające są ściśle egzogeniczne: $E(x_{it}\, \varepsilon_{js}) = 0$ dla dowolnych i, j, t, s; zmienne objaśniające mogą, ale nie muszą, być skorelowane z efektami grupowymi. Dodatkowo założono normalność rozkładu efektów grupowych i właściwego składnika losowego: $\alpha_i \sim \mathrm{N}(0, \sigma_\alpha^2)$, gdzie $\sigma_\alpha^2 \geq 0$, oraz $\varepsilon_{it} \sim \mathrm{N}(0, \sigma_\varepsilon^2)$, gdzie $\sigma_\varepsilon^2 > 0$. Zakłada się też, że $|\gamma| < 1$, tzn. że długookresowa zależność między potencjalnie niestacjonarnymi zmiennymi x_{kit} i y_{it} jest stabilna. Model Kivieta jest więc równoważny modelowi nazwanemu przez A n d e r s o n a i H s i a o [1982] modelem zależności stanowej (por. wzory (3.18) i (3.19)). Kiviet przyjmuje ponadto założenie, że wartość początkowa y_{i0} jest losowa, przy czym: $E(y_{i0}\varepsilon_{jt}) = 0$ dla dowolnych i, j, t, oraz $E(y_{i0}\alpha_j)$ jest dowolne dla dowolnych i, j.

Asymptotyczne obciążenie estymatora wewnątrzgrupowego w przypadku panelowego modelu dynamicznego było analizowane w literaturze wcześniejszej niż artykuł K i v i e t a [1995]. Odpowiednie formuły, pozwalające wyznaczyć jego wielkość w różnych sytuacjach, w szczególności w zależności od tego, czy model zawiera zmienne egzogeniczne, oraz przy różnych założeniach odnośnie do wartości początkowej y_{i0}, zaproponowane zostały przez: T r o g n o n a [1978], N i c k e l l a [1981], S e v e s t r e'a i T r o g n o n a [1985]. Dwie ostatnie są przytoczone w niniejszej publikacji i dane wzorami (3.4), (3.5) i (3.7). K i v i e t [1995] zauważył jednak, że formuły te nie są użyteczne w praktyce, a ponadto mogą dawać bardzo niedokładne wyniki w przypadku małych prób. Proponuje on własną formułę, pozwalającą wyznaczyć korektę obciążenia estymatora wewnątrzgrupowego.

Dla uproszczenia dalszego wywodu i w celu zaprezentowania formuły, pozwalającej wyznaczyć korektę Kivieta, przypomnieć trzeba używany w matematyce symbol $O(x)$. Jeśli dane są dwie funkcje: $f_1(x)$ i $f_2(x)$ określone w zbiorze D, $c > 0$ jest dowolną stałą niezależną od x oraz w każdym punkcie zbioru D spełniona jest zależność $|f_1(x)| \leq c|f_2(x)|$, to mówimy, że f_1 jest co najwyżej rzędu f_2, co oznaczamy symbolem $f_1(x) = O(f_2(x))$.

Model (3.88), zapisać można w postaci macierzowej jako:

$$\mathbf{y} = \mathbf{W}\boldsymbol{\delta} + (\mathbf{I}_N \otimes \mathbf{i}_T)\boldsymbol{\alpha} + \boldsymbol{\varepsilon} \tag{3.89}$$

gdzie: $\mathbf{y} = [y_{11}, ..., y_{1T}, ..., y_{N1}, ..., y_{NT}]^{\mathrm{T}}$, $\mathbf{W} = [\mathbf{y}_{-1} \vdots \mathbf{X}]$, $\mathbf{y}_{-1} = [y_{10}, ..., y_{1,T-1}, ..., y_{N0}, ..., y_{N,T-1}]^{\mathrm{T}}$, $\mathbf{X} = [\mathbf{X}_1, ..., \mathbf{X}_N]^{\mathrm{T}}$, $\mathbf{X}_i = [\mathbf{x}_{i1}^{\mathrm{T}}, ..., \mathbf{x}_{iT}^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{\delta} = [\boldsymbol{\gamma}\, \boldsymbol{\beta}^{\mathrm{T}}]^{\mathrm{T}}$, $\boldsymbol{\varepsilon}$ jest wektorem składników losowych, określonym analogicznie jak wektor $\mathbf{y}$,

zaś $\boldsymbol{\alpha} = [\alpha_1, ..., \alpha_N]^{\mathrm{T}}$ jest wektorem efektów grupowych. $\mathbf{I}_N$ jest macierzą jednostkową stopnia N, a $\mathbf{i}_T$ jest wektorem jedynek o wymiarach (T × 1). Na s. 61–65 artykułu K i v i e t a [1995] przedstawiono formalny dowód twierdzenia, które określa poprawkę, nazywaną w późniejszej literaturze przedmiotu poprawką Kivieta. Treść tego twierdzenia jest następująca.

Twierdzenie 1K.

Jeśli spełnione są założenia modelu (3.89)[15], oraz $\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{x}_{it}\mathbf{x}_{it}^{\mathrm{T}}$ jest zbieżne do macierzy nieosobliwej o skończonych elementach (co oznacza stacjonarność zmiennych egzogenicznych), to:

$$\begin{aligned}\mathrm{E}(\hat{\boldsymbol{\delta}}_{\mathrm{W}} - \boldsymbol{\delta}) = -\sigma_\varepsilon^2 \mathbf{D}^{-1}\Bigg(&\frac{N}{T}(\mathbf{i}_T^{\mathrm{T}}\mathbf{C}\mathbf{i}_T)[2\mathbf{q} - \overline{\mathbf{W}}^{\mathrm{T}}\mathbf{A}\overline{\mathbf{W}}\mathbf{D}^{-1}\mathbf{q}] + \\ &+ \mathrm{tr}\{\overline{\mathbf{W}}^{\mathrm{T}}(\mathbf{I}_N \otimes \mathbf{A}_T\mathbf{C}\mathbf{A}_T)\overline{\mathbf{W}}\mathbf{D}^{-1}\}\mathbf{q} + \overline{\mathbf{W}}^{\mathrm{T}}(\mathbf{I}_N \otimes \mathbf{A}_T\mathbf{C}\mathbf{A}_T)\overline{\mathbf{W}}\mathbf{D}^{-1}\mathbf{q} + \\ &+ \sigma_\varepsilon^2 N\mathbf{q}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{q}\cdot\left[-\frac{N}{T}(\mathbf{i}_T^{\mathrm{T}}\mathbf{C}\mathbf{i}_T)\mathrm{tr}\{\mathbf{C}^{\mathrm{T}}\mathbf{A}_T\mathbf{C}\} + 2\mathrm{tr}\{\mathbf{C}^{\mathrm{T}}\mathbf{A}_T\mathbf{C}\mathbf{A}_T\mathbf{C}\}\right]\mathbf{q}\Bigg) + \\ &+ O\left(N^{-1}T^{-3/2}\right)\end{aligned} \tag{3.90}$$

gdzie $\hat{\boldsymbol{\delta}}_{\mathrm{W}}$ oznacza estymator wewnątrzgrupowy wektora $\boldsymbol{\delta}$, oraz: $\mathbf{D} = \overline{\mathbf{W}}^{\mathrm{T}}\mathbf{A}\overline{\mathbf{W}} + \sigma_\varepsilon^2 N\mathrm{tr}\{\mathbf{C}^{\mathrm{T}}\mathbf{A}_T\mathbf{C}\}\mathbf{q}\mathbf{q}^{\mathrm{T}}$, $\overline{\mathbf{W}} = \mathrm{E}(\mathbf{W}) = [\mathrm{E}(\mathbf{y}_{-1}) \vdots \mathbf{X}]$, $\mathbf{A} = \mathbf{I}_N \otimes \mathbf{A}_T$, $\mathbf{A}_T = \mathbf{I}_T - \frac{1}{T}\mathbf{i}_T\mathbf{i}_T^{\mathrm{T}}$, $\mathbf{q}$ = [1, 0, ..., 0]^{T} jest wektorem o wymiarach ((K + 1) × 1), $\mathbf{C}$ jest macierzą stopnia T, określoną jako:

$$\mathbf{C} = \begin{bmatrix} 0 & \dots & \dots & \dots & \dots & 0 \\ 1 & 0 & & & & \vdots \\ \gamma & 1 & 0 & & & \\ \gamma^2 & \gamma & 1 & & & \\ \vdots & & & & & \vdots \\ \gamma^{T-2} & & & \dots & 1 & 0 \end{bmatrix}.$$

Powyższe twierdzenie głosi, że jeśli do estymacji wektora $\boldsymbol{\delta}$ parametrów modelu (3.89) zastosowany zostanie estymator wewnątrzgrupowy, skorygowany

[15] Model (3.89) spełnia oczywiście wszystkie założenia wymienione dla modelu (3.88), gdyż obie postacie są sobie równoważne.

o wartość znajdującą się po prawej stronie wzoru (3.90) bez ostatniego składnika, to obciążenie uzyskanego w ten sposób tzw. estymatora z korektą Kivieta jest rzędu $O(N^{-1}T^{-3/2})$. Ocenę nieznanej wariancji σ_ε^2 wyznaczyć można na podstawie początkowego estymatora zgodnego, np. estymatora Andersona i Hsiao.

Na s. 65 artykułu K i v i e t a [1995] sformułowane jest też twierdzenie 2, w którym podany jest sposób wyznaczania korekty estymatora $\hat{\boldsymbol{\delta}}_{\mathrm{W}}$ dla modelu (3.89) bez założenia stacjonarności zmiennych egzogenicznych. Ze względu na zawiłość zapisu matematycznego nie zostanie ono tu przytoczone.

Obciążenie estymatora $\hat{\boldsymbol{\delta}}_{\mathrm{W}}$, aproksymowane na podstawie obu twierdzeń, zostało następnie wykorzystane przez autora w eksperymentach Monte Carlo. Eksperyment Kivieta jest zaprojektowany podobnie jak eksperyment Monte Carlo A r e l l a n o i B o n d a [1991] z pewnymi różnicami, dotyczącymi założeń odnośnie do wartości początkowej y_{i0}, wariancji σ_ε^2 i σ_α^2 i rzeczywistych wartości parametrów. Różnice te są dokładnie opisane na s. 65–66.

W eksperymencie porównywane są estymatory bazujące na metodzie zmiennych instrumentalnych A n d e r s o n a i H s i a o [1981, 1982] z różnymi zbiorami instrumentów, estymatory A r e l l a n o i B o n d a [1991]: FDGMM1 (jednostopniowy) i FDGMM2 (dwustopniowy), oraz LSDVc – estymator wewnątrzgrupowy z korektą Kivieta, wyznaczoną ze wzoru (3.90). Ten ostatni estymator jest wyznaczany przez odjęcie od oceny uzyskanej na podstawie $\hat{\boldsymbol{\delta}}_{\mathrm{W}}$ oszacowanej na podstawie wzoru (3.90) wartości korekty, z wariancją σ_ε^2 wyznaczoną na podstawie reszt FDGMM1. W tab. 2–4 (K i v i e t [1995], s. 68, 69 i 71) prezentowane są szczegółowe wyniki porównań obciążenia, średniego błędu szacunku i RMSE dla estymatorów parametrów $\boldsymbol{\beta}$ i γ wyznaczonych wymienionymi powyżej metodami, przy założeniach różnych wartości rzeczywistych tych parametrów. W wyniku eksperymentu stwierdzono, że estymatory GMM wektora $\boldsymbol{\beta}$ były w zasadzie nieobciążone, ale dla parametru γ ujemne obciążenie estymatora sięgać może w zależności od scenariusza nawet 50% jego błędu szacunku. Estymatory typu Andersona i Hsiao parametru γ miały w niektórych scenariuszach lepsze własności niż GMM, szczególnie przy dużych wartościach tego parametru; odwrotna sytuacja miała zaś miejsce w przypadku estymatorów wektora $\boldsymbol{\beta}$. Estymator LSDVc natomiast, porównany za pomocą RMSE, może, według Kivieta, konkurować ze wszystkimi pozostałymi branymi od uwagę technikami estymacji.

Alternatywne propozycje procedury korygującej obciążenie estymatora wewnątrzgrupowego zawarte są w opracowaniach B u n a i C a r r e e [2006] oraz E v e r a e r t a i P o z z i e g o [2007]. W pierwszym z nich wyznaczono wielkość asymptotycznego obciążenia estymatora wewnątrzgrupowego dla skończonego T i $N \to \infty$, przy założeniu heteroskedastyczności przekrojowej, i zaproponowano metodę wyznaczania korekty. Drugie z wymienionych opracowań zawiera

metodę wyznaczania korekty na podstawie iteracyjnej procedury bootstrapowej. Wydaje się jednak, że metody te mają obecnie znacznie mniejsze znaczenie praktyczne niż metoda Kivieta.

3.3. Testowanie hipotez

Test autokorelacji Arellano–Bonda

Występowanie autokorelacji pierwszego rzędu w modelu pierwszych różnic jest zjawiskiem spodziewanym, gdyż jeśli ε_{it} są niezależne, to ich pierwsze różnice są skorelowane rzędu 1. Występowanie w tym modelu autokorelacji rzędu wyższego niż 1 oznaczałoby jednak, że warunki momentów są niespełnione, lub inaczej, że instrumenty użyte podczas estymacji GMM nie są właściwe. Arellano i Bond [1991] zaproponowali test, badający występowanie autokorelacji drugiego rzędu składnika losowego w modelu pierwszych różnic, postaci:

$$y_{it} - y_{i,t-1} = \gamma(y_{i,t-1} - y_{i,t-2}) + (\mathbf{x}_{it}^{\mathrm{T}} - \mathbf{x}_{i,t-1}^{\mathrm{T}})\boldsymbol{\beta} + (\varepsilon_{it} - \varepsilon_{i,t-1}) \tag{3.91}$$

W najbardziej zwartym zapisie macierzowym (3.91) przyjmie postać:

$$\Delta\mathbf{y} = \mathbf{W}\boldsymbol{\delta} + \Delta\boldsymbol{\varepsilon} \tag{3.92}$$

gdzie $\Delta\mathbf{y} = [y_{it} - y_{i,t-1}]$, $\Delta\boldsymbol{\varepsilon} = [\varepsilon_{it} - \varepsilon_{i,t-1}]$ są wektorami $((T-2)N \times 1)$, $\mathbf{W}$ jest macierzą obserwacji na zmiennych objaśniających modelu (3.91), $\boldsymbol{\delta}$ – wektorem parametrów strukturalnych tego modelu. Wektor reszt modelu (3.92) wyznaczyć można jako:

$$\Delta\hat{\boldsymbol{\varepsilon}} = \Delta\mathbf{y} - \mathbf{W}\hat{\boldsymbol{\delta}} = \Delta\boldsymbol{\varepsilon} - \mathbf{W}(\hat{\boldsymbol{\delta}}_{\mathrm{GMM}} - \boldsymbol{\delta}) \tag{3.93}$$

gdzie $\hat{\boldsymbol{\delta}}_{\mathrm{GMM}}$ oznaczać może estymator FDGMM lub estymator SGMM modelu (3.92), postaci (3.43).

Zgodność estymatora GMM zależy, jak zostało to stwierdzone uprzednio, od prawdziwości warunku:

$$\mathrm{E}(\Delta\varepsilon_{it}\Delta\varepsilon_{i,t-2}) = 0 \quad \text{dla } t = 3, ..., T \tag{3.94}$$

Zatem testem, weryfikującym poprawność warunków momentów, może być test, badający występowanie autokorelacji drugiego rzędu w modelu (3.92). Weryfikowana jest hipoteza zerowa o niewystępowaniu takiej autokorelacji. Statystyka empiryczna ma postać:

$$m_2 = \frac{\Delta\hat{\boldsymbol{\varepsilon}}_{-2}^{\mathrm{T}}\Delta\hat{\boldsymbol{\varepsilon}}_{*}}{\Delta\hat{\boldsymbol{\varepsilon}}^{1/2}} \tag{3.95}$$

gdzie $\Delta\hat{\boldsymbol{\varepsilon}}_{-2}$ oznacza drugie różnice wektora $\Delta\boldsymbol{\varepsilon}$, zaś elementy wektora $\Delta\hat{\boldsymbol{\varepsilon}}_{*}$ są równe elementom $\Delta\boldsymbol{\varepsilon}$, ale z pominięciem pierwszych dwóch wartości, tak aby wymiary wektorów umożliwiały wykonalność mnożenia. Statystyka m_2 ma rozkład normalny N(0, 1), co zostało wykazane w: A r e l l a n o, B o n d [1991], s. 293–294 (Appendix).

Hipoteza zerowa jest prawdziwa, jeśli składnik losowy modelu na poziomach (3.11) nie wykazuje autokorelacji w czasie, ale również jeśli składnik ten jest procesem błądzenia losowego. Rozróżnienia między tymi dwoma przypadkami można dokonać obliczając, analogicznie do wyliczania m_2, wartość statystyki m_1. W ten sposób zbadane zostanie występowanie autokorelacji pierwszego rzędu w modelu pierwszych różnic. Alternatywnie A r e l l a n o, B o n d [1991] zasugerowali stosowanie testu Hausmana, bazującego na różnicy efektywności między dwoma estymatorami: GMM i MNK modelu pierwszych różnic[16]. Jest to uzasadnione, ponieważ jeśli składnik losowy modelu na poziomach jest procesem błądzenia losowego, to oba te estymatory są zgodne.

Test Sargana (Sargana–Hansena)[17]

Test ten, w kontekście panelowych modeli dynamicznych, zaproponowany został przez A r e l l a n o, B o n d a [1991]. Konstrukcja bazuje na propozycjach S a r g a n a [1958] oraz H a n s e n a [1982]. Test bada prawdziwość warunków ponadidentyfikujących, niewykorzystanych w procesie estymacji. Weryfikowana jest hipoteza zerowa, że wykorzystane instrumenty są właściwe w sensie ich nieskorelowania ze składnikami losowymi modelu pierwszych różnic (3.92). Statystyka empiryczna ma postać:

$$s = \Delta\hat{\boldsymbol{\varepsilon}}^{\mathrm{T}}\mathbf{Z}\left[\sum_{i=1}^{N}\mathbf{Z}_i^{\mathrm{T}}\Delta\hat{\boldsymbol{\varepsilon}}\Delta\hat{\boldsymbol{\varepsilon}}^{\mathrm{T}}\mathbf{Z}_i\right]^{-1}\mathbf{Z}^{\mathrm{T}}\Delta\hat{\boldsymbol{\varepsilon}} \tag{3.96}$$

gdzie $\Delta\hat{\boldsymbol{\varepsilon}}$ jest określone wzorem (3.93), $\mathbf{Z}_i$ jest macierzą zmiennych instrumentalnych, określoną wzorem (3.27) lub (3.39), w zależności od tego, czy model jest autoregresyjny lub czy zawiera niezależne zmienne objaśniające, zaś

[16] Test Hausmana opisany jest w podrozdz. 2.3.

[17] Test Sargana dla modelu szacowanego na podstawie danych czasowych jest omówiony w dodatku II.

$\mathbf{Z} = \left[\mathbf{Z}_1^{\mathrm{T}}, ..., \mathbf{Z}_N^{\mathrm{T}}\right]^{\mathrm{T}}$. Statystyka s ma rozkład χ^2 z q stopniami swobody, gdzie q oznacza liczbę kolumn macierzy **Z**, pomniejszoną o liczbę szacowanych parametrów.

Jak zauważyli A r e l l a n o, B o n d [1991], możliwa jest sytuacja, gdy jeden z opisanych powyżej testów można do przeprowadzić, a drugi jest niewykonalny.

Zastosowanie testu Sargana jest możliwe tylko w przypadku, gdy liczba kolumn macierzy **Z** przewyższa liczbę szacowanych parametrów. Ponadto przeprowadzenie testu Sargana na podstawie reszt estymatora jednostopniowej GMM może nastąpić tylko wtedy, gdy składniki losowe mają jednakowy rozkład i są niezależne dla wszystkich i, t. W innym przypadku zmodyfikowana statystyka s, obliczona na podstawie reszt jednostopniowej GMM, nie ma rozkładu χ^2. Test autokorelacji jest natomiast wykonalny niezależnie od rozkładu składnika losowego, ponieważ rozkład statystyki m_2 pozostaje rozkładem normalnym.

Odwrotna sytuacja, kiedy test autokorelacji nie może być zastosowany, a test Sargana jest możliwy do przeprowadzenia, ma miejsce, jeśli liczba obserwacji w czasie $T < 5$. Dla przykładu, w przypadku modelu autoregresyjnego i $T = 4$ statystyka testu Sargana bada dwie liniowe kombinacje utworzone z trzech dostępnych warunków momentów: $\mathrm{E}(\Delta\varepsilon_{i3}y_{i1}) = \mathrm{E}(\Delta\varepsilon_{i4}y_{i1}) = \mathrm{E}(\Delta\varepsilon_{i4}y_{i2}) = 0$. Nie można natomiast obliczyć, z powodu braku danych, drugich różnic wektora $\Delta\boldsymbol{\varepsilon}$, potrzebnych do obliczenia wartości m_2.

B o w s h e r [2002] zauważył, że zdolność testu Sargana do wykrywania niewłaściwych warunków momentów zmniejsza się drastycznie w małych próbach wraz ze wzrostem liczby dodatkowych warunków. Ponadto w przypadku, gdy wariancja składnika losowego nie jest stała test Sargana jest obciążony na niekorzyść hipotezy zerowej.

Różnicowy test Sargana (ang. Sargan difference test)

Test Sargana, dla którego postać statystyki empirycznej dana jest wzorem (3.96), został zaproponowany do badania prawdziwości warunków ponadidentyfikujących, których spełnienie jest zakładane, jeśli metodą estymacji jest FDGMM A r e l l a n o i B o n d a [1991]. Zastosowane systemowej GMM B l u n d e l l a i B o n d a [1998] wymaga przyjęcia założenia, że spełnione są dodatkowe warunki momentów określone wzorem (3.80). Zbadanie prawdziwości tych dodatkowych warunków jest możliwe poprzez zastosowanie testu, określanego w literaturze anglojęzycznej terminami *differenced-Sargan*, *Sargan difference* lub *incremental Sargan test*[18]. Poniżej test ten nazywany będzie

[18] Możliwość zastosowania tego testu została zauważona przez A r e l l a n o, B o n d a [1991], s. 283. B o n d i W i n d m e i j e r [2005] zaproponowali jego zastosowanie do porównania SGMM i FDGMM. Praca ta zawiera ponadto na s. 10–12 szczegółową analizę własności statystycznych testu.

różnicowym testem Sargana. Jest on skonstruowany jako różnica pomiędzy wartościami statystyk empirycznych testu Sargana dla dwóch modeli: modelu „z ograniczeniami” (ang. *restricted model*) i modelu „bez ograniczeń” (ang. *unrestricted model*). Pod tym pierwszym pojęciem, w kontekście SGMM i FDGMM, rozumie się model, który spełnia założenia Blundella i Bonda [1998], a drugi z nich to model spełniający założenia Arellano i Bonda [1991]. Statystyka empiryczna ma postać:

$$ds = s_{\mathrm{SGMM}} - s_{\mathrm{FDGMM}} \tag{3.97}$$

gdzie s_{SGMM} i s_{FDGMM} oznaczają wartości empiryczne testu Sargana, obliczone na podstawie wzoru (3.96) z wektorem reszt $\Delta\hat{\boldsymbol{\varepsilon}}$ wyznaczonym odpowiednio na podstawie dwustopniowego estymatora SGMM lub FDGMM.

Jeśli hipoteza zerowa jest prawdziwa, a więc wykorzystane instrumenty są właściwe, to statystyka *ds* ma rozkład χ^2 z *r* stopniami swobody, gdzie *r* oznacza różnicę między liczbami stopni swobody dla statystyk s_{SGMM} i s_{FDGMM}. Tak określona liczba stopni swobody jest równa liczbie dodatkowych warunków momentów, spełnionych dzięki uwzględnieniu założeń Blundella i Bonda [1998].

3.4. Porównanie własności estymatorów

Wobec istnienia alternatywnych metod estymacji dynamicznych modeli panelowych oraz propozycji różnorodnych ich modyfikacji, może pojawić się pytanie, czy istnieje jedna, uniwersalna metoda, czy raczej różne metody powinny być stosowane w różnych sytuacjach. Próby odpowiedzi na to pytanie podejmowane były przez wielu autorów. Niejednokrotnie autorzy ci, proponując własny estymator, wskazują sytuacje, w których jego własności są najlepsze. Oprócz tego w literaturze znaleźć można prace, w których porównywane są własności kilku alternatywnych estymatorów. Zaliczyć do nich można np. publikacje: Windmeijera [1998], Judson i Owen [1999], Harrisa i Matyasa [2004], Alvarezа i Arellano [2003].

Wnioski, płynące z porównania własności różnych estymatorów wskazują na potrzebę stosowania różnych metod, w zależności od rodzaju próby. Na przykład Arellano i Bond [1991] wykazali, że proponowany przez nich estymator FDGMM jest bardziej efektywny niż estymator metody zmiennych instrumentalnych Andersona i Hsiao [1981, 1982]. Badali oni własności asymptotyczne estymatorów, przy ustalonej wartości T dla $N \to \infty$. Kiviet [1995], przy nieco zmodyfikowanych założeniach, wykazał przewagę estymatora Andersona–Hsiao nad estymatorem FDGMM.

J u d s o n i O w e n [1999] porównali własności różnych estymatorów dla próby typowej dla danych makroekonomicznych. Założyli oni stacjonarność zmiennych oraz wartości T od 5 do 30 i nielosowość efektów grupowych. Stosując metodę Monte Carlo, porównali własności czterech estymatorów: estymatora metody zmiennych instrumentalnych (IV) Andersona i Hsiao, estymatorów Arellano–Bonda: jednostopniowego (FDGMM1) i dwustopniowego (FDGMM2), i skorygowanego estymatora Kivieta (LSDVc). Dodatkowo przeprowadzono badanie obciążenia estymatora KMNK i estymatora wewnątrzgrupowego (WG). Własności tych estymatorów badane były dla 32 różnych kombinacji parametrów, dla czterech wartości T: 5, 10, 20, 30, wartości N równych 20 lub 100, oceny parametru przy opóźnionej zmiennej endogenicznej γ równej 0,2 lub 0,8 oraz ustalonych pozostałych parametrach. Zasadnicze wnioski płynące z eksperymentu Judson i Owen są następujące:

– obciążenie estymatora WG parametru γ przy opóźnionej zmiennej objaśnianej było zawsze większe niż obciążenie estymatora WG parametru β_k przy dowolnej zmiennej objaśniającej x_k;

– obciążenie estymatora wewnątrzgrupowego rośnie ze wzrostem wartości γ, a maleje ze wzrostem T; jednak nawet przy $T = 30$ obciążenie to może wynosić nawet 20% rzeczywistej wartości parametru;

– niezależnie od wymiaru czasowego próby, pierwiastek błędu średniokwadratowego (RMSE) jest najmniejszy dla estymatora Kivieta. Ponadto oba estymatory FDGMM1 i FDGMM2 mają mniejszy RMSE niż estymator Andersona i Hsiao;

– ze wzrostem N i T różnice między wszystkimi analizowanymi estymatorami zacierają się;

– w przypadku próby zbilansowanej autorzy zalecają stosowanie estymatora Kivieta, a w przypadku próby niezbilansowanej decyzja zależy od wymiaru T: dla $T = 10$ zalecany jest FDGMM1, dla $T = 20$ – FDGMM2 lub estymator Andersona i Hsiao, a dla $T = 30$ – estymator WG, którego własności są wówczas porównywalne z własnościami pozostałych estymatorów, a który jest prostszy do uzyskania.

Eksperymenty Monte Carlo stosowane były też przez H a r r i s a i M a t y a s a [2004] do porównania własności estymatorów GMM i IV Andersona–Hsiao w małych próbach, ze szczególnym uwzględnieniem problemu ich obciążenia. Oprócz najpopularniejszych w praktyce estymatorów jedno- i dwustopniowej FDGMM i SGMM analizowano też własności kilku ich wariantów, uzyskanych poprzez uwzględnianie różnych zbiorów instrumentów. Symulacje przeprowadzono dla $T = 4$ lub $T = 6$ i trzech różnych rzeczywistych wartości parametru autoregresyjnego γ: 0,1, 0,5, 0,9. Podstawowe wnioski są następujące:

– estymatory, które wykorzystują bardzo rozbudowaną macierz instrumentów wykazują szczególnie duże obciążenie w małych próbach. Dotyczy to m. in. estymatora A r e l l a n o i B o v e r [1995]. Wniosek ten jest zbieżny z wynikami

przedstawionymi przez Z i l i a k a [1997]. Wykorzystanie w procesie estymacji wszystkich dostępnych warunków momentów nie jest zatem zalecane, ponieważ możliwa wielkość obciążenia estymatora przeważa oczekiwany wzrost efektywności;

– jeśli liczba instrumentów przekracza liczbę zmiennych endogenicznych modelu, to otrzymany estymator ma skończone momenty. W przeciwnym wypadku nie ma takiej gwarancji;

– błędy szacunku estymatorów dwustopniowych są często niewiarygodne;

– własności wszystkich estymatorów poprawiają się dla $T \to \infty$, ale dotyczy to szczególnie estymatora Andersona–Hsiao, który nie powinien więc być stosowany w małych próbach;

– własności wszystkich estymatorów poprawiają się, gdy wartość parametru autoregresyjnego $\gamma \to 1$;

– za najbardziej uniwersalne estymatory, które generalnie mają najlepsze własności uznać można estymatory GMM, wykorzystujące nieliniowe warunki momentów, tzn. estymatory A h n a i S c h m i d t a [1995] i B l u n d e l l a i B o n d a [1998].

Asymptotyczne własności proponowanych w literaturze estymatorów były przedmiotem badania A l v a r e z a i A r e l l a n o [2003]. Porównano estymatory: WG, FDGMM, oraz estymator metody największej wiarygodności z ograniczoną informacją LIML (A l o n s o-B o r r e g o, A r e l l a n o [1996]) dla modelu autoregresyjnego, przy $T \to \infty$ i $N \to \infty$. Stwierdzono, że wszystkie wymienione estymatory są zgodne, jeśli stosunek $T/N \to c$, gdzie $0 < c \leq 2$. Fakt ten uzasadnić można następująco: w przypadku modeli autoregresyjnych duża liczba dostępnych instrumentów jest związana z tym, że $T \to \infty$; w takim zaś przypadku zbieżność estymatorów innych metod do estymatora KMNK (WG) jest pożądana, gdyż obciążenie wynikające ze współzależności zmiennych dąży do 0, jeśli $T \to \infty$[19]. Mimo to estymatory FDGMM, WG i LIML są asymptotycznie obciążone, a rząd tego obciążenia wynosi odpowiednio N, T i $2N - T$. Jeśli $T < N$, to asymptotyczne obciążenie estymatora FDGMM jest mniejsze niż estymatora WG, zaś asymptotyczne obciążenie estymatora LIML – mniejsze niż dwóch pozostałych. Dla $T = N$ obciążenia wszystkich estymatorów są identyczne.

W praktyce ekonometrycznej najczęściej stosowanymi metodami estymacji panelowych modeli dynamicznych są FDGMM i SGMM. Porównania własności tych estymatorów w konkretnych zastosowaniach pojawiają się w wielu publikacjach. Wydaje się, że wnioski streścić można jak poniżej. W szczególnym przypadku, kiedy macierz wariancji-kowariancji składnika losowego jest sferyczna, estymatory jedno- i dwustopniowej FDGMM są asymptotycznie

[19] Zaznaczyć warto, że w przypadku modeli strukturalnych zbyt duża liczba instrumentów powoduje nadmierne dopasowanie, a zbieżność estymatorów innych metod do estymatora KMNK jest niepożądana (M o r i m u n e [1983] i B e k k e r [1994] wykazali, że estymator 2MNK jest niezgodny, jeśli liczba instrumentów dąży do nieskończoności).

równoważne. W przeciwnym wypadku – dwustopniowy estymator FDGMM jest bardziej efektywny. Estymator dwustopniowy SGMM jest zawsze bardziej efektywny niż jednostopniowy SGMM. Eksperymenty Monte Carlo pokazały jednak, że wzrost efektywności jest relatywnie niewielki, a zbieżność estymatora GMM do rozkładu asymptotycznego – powolna. Ponadto w małych próbach średnie błędy szacunku estymatorów GMM są silnie obciążone w dół. Dlatego bardziej wiarygodne wyniki uzyskuje się stosując jednostopniowy estymator GMM z odpowiednio liczonymi, odpornymi na heteroskedastyczność, błędami szacunku (por. B o n d, H o e f f l e r, T e m p l e [2001]).

ROZDZIAŁ 4

PRZYKŁADY ZASTOSOWAŃ DYNAMICZNYCH MODELI PANELOWYCH W BADANIACH MIKROEKONOMICZNYCH

Początki rozwoju metodologii modelowania na podstawie danych panelowych związane są z wykorzystaniem danych indywidualnych, tzn. danych pochodzących z gospodarstw domowych lub firm (patrz podrozdz. 1.4). Uznawane obecnie za klasyczne metody estymacji modeli z efektami grupowymi (patrz podrozdz. 2.1–2.3) zostały opracowane z myślą o wykorzystaniu danych indywidualnych, w których wymiar przekrojowy jest znacznie większy niż wymiar czasowy. Współcześnie zakres zastosowania danych panelowych bardzo się rozszerzył i obejmuje również dane ze szczebla mezo- i makroekonomicznego, o czym mowa będzie w kolejnych rozdziałach pracy.

W niniejszym rozdziale przedstawiono przykłady badań światowych z obszaru ekonomii przemysłowej i finansów przedsiębiorstw, prowadzonych na podstawie dynamicznych modeli panelowych. Przykłady starano się dobrać tak, aby obrazowały korzyści badawcze płynące z zastosowania mikropaneli dla analizy ważnych aspektów funkcjonowania przedsiębiorstw: inwestycji, produktywności, zatrudnienia, innowacji, czynników finansowych, transferu kwalifikacji czy struktury zarządu firmy. Ponieważ problemy badawcze rozważane w omawianych pracach obejmują różnorodne aspekty funkcjonowania przedsiębiorstw, zostały one zestawione w formie tabelarycznej.

Tabela 4.1

Przykłady analiz działalności przedsiębiorstw na podstawie dynamicznych modeli panelowych

Problem badawczy	Opracowania
Czynniki determinujące decyzje inwestycyjne	**Bond**, **Elton**, **Mairesse**, **Mulkay** [**1997**]; **Bond** [**2002**]; Bo [2007]
Produktywność i rentowność	**Blundell**, **Bond** [**2000**]; Benito, Hernando [2007]; Casu, Girardone [2007]; **Loshkin**, **Belderbos**, **Carree** [**2008**]; Ordine, Rose [2008]
Zatrudnienie	**Arellano**, **Bond** [**1991**]; **Blundell**, **Bond** [**1998**]; Shehzad, De Haan, Scholtens [2009]
Działalność innowacyjna	**Lachenmaier**, **Rottmann** [**2007**]; Brown, Fazzari, Petersen [2009]

Tabela 4.1 (cd.)

Problem badawczy	Opracowania
Wynagrodzenia	**Bover**, **Bentolila**, **Arellano** [**2001**]; **Andini** [**2009**]
Wpływ struktury właścicielskiej i kapitałowej na funkcjonowanie firmy	**Criscuolo** [**2003**]; **Drobetz**, **Wanzenried** [**2006**]; **Guest** [**2009**]
Transfer kwalifikacji i BiR, mobilność pracowników	**Gorg**, **Strobl** [**2002**]; Bloom, Schankerman, Van Reenen [2005]; **Higon** [**2007**]

Uwaga: nazwiska autorów opracowań omawianych w tym rozdziale wyróżniono czcionką półgrubą.

Źródło: opracowanie własne.

W tab. 4.1, oprócz opracowań omawianych w niniejszym rozdziale zawarto odwołania do innych prac, które mogą być inspirujące dla Czytelnika zainteresowanego problematyką ekonomii przemysłowej i funkcjonowania przedsiębiorstw.

4.1. Inwestycje w firmach

Wiele miejsca w empirycznych studiach mikroekonomicznych zajmują analizy roli, jaką w decyzjach inwestycyjnych firm odgrywają makroekonomiczne czynniki finansowe. Większość z tych analiz pozwala na wniosek, że takie zmienne, jak przepływy pieniężne, istotnie wpływają na wydatki inwestycyjne. Przy braku założenia o doskonałym rynku kapitałowym, w którym decyzje finansowe i decyzje inwestycyjne są niezależne, możliwe jest, że różne systemy finansowe na szczeblu krajowym mogą w różny sposób wpływać na inwestycje w firmach. Na przykład dużo miejsca poświęca się w literaturze przedmiotu analizie różnic między anglosaskim systemem opartym na rynku (ang. *market-based*) a niemieckim lub japońskim systemem opartym na bankach (ang. *bank-based*). Celem opracowania Bonda, Elton, Mairesse i Mulkay'a [1997] jest odpowiedź na pytanie, czy zróżnicowanie wpływu wywieranego na decyzje inwestycyjne firm przez ograniczenia finansowe może wynikać z różnic między systemami finansowymi poszczególnych krajów. Autorzy oszacowali trzy alternatywne specyfikacje modelu inwestycji: model akceleracyjny, model korekty błędem i równanie Eulera, uwzględniające dodatkowo zmienne finansowe: przepływy pieniężne i składniki zysku. Specyfikacje tych równań są następujące:

– model akceleracyjny:

$$\frac{I_{it}}{K_{i,t-1}} = \rho \frac{I_{i,t-1}}{K_{i,t-2}} + \beta_0 \Delta y_{it} + \beta_1 \Delta y_{i,t-1} + d_t + \eta_i + \nu_{it} \tag{4.1}$$

gdzie I_{it} oznacza inwestycje w firmie i w roku t, K_{it} – zasób kapitału, d_t – efekt czasowy, η_{it} – efekt grupowy,

– model korekty błędem:

$$\frac{I_{it}}{K_{i,t-1}} = \rho \frac{I_{i,t-1}}{K_{i,t-2}} + \beta_0 \Delta y_{it} + \beta_1 \Delta y_{i,t-1} + \phi(k_{i,t-2} - y_{i,t-2}) + d_t + \eta_i + \nu_{it} \qquad (4.2)$$

gdzie dodatkowo $k_{it} = \ln(K_{it})$,

– równanie Eulera:

$$\left(\frac{I}{K}\right)_{i,t+1} = \beta_1 \left(\frac{I}{K}\right)_{it} - \beta_2 \left(\frac{I}{K}\right)_{it}^2 - \beta_3 \left(\frac{\Pi}{K}\right)_{it} + \beta_4 \left(\frac{Y}{K}\right)_{it} + d_{t+1} + \eta_i + \nu_{i,t+1} \qquad (4.3)$$

gdzie $\Pi_{it} = p_{it} F(K_{it}, L_{it}) - p_{it} G(K_{it}, L_{it}) - w_{it} L_{it}$, p_{it} – cena produkcji, $F(K_{it}, L_{it})$ – funkcja produkcji brutto, $G(K_{it}, L_{it})$ – funkcja kosztów dostosowawczych, L_{it} – czynniki produkcji, w_{it} – cena czynników produkcji.

Autorzy przykładają szczególna wagę do porównania wyników estymacji tego samego modelu dla różnych krajów. Mniejsze znaczenie ma porównanie wyników estymacji różnych modeli dla danego kraju. Estymacji modeli (4.1)–(4.3) dokonano na podstawie danych panelowych z lat 1979–1989 dla 228 firm sektora wytwórczego z obszaru Wielkiej Brytanii, Belgii, Francji i Niemiec. Do estymacji modeli zastosowano FDGMM[1].

Podstawowym wnioskiem, niezależnym od specyfikacji modelu, jest to, że makroekonomiczne zmienne finansowe mają bardziej istotny wpływ (statystycznie i ilościowo) na inwestycje w Wielkiej Brytanii niż w pozostałych trzech krajach. Oznacza to zdaniem autorów, że wpływ ograniczeń finansowych na inwestycje może być relatywnie silniejszy w opartym na rynku systemie brytyjskim. Zmienne finansowe wpływają przy tym na inwestycje we wszystkich firmach, niezależnie od kraju. Autorzy zauważają też, że prosta analiza czasowa procesów inwestycyjnych pokazuje znaczące podobieństwo w ich kształtowaniu się w poszczególnych krajach.

B o n d [2002] prezentuje przykład zastosowania alternatywnych estymatorów do oszacowania parametrów autoregresyjnego modelu stopy inwestycji dla firm brytyjskich. Model stopy inwestycji jest szacowany na podstawie 4966

[1] SGMM zaproponowano w roku 1998, a więc później niż ukazało się omawiane opracowanie.

obserwacji, pochodzących z 703 firm, dla których dostępne były dane roczne za co najmniej 4 kolejne lata w okresie 1988–2000. Model ma postać:

$$\left(\frac{I_{it}}{K_{it}}\right) = c_t + \gamma\left(\frac{I_{i,t-1}}{K_{i,t-1}}\right) + (\alpha_i + \varepsilon_{it}) \tag{4.4}$$

gdzie I_{it} oznacza wydatki inwestycyjne brutto, K_{it} – kapitał akcyjny netto i koszt odtworzenia. Stosunek I_{it}/K_{it} jest zatem przybliżoną miarą stopy wzrostu kapitału akcyjnego netto plus stopy amortyzacji. Zmienność w czasie wyrazu wolnego c_t odzwierciedlać ma jednakowe dla wszystkich firm składowe stopy inwestycji, wynikające z cykliczności lub tendencji rozwojowej.

Do oszacowania modelu zastosowane zostały cztery alternatywne estymatory: KMNK, wewnątrzgrupowy (WG), 2MNK dla modelu pierwszych różnic i FDGMM. Wyniki prezentowane są w pracy B o n d a [2000] (tab. 1, s. 31). Oceny parametru γ wahają się od –0,01 dla estymatora WG, do 0,27, dla estymatora KMNK. Jest to duża rozbieżność, która potwierdza regułę, że w przypadku panelowych modeli dynamicznych estymator KMNK jest obciążony w górę, a WG – w dół.

Ocena parametru autoregresyjnego uzyskana na podstawie 2MNK wynosi 0,1626. Estymator 2MNK dla modelu pierwszych różnic jest wyznaczony przy zastosowaniu zmiennej instrumentalnej $(I/K)_{i,t-2}$. Ponieważ model jest jednoznacznie identyfikowalny, nie można zastosować testu Sargana, zatem oceny jego jakości dokonano na podstawie testu autokorelacji Arellano–Bonda[2]. Przy standardowym założeniu, że składniki losowe ε_{it} nie wykazują korelacji w czasie, w prawidłowo wyspecyfikowanym modelu należy się spodziewać, że przyrosty $\Delta\varepsilon_{it}$ będą wykazywały istotną ujemną korelację pierwszego rzędu i brak istotnej korelacji drugiego rzędu. Wynik testu Arellano–Bonda potwierdza prawidłową specyfikację modelu.

Kolejną zastosowaną metodą estymacji modelu (4.4) jest jednostopniowa FDGMM. Autor prezentuje wyniki uzyskane dla dwóch różnych zbiorów instrumentów: zbioru ograniczonego do $(I/K)_{i,t-2}$ i $(I/K)_{i,t-3}$, oraz pełnego zbioru zmiennych instrumentalnych, wynikającego z założeń A r e l l a n o i B o n d a [1991]. Macierz $\mathbf{Z}_i$ wykorzystana w tym drugim przypadku liczy aż 78 kolumn. Powodem zastosowania FDGMM z ograniczonym zbiorem instrumentów (oprócz FDGMM z pełnym zbiorem instrumentów) jest to, że estymatory GMM wyznaczane na podstawie zbyt wielu warunków momentów mogą być obciążone w małych próbach, a przyczyną tego obciążenia (ang. *overfitting bias*) może być właśnie nadmiar wykorzystanych warunków[3]. Porównanie przedstawionych

[2] Konstrukcje testów Sargana i testu Arellano–Bonda są omówione w podrozdz. 3.3.

[3] Więcej na temat wpływu nadmiernej ilości warunków momentów na obciążenie estymatorów znaleźć można w rozdz. 3.

poniżej wyników estymacji pokazuje jednak, że w omawianym modelu ten problem nie wystąpił. Z drugiej jednak strony utrata efektywności spowodowana przez pominięcie zmiennych instrumentalnych, którymi są bardziej odległe w czasie opóźnienia, okazała się bardzo niewielka.

Mimo że FDGMM z ograniczonym zbiorem instrumentów nie wykorzystuje pełnej macierzy zmiennych instrumentalnych $\mathbf{Z}_i$ danej wzorem (3.27), zauważalna jest pewna poprawa precyzji estymacji w stosunku do 2MNK (błąd szacunku spada z 0,0362 do 0,0327, przy ocenie parametru γ równej 0,1593). Macierz $\mathbf{Z}_i$ wykorzystana w 2MNK składa się z jednej kolumny postaci $[(I/K)_{i1}\ (I/K)_{i2}\ \ldots\ (I/K)_{i,T-2}]^{\mathrm{T}}$, a FDGMM, nawet w ograniczonej wersji, wykorzystuje macierz liczącą 23 kolumny. Wzrost efektywności wynikać może więc z wykorzystania dodatkowych warunków momentów. Model jest obecnie identyfikowalny z nadmiarem, zatem do oceny jego jakości oprócz testu Arellano–Bonda zastosowano test Sargana. Wyniki obu testów pokazały, że warunki ponadidentyfikujące są właściwe, a więc model jest poprawny.

W kolejnym kroku model (4.4) oszacowano jednostopniową FDGMM z pełnym zbiorem zmiennych instrumentalnych. Wykorzystanie dodatkowych warunków momentów przyniosło dalszą poprawę precyzji estymacji. Ocena parametru γ wynosi obecnie 0,156, a błąd jego szacunku 0,0318. Testy Sargana i Arellano–Bonda potwierdziły hipotezę o poprawności modelu.

Dodatkowo model (4.4) estymowano za pomocą dwustopniowej FDGMM z pełnym zbiorem instrumentów. Wyznaczona za pomocą tej metody ocena parametru γ wynosi 0,1806. Asymptotyczny błąd szacunku obliczony według wzoru (3.35) ma wartość 0,0227, a więc jest o ok. 30% mniejszy niż błąd estymatora jednostopniowego, przy zbliżonych wartościach ocen parametru. Autor stwierdza, że tak duży wzrost efektywności estymacji w wyniku zastosowania estymatora dwustopniowego jest nieuzasadniony i nigdy nie był odnotowany w symulacjach. Błąd obliczony przy zastosowaniu korekty Windmeijera na podstawie wzoru (3.46) ma wartość 0,0314. Jest więc praktycznie taki sam, jak błąd estymatora jednostopniowego z pełnym zbiorem instrumentów. Rezultat ten potwierdza udowodnioną w literaturze tezę (por. np. B o n d, W i n d m e i j e r [2005]), że praktyczne wnioskowanie na asymptotycznej macierzy wariancji-kowariancji dwustopniowego estymatora GMM jest ryzykowne.

4.2. Produkcja i produktywność w firmach

Istotnym czynnikiem wpływającym na produktywność w firmach jest wykorzystanie nowych technologii i innowacji. W dobie łatwego transferu wiedzy i technologii istotnego znaczenia nabiera określenie, które technologie będą rozwijane lokalnie, a które kupowane z zewnątrz. Część autorów zauważa

rosnącą tendencję do czerpania technologii z zewnątrz, przy dodatkowym zmniejszaniu się liczby wewnętrznych komórek badawczo-rozwojowych, szczególnie w branżach silnie zależnych od nakładów na BiR. Celem pracy Loshkina, Belderbosa i Carree [2008] jest zbadanie łącznego wpływu zewnętrznych i wewnętrznych nakładów na BiR na produktywność w innowacyjnych firmach w Holandii. Jest to jedna z pierwszych prac z tej dziedziny wykorzystujących dane panelowe na szczeblu firm. Ten rodzaj danych umożliwia, w przeciwieństwie do danych przekrojowych, rozdzielenie efektów komplementarności od stałych lub zmiennych w czasie skutków heterogeniczności obiektów (firm). Analiza bazuje na rozszerzonej funkcji Cobba–Douglasa, w której oprócz nakładów pracy *L* i kapitału *C* uwzględniono wpływ zasobu wiedzy *K* na produkcję *Y*. Ostateczna postać modelu jest następująca:

$$\begin{aligned} q_{it} &= (1+\theta)q_{i,t-1} + (\beta-1)\Delta l_{it} + \delta\Delta i_{it} + \\ &+ \varphi\left[\eta_1 r_{i,t-1}^{\text{int}} + \eta_2 r_{i,t-1}^{\text{ext}} + \eta_3 (r_{i,t-1}^{\text{int}})^2 + \eta_4 (r_{i,t-1}^{\text{ext}})^2 + \eta_5 r_{i,t-1}^{\text{int}} r_{i,t-1}^{\text{ext}}\right] + \lambda_t + \mu_i + \nu_{it} \end{aligned} \tag{4.5}$$

gdzie $q_{it} = \log Y_{it} - \log L_{it}$, $l_{it} = \log L_{it}$, $i_{it} = \log I_{it}$, I_{it} – inwestycje kapitałowe w środki trwałe, będące aproksymantą inwestycji kapitałowych C_{it}, μ_i i λ_t są efektami grupowymi i czasowymi zaś, $r_{i,t-1}^{\text{int}}$ i $r_{i,t-1}^{\text{ext}}$ określone są związkiem:

$$\frac{\Delta K_{it}}{Y_{i,t-1}} = f\left(\frac{R_{i,t-1}^{\text{int}}}{Y_{i,t-1}}, \frac{R_{i,t-1}^{\text{ext}}}{Y_{i,t-1}}\right) = f\left(r_{i,t-1}^{\text{int}}, r_{i,t-1}^{\text{ext}}\right) \tag{4.6}$$

gdzie $R_{i,t-1}^{\text{int}}$ i $R_{i,t-1}^{\text{ext}}$ są odpowiednio zewnętrznymi i wewnętrznymi nakładami na BiR.

Model (4.5) estymowany był za pomocą FDGMM i SGMM oraz za pomocą estymatorów Hsiao, Pesarana i Tahmiscioglu [2002], bazujących na MNW. Estymacji dokonano na podstawie panelu składającego się z 304 firm wytwórczych z lat 1996–2001, należących do 12 działów. Rezultaty estymacji modeli typu (4.5) były jednakowe, niezależnie od metody estymacji. Wykazały one, że połączenie nakładów zewnętrznych i wewnętrznych na BIR znacząco przyczynia się do wzrostu produktywności. Potwierdza to słuszność łączenia tych dwóch źródeł finansowania innowacji i dostarcza dowodów komplementarności strategii innowacyjnych. Ponadto pozytywny wpływ zewnętrznych nakładów na BiR występuje tylko w przypadku ponoszenia wystarczająco wysokich nakładów wewnętrznych.

W pracy B l u n d e l l a i B o n d a [2000] dynamiczny model panelowy zastosowano do estymacji modelu kształtowania się produkcji w firmach. Przykład ten pokazuje, dlaczego specyfikacja dynamiczna może być konieczna dla zapewnienia identyfikowalności pewnych parametrów, mimo że dynamika jako taka nie jest centralnym punktem analizy. Punktem wyjścia jest funkcja produkcji Cobba–Douglasa postaci

$$\begin{aligned} &y_{it} = \beta_n n_{it} + \beta_k k_{it} + \lambda_t + (\alpha_i + \varepsilon_{it} + m_{it}) \\ &\varepsilon_{it} = \gamma \varepsilon_{i,t-1} + e_{it}, \ |\gamma| < 1 \\ &e_{it}, \ \ m_{it} \sim \mathrm{MA}(0) \end{aligned} \tag{4.7}$$

gdzie y_{it}, n_{it}, k_{it} oznaczają kolejno logarytmy: sprzedaży, zatrudnienia i kapitału akcyjnego firmy i w roku t, λ_t jest specyficznym dla danego roku wyrazem wolnym (jednakowym dla wszystkich firm), odzwierciedlającym postęp technologiczny, β_n, β_k, γ są parametrami. Składowe składnika losowego to: efekty grupowe α_i, potencjalnie autoregresyjny „właściwy” składnik losowy ε_{it}, oraz składowa m_{it} odzwierciedlająca błędy pomiaru[4], o której zakłada się, że nie wykazuje autokorelacji. Zakłada się, że zatrudnienie i kapitał mogą być skorelowane zarówno z efektami grupowymi α_i, jak i zakłóceniami e_{it} i błędami pomiaru m_{it}. Przy tych założeniach nie można wskazać właściwych warunków momentów dla modelu statycznego postaci (4.7), o ile składnik losowy ε_{it} rzeczywiście jest autoregresyjny, tzn. o ile $\gamma \neq 0$. Model (4.7) zapisać można jednak równoważnie w postaci:

$$\begin{aligned} y_{it} = \beta_n n_{it} - \gamma \beta_n n_{i,t-1} + \beta_k k_{it} - \gamma \beta_k k_{i,t-1} + \gamma y_{i,t-1} + \\ + (\lambda_t - \gamma \lambda_{t-1}) + (\alpha_i (1-\gamma) + e_{it} + m_{it} - \gamma m_{i,t-1}) \end{aligned} \tag{4.8}$$

lub w postaci:

$$y_{it} = \pi_1 n_{it} + \pi_2 n_{i,t-1} + \pi_3 k_{it} + \pi_4 k_{i,t-1} + \pi_5 y_{i,t-1} + \lambda_t^* + (\alpha_i^* + w_{it}) \tag{4.9}$$

z dwoma nieliniowymi restrykcjami wspólnego czynnika: $\pi_2 = -\pi_1 \pi_5$ i $\pi_4 = -\pi_3 \pi_5$. Ponieważ składnik losowy $(\varepsilon_{it} + m_{it})$ modelu (4.7) wykazuje autokorelację, to składnik losowy w_{it} modelu (4.9) może: a) nie wykazywać autokorelacji w przypadku, gdy nie występują błędy pomiaru, tzn. $w_{it} = e_{it}$ oraz $D^2(m_{it}) = 0$ lub b) $w_{it} \sim \mathrm{MA}(1)$ w przypadku, gdy w niektórych szeregach

[4] Na temat możliwości redukcji skutków błędu pomiaru zmiennych przy zastosowaniu dynamicznych modeli panelowych czytaj w podrozdz. 4.8.

występują niesystematyczne (ang. *transient*) błędy pomiaru. W obu przypadkach możliwe jest uzyskanie zgodnych estymatorów wektora parametrów $\boldsymbol{\pi} = [\pi_1, ..., \pi_5]$ za pomocą GMM. Po wyznaczeniu zgodnych estymatorów wektora $\boldsymbol{\pi}$ i wariancji $D^2(\boldsymbol{\pi})$, możliwe jest nałożenie i testowanie restrykcji wspólnego czynnika, postulowanych w modelu (4.9) i wyznaczenie zgodnych estymatorów parametrów β_n, β_k, γ.

B l u n d e l l i B o n d [2000] oszacowali parametry modelu (4.7) na podstawie zrównoważonego panelu danych z 509 firm prowadzących działalność badawczo-rozwojową w USA w latach 1982–1989. Kapitał akcyjny i zatrudnienie były mierzone na koniec roku rozrachunkowego w danej firmie.

Szeregi obserwacji na zmiennych y_{it}, n_{it}, k_{it} charakteryzowały się wysokim stopniem trwałości, co zostało wykazane poprzez konstrukcję i estymację modeli AR(1) dla tych zmiennych. Oceny MNK parametru autoregresyjnego dla wszystkich szeregów wyniosły 0,99. Oceny wyznaczone na podstawie SGMM są niewiele niższe – od 0,92 do 0,96. Zaskakujące zdaniem autorów jest to, że ocena FDGMM parametru przy zmiennej n_{it} jest niemal identyczna (0,920), jak ocena SGMM (0,923). Dla pozostałych dwóch zmiennych potwierdzone zostało przypuszczenie, że w związku z ich wysokim stopniem trwałości oceny FDGMM są obciążone w dół – wynoszą one 0,768 dla zmiennej k_{it} i 0,775 – y_{it}.

Autorzy zaprezentowali wyniki estymacji modeli (4.7) i (4.9) uzyskane jednostopniowymi FDGMM i SGMM. FDGMM stosowano dwukrotnie: z opóźnionymi o dwa okresy poziomami jako zmiennymi instrumentalnymi, oraz powtórnie, z instrumentami, którymi są poziomy opóźnione o trzy okresy. Właściwość instrumentów w pierwszym przypadku została odrzucona przez test Sargana. Odrzucenie to wynika z występowania błędów pomiaru m_{it}. Instrumenty w drugim przypadku są właściwe, a restrykcje wspólnego czynnika spełnione. Jednakże oceny wszystkich parametrów (w tym parametru γ przy opóźnionej zmiennej objaśnianej) są zbliżone do ocen estymatora wewnątrzgrupowego, co sugeruje możliwość wystąpienia obciążenia wynikającego z małej próby i problemu słabych instrumentów. Kolejną metodą estymacji była SGMM. Właściwość poziomów opóźnionych o trzy okresy, jako instrumentów dla równań na przyrostach, i przyrostów opóźnionych o dwa okresy, jako instrumentów dla równań na poziomach, jest na granicy istotności w teście Sargana. Jednakże wynik różnicowego testu Sargana, badającego poprawność dodatkowych warunków momentów, uzyskanych dzięki wprowadzeniu instrumentów dla równań na poziomach, potwierdza właściwość tych instrumentów na poziomie istotności 0,1. Ponadto o poprawności SGMM świadczy fakt, że ocena parametru γ uzyskana tą metodą jest wyższa niż ocena FDGMM i niższa niż ocena KMNK. Restrykcje wspólnego czynnika są również spełnione. Dlatego SGMM uznano za najwłaściwszą metodę estymacji rozważanego modelu.

4.3. Modelowanie zatrudnienia w przedsiębiorstwach

Prace Arellano i Bonda [1991] oraz Blundella i Bonda [1998] oprócz propozycji nowych metod estymacji zawierają przykłady praktycznego zastosowania FDGMM i SGMM. W obu przykładach modelowane jest zatrudnienie, a zbiór danych jest ten sam. Przykłady te zostaną omówione poniżej, ponieważ pokazują ewolucję wyników w zależności od metody estymacji.

Model Arellano i Bonda [1991] estymowano na podstawie danych ze 140 firm brytyjskich, dla których dostępne były dane z co najmniej 7 kolejnych lat z okresu 1976–1984. Panel jest niezbilansowany w dwojakim sensie: liczba obserwacji w czasie jest różna dla różnych firm oraz obserwacje pochodzą z różnych okresów. Firmy, należące do działu „przemysł" (ang. *manufacturing*), zaklasyfikowano dodatkowo do jednej z dziewięciu głównych gałęzi przemysłu. Ogólna postać modelu jest następująca:

$$\begin{aligned} n_{it} &= \gamma_1 n_{i,t-1} + \gamma_2 n_{i,t-2} + \beta_1 w_{it} + \beta_2 w_{i,t-1} + \theta_1 k_{it} + \theta_2 k_{i,t-1} + \theta_3 k_{i,t-2} + \\ &+ \delta_1 ys_{it} + \delta_2 ys_{i,t-1} + \delta_3 ys_{i,t-2} + \alpha_i + \lambda_t + \varepsilon_{it} \end{aligned} \tag{4.10}$$

gdzie n_{it} oznacza logarytm zatrudnienia w firmie i na koniec roku t, w_{it} – logarytm przeciętnego wynagrodzenia (urealnionego za pomocą deflatora wartości dodanej dla przemysłu), k_{it} – logarytm nakładów kapitałowych brutto, ys_{it} – logarytm produkcji gałęzi przemysłu, do której należy dana firma, mierzony jako wartość dodana wytworzona w tej gałęzi, α_i – efekty grupowe, λ_t – efekty czasowe. Założeniem ekonomicznym, leżącym u podstaw konstrukcji modelu (4.10) jest to, że zatrudnienie kształtuje się zgodnie z log-liniową funkcją popytu na pracę, przy czym w krótkim okresie różni się ono od modelowego. Zmienne k_{it} i w_{it} oraz ich opóźnienia są uwzględnione w celu modelowania krótkookresowych mechanizmów dostosowawczych zatrudnienia. Zmienna ys_{it} wprowadzona jest w celu uwzględnienia zmian popytu danej gałęzi, a zagregowane zmiany popytu są uwzględnione poprzez wprowadzenie efektów czasowych.

Model (4.10) oszacowano za pomocą jedno- i dwustopniowej FDGMM, przy założeniu, że wszystkie zmienne objaśniające, oprócz opóźnień zmiennej objaśnianej, są ściśle egzogeniczne[5]. Oceny parametrów uzyskane metodą jedno- i dwustopniową różnią się nieznacznie. Błędy szacunku parametrów dla metody dwustopniowej, obliczone według wzoru (3.35), są jednak ok. 30% mniejsze niż dla metody jednostopniowej. Autorzy przypuszczają, że ten wzrost efektywności może być pozorny, ponieważ w rzeczywistości może on świadczyć o obciążeniu

[5] Szczegółowe wyniki estymacji omawianych poniżej modeli są zawarte w tekście oryginalnym w tab. 4 na s. 290.

w dół estymatorów wariancji estymatorów dwustopniowej FDGMM, właściwym dla małej próby. Wyniki testów statystycznych są niejednoznaczne. Za pomocą testu autokorelacji Arellano–Bonda stwierdzono poprawność warunków momentów wykorzystanych w obu metodach. Taki sam wynik dał test Sargana dla metody dwustopniowej. Test Sargana dla metody jednostopniowej i różnicowy test Sargana odrzuciły hipotezę zerową o poprawności warunków momentów. Według autorów, przyczyny takiego rozstrzygnięcia mogą być dwie. Jedną z nich jest stwierdzona metodą symulacji generalna tendencja testu Sargana do błędnego odrzucania H_0 w przypadku występowania heteroskedastyczności, a drugą – niepoprawność założenia o ścisłej egzogeniczności wynagrodzeń i kapitału.

Kolejnym etapem analizy były próby modyfikacji modelu (4.10). Pierwszą z nich było pominiecie opóźnień czasowych, które w modelu wyjściowym były nieistotne, tzn. $k_{i,t-1}$, $k_{i,t-2}$, $ys_{i,t-2}$. Tak zmodyfikowany model przyjmuje postać:

$$n_{it} = \gamma_1 n_{i,t-1} + \gamma_2 n_{i,t-2} + \beta_1 w_{it} + \beta_2 w_{i,t-1} + \theta_1 k_{it} + \delta_1 ys_{it} + \delta_2 ys_{i,t-1} + \alpha_i + \lambda_t + \varepsilon_{it} \tag{4.11}$$

Nie uzyskano jednak w ten sposób znaczących zmian wyników. Następnym krokiem było potraktowanie wynagrodzeń i kapitału w modelu (4.11) jako zmiennych endogenicznych. Jako instrumenty dla zmiennych w_{it}, $w_{i,t-1}$ i k_{it} zastosowano opóźnienia zmiennych w i k z okresu $t - 2$ i wcześniejsze. Jako dodatkowe instrumenty wykorzystano opóźnienia realnych wartości produkcji sprzedanej w danej firmie i zapasów. Postać macierzy zmiennych instrumentalnych jest wskazana przez autorów w przypisach do tab. 4 na s. 290. Wyniki estymacji i zastosowanych testów statystycznych wskazują jednoznacznie, że traktowanie wynagrodzeń i kapitału jako zmiennych egzogenicznych było nieuzasadnione. Zarówno test Arellano–Bonda, jak i testy Sargana stwierdzają poprawność warunków momentów wykorzystanych do estymacji modelu (4.11).

A r e l l a n o i B o n d [1991] podjęli dodatkowo próbę oszacowania modelu (4.10) za pomocą metody zmiennych instrumentalnych Andersona i Hsiao oraz estymatora wewnątrzgrupowego (WG). Wyniki zawarte są w tekście oryginalnym w tab. 5 na s. 292. Estymatory Andersona i Hsiao okazały się zdecydowanie mniej efektywne niż estymatory FDGMM. Oceny parametrów uzyskane na podstawie WG były zbliżone do ocen FDGMM, choć ocena WG parametru autoregresyjnego była nieco wyższa niż FDGMM. Należy jednak zauważyć, że metoda WG nie daje możliwości uwzględnienia endogeniczności zmiennych objaśniających, a ponadto w świetle późniejszych prac wartości bliskie ocenom WG świadczyć mogą o obciążeniu estymatora.

Na zakończenie Arellano i Bond podkreślają, że FDGMM jest znacznie efektywniejszą metodą estymacji panelowych modeli dynamicznych niż prostsze

metody zmiennych instrumentalnych (w szczególności metoda Andersona–Hsiao). Zwracają jednak uwagę na tendencję testu Sargana do błędnego odrzucania H_0, w przypadku występowania heteroskedastyczności, oraz na problem obciążenia w dół estymatora asymptotycznej wariancji dwustopniowego estymatora FDGMM. Korekta, która powinna być uwzględniona dla małych prób przy obliczaniu estymowanej asymptotycznej wariancji dwustopniowego estymatora GMM, została zaproponowana w 2005 r. przez F. Windmeijera. Sposób jej wyznaczenia jest dany wzorem (3.46) w pkt 3.2.3. Jednym z przykładów zawartych w pracy Windmeijera [2005] jest próba zastosowania korekty (3.46) do oszacowania wariancji estymatora GMM modelu (4.11). Praca Windmeijera [2005] (tab. 2, s. 42) zawiera wyniki estymacji modelu (4.11) za pomocą jedno- i dwustopniowego estymatora FDGMM. Dla metody dwustopniowej obliczone są błędy szacunku parametrów: standardowe – asymptotyczne, dane wzorem (3.35), oraz na podstawie korekty Windmeijera (3.46). Błędy (3.35) dla estymatora dwustopniowego są znacznie mniejsze niż błędy estymatora jednostopniowego, podczas gdy błędy obliczone na podstawie (3.46) są im w zasadzie równe, z nieznacznymi odchyleniami w górę lub w dół. Windmeijer [2005], podobnie jak Arellano i Bond [1991], przypisuje pozorny wzrost efektywności estymacji dwustopniowej FDGMM temu, że estymatory (3.35) są obciążone w dół w małych próbach.

Blundell i Bond [1998] zilustrowali korzyści z zastosowania SGMM przykładem modelu zatrudnienia szacowanego na podstawie tych samych danych, które Arellano i Bond [1991] wykorzystali do estymacji FDGMM. Postać modelu, będącego dynamicznym równaniem popytu na pracę, jest następująca (oznaczenia zmiennych, takie jak w modelu (4.10)):

$$n_{it} = \gamma n_{i,t-1} + \beta_1 w_{it} + \beta_2 w_{i,t-1} + \theta_1 k_{it} + \theta_2 k_{i,t-1} + u_{it}, \;\; u_{it} = \alpha_i + \varepsilon_{it} \qquad (4.12)$$

Model (4.12) jest przez autorów interpretowany ekonomicznie jako statyczny model popytu na pracę typu Cobba–Douglasa, który został zdynamizowany dla uwzględnienia autoregresji w składniku losowym. Ocena parametru autoregresyjnego może być interpretowana jako ocena tempa procesów dostosowawczych. Postać możliwych do wykorzystania warunków momentów zależy od założeń, dotyczących korelacji między x_{it} i u_{it}. Ponieważ, podobnie jak Arellano i Bond [1991], autorzy traktują wynagrodzenia i kapitał jako zmienne endogeniczne, to warunki momentów wykorzystane w SGMM mają postać (3.40) i (3.87). Warunek (3.80), równoważny założeniu, dotyczącemu warunków początkowych jest w rozważanym tu przypadku spełniony, gdyż, jak stwierdzają Blundell i Bond [1998], jedynym powodem, dla którego rok 1976 jest pierwszym okresem próby, jest brak wcześniejszych danych statystycznych. Nie ma zatem żadnych powodów, by obserwację początkową traktować w szczegól-

ny sposób: „there is nothing special about the first observation on the firms in this sample” (B l u n d e l l, B o n d [1998], s. 136).

Model (4.12) oszacowano na podstawie dwóch prób. Dłuższa obejmuje okres 1976–1984 i jest identyczna z próbą wykorzystaną przez A r e l l a n o i B o n d a [1991], a krótsza obejmuje okres od 1979–1984 i jest niezbilansowanym panelem 4 do 6 obserwacji dla 140 firm. Model jest szacowany za pomocą FDGMM i SGMM, przy czym są to metody jednostopniowe, uznane za bardziej wiarygodne ze względu na obciążenie ocen wariancji estymatorów właściwe dla metod dwustopniowych (korekta Widmeijera opracowana została dopiero w późniejszych latach). W przypadku krótszej próby ocena FDGMM parametru autoregresyjnego wynosi 0,4, a elastyczność zatrudnienia względem kapitału równa jest 0,2, co sugeruje zbyt niskie korzyści skali. Ocena SGMM parametru autoregresyjnego wynosi 0,86, a elastyczność zatrudnienia względem kapitału – 0,6, co jest bardziej sensowne. Autorzy zwracają ponadto uwagę na duży wzrost efektywności estymacji w wyniku zastosowania SGMM. Właściwość dodatkowych warunków momentów zbadano testem Sargana i testem Arellano–Bonda. W przypadku dłuższej próby wyniki FDGMM i SGMM są bardziej zbliżone – elastyczność zatrudnienia względem kapitału wynosi ok. 0,8 dla obu estymatorów, jednak SGMM pozostaje bardziej efektywny. Blundell i Bond konkludują, że dla krótkich prób, szczególnie jeśli szeregi charakteryzują się wysokim stopniem trwałości, estymatory FDGMM są obciążone, a wykorzystanie w SGMM dodatkowych warunków momentów daje możliwość redukcji tego obciążenia i zwiększenia precyzji estymatorów.

4.4. Innowacje a zatrudnienie w przedsiębiorstwach

W literaturze teoretycznej istnieją rozbieżności dotyczące kierunku wpływu innowacji na zatrudnienie w firmach. Istotne dla zrozumienia źródeł tych rozbieżności jest rozróżnienie między innowacjami produktowymi a innowacjami procesowymi. Innowacje produktowe kreują popyt (na nowe produkty). Ten wzrost popytu pozwala zaś firmom zwiększyć zatrudnienie. Z drugiej strony efektem pośrednim wprowadzenia nowego produktu jest osiągniecie przez firmę pozycji monopolisty na rynku (przynajmniej w początkowym okresie) i wykorzystywanie tej pozycji do maksymalizacji zysku m. in. poprzez redukcję zatrudnienia. W przypadku innowacji procesowych efektem bezpośrednim ich wprowadzenia jest spadek zatrudnienia przy tym samym poziomie produkcji. Wynika to bezpośrednio z definicji innowacji procesowej jako ulepszenia procesu produkcyjnego prowadzącego do wzrostu produktywności czynników, m. in. pracy. Jeśli jednak firma przełoży wspomniany wzrost produktywności na spadek cen produktów, to popyt na nie może wzrosnąć.

W odpowiedzi wzrośnie produkcja, co może spowodować zatrudnienie nowych pracowników.

Większość prac empirycznych badających wpływ innowacji na zatrudnienie dowodzi pozytywnego wpływu innowacji produktowych, natomiast w ocenie kierunku wpływu innowacji procesowych na zatrudnienie brak jest jednoznaczności. W pracy Lachenmaiera i Rottmanna [2007] badano wpływ innowacji na zatrudnienie na podstawie danych z 1073 firm niemieckiego przemysłu wytwórczego w latach 1983–2002. Konstrukcja i estymacja dynamicznego modelu panelowego na podstawie danych charakteryzujących się tak długim (dwudziestoletnim) wymiarem czasowym umożliwia kontrolowanie nieobserwowalnej heterogeniczności obiektów, możliwej endogeniczności innowacji oraz dynamiki procesu dostosowania zatrudnienia. Dodatkowo autorzy dysponowali bardzo szczegółowymi danymi dotyczącymi innowacji. Dane te dotyczyły nakładów na innowacje, wyników (innowacji wprowadzonych) rozdzielanych następnie na innowacje procesowe i produkcyjne, a także dawały możliwość określenia „poziomu znaczenia" (wagi) innowacji. Wynik innowacji mierzono poprzez określenie, czy została ona wprowadzona do produkcji, a poziom znaczenia innowacji poprzez określenie czy wymagała ona zaangażowania nakładów na BiR lub czy złożono dla niej wniosek patentowy. Model estymowany przez Lachenmaiera i Rottmanna [2007] ma postać:

$$n_{it} = \beta_1 n_{i,t-1} + \beta_2 n_{i,t-2} + \beta_3 I_{it}^{\mathrm{Pd}} + \beta_4 I_{i,t-1}^{\mathrm{Pd}} + \beta_5 I_{i,t-2}^{\mathrm{Pd}} + \beta_6 I_{it}^{\mathrm{Pc}} + \beta_7 I_{i,t-1}^{\mathrm{Pc}} + \beta_8 I_{i,t-2}^{\mathrm{Pc}} + \\ + \beta_9 w_{it} + \beta_{10} d_{it} + \gamma_i + \varepsilon_{it} \tag{4.13}$$

gdzie n_{it} oznacza logarytm zatrudnienia, I_{it}^{Pd} – innowacje produktowe, I_{it}^{Pc} – innowacje procesowe, w_{it} – stawkę godzinową, d_{it} – wartość dodaną brutto w przemyśle i w roku t; γ_i są efektami grupowymi. Zmienna d_{it} jest wprowadzona jako aproksymanta popytu na wyroby danego przemysłu.

Model (4.13) estymowano za pomocą SGMM z korektą Windmeijera. Innowacje produktowe i procesowe potraktowano jako zmienne endogeniczne, co zostało uzasadnione wynikiem testu Sargana. Autorzy zaprezentowali wyniki estymacji kilku alternatywnych specyfikacji modelu podstawowego postaci (4.13). Najważniejsze wyniki przedstawione są poniżej.

Słaby dodatni wpływ innowacji produktowych na zatrudnienie jest istotny, ale dotyczy to tylko zmiennej opóźnionej o dwa okresy. Wniosek taki, sprzeczny z teorią autorzy uzasadniają naturą danych, obejmujących również bardzo niewielkie zmiany innowacyjne. Innowacje procesowe wpływają natomiast na zatrudnienie silnie dodatnio.

Kolejny etap badania polegał na zastąpieniu innowacji ogółem (zmiennych I_{it}^{Pd} i I_{it}^{Pc}) przez inne miary innowacji: 1) innowacje, które wymagają zaangażowania sektora BiR, oraz 2) innowacje, dla których złożono wnioski patentowe. Wprowadzenie tych miar stanowi próbę zbadania wpływu wywieranego na zatrudnienie przez innowacje o większej wadze. W przypadku innowacji typu 1) wnioski są takie same, jak dla innowacji ogółem. W przypadku 2) innowacje produktowe wywierają silny, dodatni wpływ na zatrudnienie, przy czym dotyczy to wartości bieżącej i obu opóźnień. Innowacje procesowe są natomiast zmienną nieistotną.

Ostatni etap badania polegał na zastąpieniu zmiennych I_{it}^{Pd} i I_{it}^{Pc}, mierzących wyniki innowacji, nakładami na innowacje. Wpływ tej zmiennej jest dodatni, przy czym jest on istotny dla opóźnienia o dwa okresy. Jest to naturalne ze względu na czas, jaki musi nastąpić między poniesieniem nakładów a wdrożeniem innowacji i następującą po nim reakcją zatrudnienia.

Podsumowując, autorzy stwierdzili, że prawie wszystkie miary innowacji wskazują na pozytywny efekt tej wartości na zatrudnienie.

4.5. Kształtowanie się wynagrodzeń w firmach

Zgodnie ze standardowym modelem zaproponowanym przez Mincera [1974] logarytm obserwowalnego wynagrodzenia w okresie t (w_t) zależy od lat nauki (s), potencjalnego doświadczenia na rynku pracy (z) i kwadratu tej zmiennej (z^2):

$$\ln w_t \approx \alpha + \beta s + \gamma z + \phi z^2 \tag{4.14}$$

Równanie Mincera jest oparte na założeniu, że w dowolnym okresie t logarytm obserwowalnego wynagrodzenia pracownika jest równy pieniężnej wartości jego produktywności netto, mierzonej jako jego potencjalne wynagrodzenie netto (potencjalne wynagrodzenie po odjęciu kosztów inwestycji w uzupełnianie wykształcenia). Andini [2009] wykazał, wykorzystując indywidualne dane panelowe, że założenie to nie jest spełnione. Zaproponował on podejście dynamiczne, w którym dopuszcza się proces dostosowania wynagrodzenia obserwowalnego i potencjalnego w czasie. Proces dostosowania opisać można za pomocą modelu:

$$\ln w_t - \ln w_{t-1} = \rho(\ln npe_t - \ln w_{t-1}) \tag{4.15}$$

gdzie npe_t oznacza potencjalne wynagrodzenie netto w okresie t, a $\rho \in [0,1]$ mierzy szybkość dostosowania (ρ = 1 oznacza spełnienie założenia Mincera: $\ln w_t = \ln npe_t$). Dynamiczna wersja równania Mincera (4.14) ma postać:

$$\ln w_{it} \approx \nu_0 + \nu_1 w_{i,t-1} + \nu_2 s_i + \nu_3 z_i + \nu_4 z_i^2 \tag{4.16}$$

gdzie indeks i oznacza pracownika, a między parametrami zachodzą związki: $\nu_0 = \rho\alpha$, $\nu_1 = 1 - \rho$, $\nu_2 = \rho\beta$, $\nu_3 = \rho\delta$, $\nu_4 = \rho\varphi$.

Model (4.16) estymowano za pomocą SGMM Blundella i Bonda na podstawie danych pochodzących z bazy ECHP (patrz podrozdz. 1.2). Zastosowanie SGMM umożliwiło uwzględnienie endogeniczności zmiennych objaśniających, heterogeniczności obiektów i efektów czasowych. Baza danych obejmowała mężczyzn zatrudnionych w Belgii (6873 obserwacji), Danii (2053) i Finlandii (2341) w wieku 18–65 lat, w latach 1994–2001. Sposób konstrukcji zmiennych w, s i z na podstawie szeregów dostępnych w ECHP opisany jest w: Andini [2009] (s. 18). Oszacowane tempo dostosowania ρ waha się od 0,58 dla Danii do 0,782 dla Finlandii. Zatem podstawowym wnioskiem uzyskanym na podstawie modelu (4.16) jest odrzucenie założenia Mincera, co oznacza, że $\ln w_t \neq \ln npe_t$.

Następnym etapem badań było zbadanie konsekwencji wynikających z przyjęcia dynamicznego modelu (4.16) dla określenia korzyści z wykształcenia. Wykazano, że korzyści te w kategoriach wynagrodzenia obserwowalnego zależą od doświadczenia na rynku pracy. Ponadto korzyści te rosną w czasie w początkowym okresie zatrudnienia, a rzeczywiste korzyści przy wejściu na rynek pracy mogą kształtować się znacznie poniżej korzyści potencjalnych. Dodatkowo wykazano, że dynamiczna wersja modelu Mincera może być traktowana jako rozwiązanie prostego modelu negocjacji płac.

W latach 80. XX w. nastąpił znaczny wzrost nierówności płacowych w USA i Wielkiej Brytanii. W innych krajach, np. Francji i Niemczech, zmiany były niewielkie. Wzrost płac w krajach anglosaskich przypisuje się rosnącym korzyściom z wykształcenia i doświadczenia pracowników, ale obserwuje się także większe różnice w wynagrodzeniach pracowników z podobnym wykształceniem i stażem. Wśród czynników sprzyjających zróżnicowaniu płac wskazuje się na zmiany demograficzne, zmiany jakości pracy, zmiany technologiczne, rosnące korzyści z nieobserwowalnych zdolności i zmiany instytucjonalne. Analizę zróżnicowania płac w Hiszpanii przeprowadzili Bover, Bentolila i Arellano [2001]. Dysponowali oni niezbilansowanym panelem 140 545 obserwacji dla 32 291 zatrudnionych mężczyzn w latach 1980–1987. Kluczową zmienną jest poziom umiejętności. Baza nie zawiera danych dotyczących liczby lat kształcenia, ale dostępne są dane o rodzaju wykonywanej pracy. W klasyfikacji wyróżnia się 10 grup zawodowych, spośród których tylko trzy wymagają jednoznacznie określonego poziomu wykształcenia. Analizę skoncentrowano

więc na tych grupach: a) powyżej 5 lat edukacji wyższej, b) powyżej 3 lat edukacji wyższej c) bez wykształcenia wyższego. Pozostałych pracowników o poziomie wykształcenia pomiędzy b) a c) zakwalifikowano do grupy d) średnio wykształconych. Drugą ważną zmienną jest wielkość firmy. Wyróżniono trzy kategorie: małe (do 100 pracowników), średnie (101–1000 pracowników) i duże (powyżej 1000 pracowników).

W pierwszej części pracy autorzy skoncentrowali się na ewolucji korzyści z umiejętności i doświadczenia według sektorów i firm w czasie. Estymowano wspólną składową zarobków (zarobki podstawowe), korzyści z umiejętności dla dwóch najwyżej wykwalifikowanych grup stosunku do pracowników niewykwalifikowanych, a także korzyści z doświadczenia według sektorów i lat.

W drugiej części podjęto próbę wyjaśnienia zróżnicowania oszacowanych uprzednio korzyści. Analizę oparto na modelu negocjacji płac w firmach zatrudniających pracowników wykwalifikowanych i niewykwalifikowanych, działających na niedoskonale konkurencyjnym rynku czynników. Stwierdzono, że płace i korzyści z wykształcenia zależą od siły negocjacyjnej związków zawodowych, zmian technologicznych, stopnia konkurencyjności produkcji i elastyczności płac pracowników wykwalifikowanych. Następnie oszacowano model, w którym uzyskane uprzednio oceny zarobków podstawowych, korzyści z umiejętności i korzyści z doświadczenia zależały od zbioru zmiennych sektorowych i krajowych. Zmienne sektorowe dotyczą: przynależności do związków zawodowych, umów zbiorowych, udziału zatrudnienia w sektorze publicznym, wydatków na BiR, a zmienne krajowe – udziału bezrobocia długookresowego, stopy bezrobocia według wykształcenia i struktury siły roboczej. Podstawowe wnioski są następujące: aktywność związków zawodowych zmniejsza zróżnicowanie płac, a negocjacje na poziomie firm zwiększają je; długotrwałe bezrobocie przyczynia się do spadku płac pracowników niewykwalifikowanych, ale zwiększa korzyści z umiejętności; inwestycje w BiR powodują wzrost płac. Stwierdzono także wzrost zróżnicowania płac wewnątrz grup o jednakowym poziomie wykształcenia, które zależy dodatnio m. in. od udziału pracowników wykwalifikowanych w populacji i od stanu cyklu gospodarczego. Korzyści z doświadczenia rosły w okresie próby; stwierdzono przy tym, że pozytywny wpływ wywierają udział osób młodych i udział zatrudnienia w sektorze publicznym, a stan cyklu ma wpływ negatywny.

4.6. Struktura właścicielska i kapitałowa firmy

Struktura właścicielska i kapitałowa firmy to istotne czynniki, wpływające na jej wyniki produkcyjne i finansowe. We współczesnej teorii ekonomicznej stwierdza się, że firmy mogą wpływać na swoją wartość rynkową i perspektywy na przyszłość poprzez kształtowanie optymalnego stosunku między zadłużeniem

a kapitałem akcyjnym (ang. *equity*). Poniżej zaprezentowane zostaną przykłady badań empirycznych poruszających tę problematykę.

W pracy C r i s c u o l o [2003] analizowany jest związek między strukturą właścicielską firm a ich produktywnością na podstawie danych panelowych, pochodzących z firm działających w Wielkiej Brytanii. Celem pracy było zbadanie, czy prawdziwe są hipotezy, że obecność wśród udziałowców posiadaczy dużych pakietów akcji pozytywnie wpływa na wyniki firmy oraz że istnieje zależność między obecnością akcjonariuszy instytucjonalnych a produktywnością. Podkreślono przy tym możliwości analizy, prowadzonej na podstawie danych panelowych. W cytowanych w tekście oryginalnym badaniach związku produktywności ze strukturą akcjonariuszy, prowadzonych na podstawie danych przekrojowych, konieczne było przyjęcie założenia, że struktura właścicieli akcji pozostaje niezmienna. Wykorzystanie danych panelowych, w tym informacji o zmianach struktury akcjonariuszy w okresie próby, umożliwia analizę przyczynowego związku między własnością a produktywnością, a także uwzględnienie w procesie estymacji endogeniczności nakładów oraz rozważenie możliwej endogeniczności struktury właścicielskiej.

Wykorzystane w badaniu dane pochodzą z lat 1985–1997 i tworzą niezbilansowany panel 2395 obserwacji pochodzących z 268 firm. Na uwagę zasługuje szczegółowość zebranych danych, opisujących strukturę właścicielską. Akcjonariusze podzieleni zostali na osiem grup. Są to: 1) osoby prywatne, 2) fundusze emerytalne, 3) instytucje ubezpieczeniowe, 4) inne instytucje finansowe (kategorie 2), 3), 4) są w badaniu empirycznym zgrupowane w jedną: instytucje finansowe) 5) prywatni klienci banków, 6) instytucje niefinansowe, 7) Sepon Ltd.[6], 8) inni udziałowcy. Dane zawierają informacje dotyczące tożsamości wszystkich akcjonariuszy posiadających udziały przewyższające 0,25% wartości firmy, o ile jest ich co najwyżej 99, a jeśli ich liczba jest większa, to dotyczą one 99 najbardziej znaczących akcjonariuszy.

W celu odpowiedniej specyfikacji równań, umożliwiającej zbadanie postawionych hipotez, autorka posługuje się różnymi miarami związanymi ze strukturą właścicielską, sytuacją finansową firmy i konkurencyjnością produktów na rynku. Dla opisu sytuacji finansowej oraz dla opisu konkurencyjności wymienione są po cztery zmienne. Bardziej szczegółowo rozważana jest struktura właścicielska. Wskazane są dwie zmienne, opisujące tożsamość posiadaczy akcji, trzy miary koncentracji własności i trzy dodatkowe zmienne, opisujące zależności między tożsamością akcjonariusza a wielkością posiadanego udziału. Spośród miar koncentracji własności największe znaczenie w dalszej analizie mają:

---

[6] Giełdowa spółka kliringowa, wskazująca pakiet akcji, będący w trakcie sprzedaży (ang. *the Stock Exhange clearing company, indicating a shareholding that is in the process of being sold*).

TOP5 – skumulowany procent akcji posiadanych przez pięciu największych udziałowców,

INST10 – skumulowany udział akcji posiadanych przez instytucje, zaliczające się do dziesięciu największych udziałowców,

IND10 – skumulowany udział akcji posiadanych przez osoby prywatne, zaliczające się do dziesięciu największych udziałowców.

Model produkcji skonstruowano w oparciu o uogólnioną funkcję produkcji Cobba–Douglasa. Wykorzystano przy tym wskazaną przez Blundella i Bonda [1998][7] możliwość wykorzystania autoregresyjnej struktury składnika losowego do uwzględnienia opóźnień dostosowawczych procesu produkcji i pewnych nieobserwowalnych charakterystyk tego procesu, posiadających wysoki stopień trwałości. Estymowane równanie ma postać:

$$y_{it} = \beta_0 + \beta_1 n_{it} + \beta_2 k_{it} + \beta_3 own_{it} + \alpha_i + \lambda_t + \varepsilon_{it} + m_{it} \tag{4.17}$$

gdzie: $\varepsilon_{it} = \rho\varepsilon_{i,t-1} + e_{it}$, $|\rho| < 1$, y_{it} – logarytm produkcji firmy i w okresie t, n_{it} – logarytm liczby zatrudnionych, k_{it} – logarytm zasobu majątku, own_{it} – miara koncentracji własności i/lub tożsamości udziałowców, α_i – efekty grupowe, λ_t – efekty czasowe, ε_{it} – składnik losowy, m_{it} – błąd pomiaru (wolny od autokorelacji). Po przekształceniach model (4.17) przyjmuje postać:

$$\begin{aligned} y_{it} = {} & \beta_1 n_{it} - \rho\beta_1 n_{i,t-1} + \beta_2 k_{it} - \rho\beta_2 k_{i,t-1} + \beta_3 own_{it} - \rho\beta_3 own_{i,t-1} + \\ & + \rho y_{i,t-1} + \alpha_i(1-\rho) + (\lambda_t - \rho\lambda_{i,t-1}) + e_{it} + m_{it} - \rho m_{i,t-1} \end{aligned} \tag{4.18}$$

lub

$$y_{it} = \pi_1 n_{it} + \pi_2 n_{i,t-1} + \pi_3 k_{it} + \pi_4 k_{i,t-1} + \pi_5 own_{it} + \pi_6 own_{i,t-1} + \pi_7 y_{i,t-1} + \alpha_t^* + \lambda_i^* + \omega_{it} \tag{4.19}$$

Powyższe modele są równoważne ze względu na następujące restrykcje wspólnego czynnika: $\pi_2 = \pi_1 \pi_7$, $\pi_4 = \pi_3 \pi_7$, $\pi_6 = \pi_5 \pi_7$.

Liczba zatrudnionych i zasób majątku są potencjalnie skorelowane z efektami grupowymi i czasowymi, a ponadto obciążone błędami pomiaru, dlatego traktowane są w modelu jako zmienne endogeniczne. Z tego powodu do estymacji modelu (4.18) zastosowana została SGMM. Zdecydowano się przy tym na wybór wersji jednostopniowej, ze względu na możliwe obciążenie estymatorów wariancji ocen parametrów metody dwustopniowej (korekta Windmeijera pochodzi dopiero z roku 2005). Struktura własności udziałów (zmienna *own*) może być zastępowana różnymi instrumentami, w zależności od

[7] Por. równanie (4.12).

tego, czy traktowana jest jako zmienna endogeniczna, słabo egzogeniczna czy egzogeniczna. Właściwość zmiennych instrumentalnych przy różnych założeniach badano testem Sargana lub testem Hausmana.

Pierwszym etapem badania była analiza wpływu struktury właścicielskiej na łączną produktywność czynników produkcji. Zmienną objaśnianą w modelu (4.18) był logarytm wartości dodanej urealnionej za pomocą indeksu cen producenta (dla przemysłu) lub indeksu cen detalicznych (dla usług). Stosowano cztery alternatywne miary struktury własności: udział akcji posiadanych przez instytucje finansowe, udział akcji posiadanych przez osoby prywatne, IND10 i TOP5. Wykazano, że zwiększenie udziału akcji posiadanych przez instytucje finansowe wywiera dodatni wpływ na produktywność: współczynnik długookresowy wynosi 11%, zaś efekt zwiększania udziału akcji posiadanych przez osoby prywatne jest odwrotny (–12% w długim okresie). Ponieważ w okresie próby, tzn. w latach 1985–1997, nastąpił wzrost średniego udziału instytucji finansowych w ogólnej liczbie posiadaczy akcji o 17%, to z punktu widzenia ekonomicznego wynik ten oznacza 1,9-procentowy wzrost produktywności. Wyniki estymacji modeli, w których jako miarę struktury właścicielskiej zastosowano jedną z miar IND10 lub TOP5, pokazały, że założenie o słabej egzogeniczności miar koncentracji jest problematyczne. Na możliwość, że nie jest ono spełnione, wskazywać mogą w zależności od użytej miary: statystyczna nieistotność zmiennej lub negatywny wynik różnicowego testu Sargana. Zatem kolejnym etapem analizy było zbadanie, czy proponowane miary struktury właścicielskiej powinny być traktowane endogenicznie.

Endogeniczność zmiennych INST10 oraz TOP5 (miar struktury właścicielskiej) jest możliwa do zbadania za pomocą różnicowego testu Sargana, gdyż zbiór warunków momentów wykorzystanych przy słabszym założeniu (o endogeniczności danej miary) zawiera się w zbiorze warunków wykorzystanych przy silniejszym założeniu (o słabej egzogeniczności). C. Criscuolo [2003] przyjmowała alternatywne założenia dotyczące egzogeniczności lub endogeniczności INST10 oraz TOP5, traktując przy tym pozostałe zmienne endogenicznie i różnicując założenia o korelacji między efektami grupowymi a rozważaną miarą struktury. Z porównania oszacowanych modeli autorka wyciąga dwa podstawowe wnioski. Po pierwsze, korelacja między efektami grupowymi a miarą INST10 (lub TOP5) powoduje obciążenie w górę oceny parametru przy mierze struktury. Sugeruje to, że akcjonariusze instytucjonalni systematycznie inwestują w akcje spółek o wyższej produktywności. Po drugie, nie można odrzucić hipotezy o słabej egzogeniczności miary INST10 ani hipotezy o endogeniczności miary TOP5. W pracy Criscuolo [2003] wykazano dodatkowo, że pozytywny związek między wzrostem udziałów posiadanych przez instytucje a produktywnością nie zależy od wielkości firmy ani koncentracji kapitału. Badanie to przeprowadzono włączając do modelu dodatkowe zmienne: liczbę zatrudnionych oraz, jednocześnie z miarą udziału akcjonariuszy instytucjonalnych INST10, miarę TOP5.

Badanie Drobetza i Wanzenried [2006], dotyczące determinant zmieniającej się w czasie docelowej struktury kapitału, jest przykładem zastosowania modelu dynamicznych dostosowań. Wpisuje się ono w jeden z ważniejszych nurtów współczesnej ekonomii finansowej, jakim są badania nad optymalną strukturą kapitału firmy. Ponieważ struktura kapitałowa firm może się przejściowo różnić od docelowej, co związane jest z kosztami dostosowań, proces dostosowań został zendogenizowany. Dokładniej, analizowano wpływ czynników specyficznych dla firm, jak również czynników makroekonomicznych na tempo dostosowań do docelowego współczynnika zadłużenia. Próba obejmuje 90 firm szwajcarskich z lat 1991–2001. Zmienną objaśnianą modelu Drobetza i Wanzenried jest obserwowalny poziom zadłużenia LV_{it} (ang. *debt level*, *leverage ratio*), który (przy braku zakłóceń) powinien być równy docelowemu poziomowi długu danej firmy LV_{it}^*, wyznaczonemu jako: $LV_{it}^* = \sum_{j=1}^{L} \alpha_j X_{jit}$, gdzie X_{jit} są wybranymi zmiennymi objaśniającymi wykorzystywanymi w przekrojowych badaniach struktury kapitałowej. Model ma postać:

$$LV_{it} = (1-\beta_0)LV_{i,t-1} - \beta_1 Z_{it} LV_{i,t-1} + \beta_0 \sum_{j=1}^{L} \alpha_j X_{jit} + \beta_1 \sum_{j=1}^{L} \alpha_j Z_{it} X_{jit} + d_t + \eta_i + u_{it} \tag{4.20}$$

gdzie Z_{it} określa tempo dostosowań, η_i są efektami grupowymi, d_t – efektami czasowymi. Model (4.20) oszacowano przy założeniu endogeniczności wszystkich zmiennych objaśniających za pomocą jednostopniowej FDGMM.

Wykorzystanymi w modelu (4.20) zmiennymi, determinującymi strukturę kapitału są: *tang* – stosunek środków trwałych do aktywów ogółem, *size* – wielkość firmy, *growth* – możliwości inwestycyjne, aproksymowane stosunkiem wartości rynkowej do księgowej, i *ROA* – rentowność. Wśród determinant tempa dostosowania do optymalnej struktury kapitałowej autorzy wymieniają dwie grupy czynników: specyficzne dla firm i makroekonomiczne. Na pierwszą grupę czynników składają się zmienne: *growth*, *size* oraz $dist = \left| LV_{it}^* - LV_{it} \right|$ – odległość między obserwowalną i docelową dźwignią. Czynniki makroekonomiczne to różnice między kursami kupna i sprzedaży (ang. *term spread*, *TED spread* i *default spread*) oraz krótkoterminowa stopa procentowa.

Wykazano, że szybsze procesy dostosowawcze zachodzą w firmach szybciej rozwijających się oraz tych, których struktura kapitałowa bardziej odbiega od optymalnej. Ponadto wykazano, że tempo dostosowania jest związane z cyklem gospodarczym – jest ono wyższe, gdy perspektywy ekonomiczne są korzystne. Jednakże nie wykryto systematycznych różnic w tempie dostosowań po podzieleniu firm na borykające się z problemami finansowymi i niemające

takich problemów, co skłoniło autorów do postawienia hipotezy, że wrażliwość tempa dostosowania na zmienne opisujące cykl koniunkturalny jest wyższa dla firm, w których nie występują problemy finansowe. Autorzy wskazują też na ograniczenia ich pracy, którym jest niewielka liczebność próby. Może ona powodować niemożność rozróżnienia tempa dostosowania miedzy firmami doświadczającymi problemów finansowych i niemającymi takich problemów.

W literaturze empirycznej dotyczącej finansów przedsiębiorstw istotną rolę odgrywają również badania nad efektami fuzji lub przejęć firm. Wykazano silny związek między wielkością firmy a wynagrodzeniami zarządu oraz słabszy związek między wynagrodzeniami a wynikami firm. W konsekwencji większość przejęć nie poprawia rentowności ani zysków akcjonariuszy firmy przejmującej. Rezultaty te dały asumpt do badań nad wpływem fuzji lub przejęcia na wynagrodzenia managerów. Literatura przedmiotu na ogół koncentruje się na przejęciach określonego typu, tzn. firm krajowych sektora publicznego. W pracy G u e s t a [2009] analizie poddano zarówno fuzje i przejęcia międzynarodowe, jak i krajowe firm prywatnych i państwowych. Badanie przeprowadzono na podstawie 4528 przejęć dokonanych przez 1408 firm w okresie 1984–2001. Przejęcie zdefiniowano jako sytuację, gdy oferent posiada mniej niż 50% akcji przed dokonaniem zakupu, a po jego dokonaniu jego udział wzrasta do co najmniej 50%. Podstawowy model, na podstawie którego przeprowadzono wnioskowanie, ma postać:

$$\ln(pay_{it}) = \beta_0 + \beta_1 \ln(pay_{i,t-1}) + \beta_2 \ln(sales_{it}) + \beta_3 ROA_{it} + \beta_4 MTBV_{it} + \beta_5 acq_{it} + \\ + \mu_i + \lambda_t + \varepsilon_{it} \tag{4.21}$$

gdzie pay_{it} oznacza najwyższe wynagrodzenie pobierane przez dyrektora firmy i w okresie t, $sales_{it}$ – wartość sprzedaży firmy i odzwierciedla wielkość firmy, ROA_{it} – stosunek wynagrodzeń brutto do wartości aktywów ogółem, $MTBV_{it}$ – stosunek wartości rynkowej do wartości księgowej, acq_{it} – zbiór zmiennych zero-jedynkowych dla roku poprzedzającego i następującego po przejęciu, μ_i, λ_t – efekty: grupowy i czasowy.

Podstawowe wnioski z przeprowadzonych badań są następujące:

– przejęcie skutkuje znaczącym wzrostem płac managerów w pierwszym roku po jego dokonaniu. Wzrost ten jest jednak przejściowy i jest równoważony przez podobny spadek w drugim roku;

– nie stwierdzono, aby zmiany płac zależały od kraju, w którym działa firma przejmowana, ani od jej formy organizacyjnej, jednakże pierwsze przejęcie międzynarodowe ma znacznie większy wpływ na płace managerów niż kolejne takie przejęcia;

– zmiany płac są znacznie większe, kiedy firma przejmowana charakteryzuje się relatywnie wysokim poziomem płac, natomiast nie obserwuje się takiej prawidłowości, gdy firma ta funkcjonuje w kraju o ogólnym wysokim poziomie płac;

– nawet złe przejęcia skutkują wzrostem płac, chociaż wzrosty te są niwelowane spadkiem wartości akcji, również tych posiadanych przez zarząd.

4.7. Mobilność pracowników i transfer kwalifikacji

Celem artykułu H. Gorga i E. Strobla [2005] jest zbadanie różnic w produktywności firm, w zależności od rodzaju doświadczenia zawodowego właściciela, zdobytego przed założeniem własnej firmy. W szczególności autorzy weryfikują tezę, że firmy, których właściciele bezpośrednio przed ich założeniem pracowali w korporacjach międzynarodowych tej samej branży, wykazują większy wzrost produktywności niż pozostałe firmy.

Uważa się, że korporacje międzynarodowe z racji posiadanych środków finansowych mają największy dostęp do nowoczesnych technologii produkcji, technik marketingowych i metod zarządzania oraz oferują najlepsze przeszkolenie swoich pracowników. Zgromadzony w nich kapitał wiedzy i zasad działania jest w pewnym stopniu przejmowany przez lokalne firmy działające w kraju, w którym ma swoją siedzibę dana korporacja. Transfer ten odbywa się trzema sposobami:

– firmy lokalne „uczą się” przez naśladowanie,

– występuje „efekt współzawodnictwa” – lokalne firmy są zmuszone do zwiększania swojej efektywności, aby utrzymać się na rynku,

– następuje przepływ pracowników, którzy wiedzę i doświadczenie nabyte w korporacjach przenoszą do firm lokalnych, których stają się pracownikami lub właścicielami. Wykorzystując zdobyte uprzednio umiejętności, są w stanie lepiej funkcjonować na rynku.

H. Gorg i E. Strobl badali, czy firmy, których właściciele pracowali lub zostali przeszkoleni w korporacjach międzynarodowych wykazują większą produktywność w stosunku do firm prowadzonych przez osoby nieposiadające takiego doświadczenia. Korzystają oni z bazy danych Regional Programme for Enterprise Development zawierającej roczne dane o firmach działających w Ghanie. Pochodzą one z lat 1991–1997 i dotyczą firm działających w czterech wybranych sektorach, w których zatrudnionych jest ok. 70% pracujących w przemyśle, w tym kraju. Zbiór zawiera dane dotyczące m. in.: wielkości produkcji, funduszu płac, wytworzonej wartości dodanej, wielkości zatrudnienia, dokładne informacje o dotychczasowym doświadczeniu zawodowym właściciela lub dyrektora i jego latach nauki. Ankietowano 204 właścicieli lub dyrektorów o różnym doświadczeniu zawodowym.

Pierwszym etapem analizy było porównanie firm prowadzonych przez osoby z doświadczeniem zdobytym w korporacjach międzynarodowych i osoby bez takiego doświadczenia na podstawie kilku miar mówiących o kondycji firmy. Firmy prowadzone przez osoby posiadające takie doświadczenie charakteryzowały się:

– wyższymi wartościami: produkcji na 1 zatrudnionego, dodanej na 1 zatrudnionego, produkcji, dodanej, zatrudnienia, nakładów kapitałowych i materiałowych na 1 zatrudnionego (różnice istotne dla $\alpha = 0{,}01$),

– niższymi płacami i niższą łączną produktywnością czynników produkcji (*TFP*)[8] (różnice istotne dla $\alpha = 0{,}1$).

Ekonometryczna analiza zjawiska transferu kwalifikacji oparta została o estymację kilku wariantów równania produkcji, którego podstawowa specyfikacja jest panelowym modelem dynamicznym ze zmiennymi sztucznymi i ma postać:

$$dy_{it} = \beta_0 + \beta y_{i,t-1} + \beta_2 FT_{it} + \beta_3 FE_{it} + \beta_4 FE_{is} + \beta_5 CO_{it} + \beta_6 CF_{it} + \beta_7 D_j + \beta_8 T_t + \varepsilon_{it} \tag{4.22}$$

gdzie:

dy – tempo wzrostu miary produktywności *y* (w różnicach logarytmów);

CO – wektor zmiennych charakteryzujących właściciela (lata nauki, oraz zmienna zero-jedynkowa przyjmująca wartość 1, jeśli właściciel posiada doświadczenie w pracy dla korporacji międzynarodowej);

CF – wektor zmiennych charakteryzujących firmę (zależny od specyfikacji równania);

FT_i, FE_{ij}, FE_{is} – zmienne zero-jedynkowe, opisujące efekt transferu doświadczeń z korporacji międzynarodowych do własnej firmy *i*, działającej w branży *j*:

$FT_i = 1$, gdy właściciel został przeszkolony przez korporację międzynarodową;

$FE_{ij} = 1$, gdy właściciel posiada doświadczenie w pracy dla korporacji tej samej branży;

$FE_{is} = 1$, gdy właściciel posiada doświadczenie w pracy dla korporacji branży *s*, $s \neq j$;

D, *T* – efekty: grupowe i czasowe;

ε_{it} – składnik losowy.

Estymowano pięć różnych specyfikacji równania (4.22), w których zmiennymi objaśnianymi były: dwie miary produktywności pracy (produkcja na

---

[8] Przyjęto założenie, że *TFP* jest generowany przez proces AR(1): $TFP_{it} = \rho TFP_{i,t-1} + \delta D_t + f_i + v_{it}$, gdzie D_t wyraża efekt specyficzny dla czasu, f_i jest efektem specyficznym dla *i*-tej firmy, v_{it} jest składnikiem losowym.

1 zatrudnionego *q/l*, wartość dodana na 1 zatrudnionego *vad/l*), wartość *TFP* oraz dwie determinanty *TFP* (*q*, *vad*). We wszystkich równaniach objaśniane były przyrosty logarytmów odpowiednich zmiennych. Wyniki jednoznacznie wskazały, że doświadczenie zdobyte przez właściciela lub dyrektora podczas pracy w korporacji tej samej branży dodatnio wpływa na produktywność własnej firmy, natomiast wpływ samego przeszkolenia w takiej korporacji nie jest statystycznie istotny. Praca w korporacji innej branży wywiera natomiast ujemny wpływ na produktywność własnej firmy.

Następnie zbiór zmiennych objaśniających został rozszerzony o nowe zmienne mierzące zdolności właściciela (dyrektora). Są to: lata nauki s_i oraz iloczyny: $FT_i \cdot s_i$, $FE_{ij} \cdot s_i$, $FE_{is} \cdot s_i$, które mają wychwycić związek między poziomem wykształcenia a umiejętnością praktycznego wykorzystania doświadczenia nabytego w korporacjach. Wprowadzono również zmienną zero-jedynkową przyjmującą wartość 1, gdy właściciel (dyrektor) pracował uprzednio w korporacji w tej samej branży. Na podstawie wyników reestymowanych równań stwierdzono, że:

1) lata nauki dodatnio wpływają na zmienne objaśniające we wszystkich pięciu równaniach, ale tylko w dwóch przypadkach wpływ ten jest statystycznie istotny;

2) doświadczenie zdobyte podczas pracy w korporacji w tej samej branży dodatnio wpływa na produktywność własnej firmy, natomiast parametr przy zmiennej $FE_{ij} \cdot s_i$ jest ujemny. Świadczy to według autorów o tym, że osoby słabiej wykształcone mogą więcej skorzystać z doświadczenia zdobytego podczas pracy, zaś lepiej wykształceni przedsiębiorcy mogą tylko w niewielkim stopniu uzupełnić swoją wiedzę;

3) wpływ doświadczenia zdobytego podczas pracy w korporacji innej branży na produktywność własnej firmy jest ujemny, tak samo jak parametr przy zmiennej $FE_{is} \cdot s_i$;

4) odbycie wyłącznie przeszkolenia niepołączonego z pracą dla korporacji nie ma wpływu na produktywność.

4.8. Problem błędów pomiaru zmiennych

W badaniach empirycznych dość często spotykanym problemem są konsekwencje błędu pomiaru zmiennych. Praktyczne próby zastosowania metod, które umożliwiłyby poprawną estymację modeli zawierających zmienne obarczone błędami pomiaru, przedstawione były w podrozdz. 4.2. W kontekście modeli panelowych, zawierających opóźnioną zmienną objaśnianą i nieobserwowalne, stałe w czasie efekty grupowe, właściwymi metodami są te, które bazują na GMM. Mogą one być stosowane w szczególności wówczas, gdy występują

niesystematyczne, tzn. zmienne w czasie (ang. *transient*) błędy pomiaru zmiennych. Istotne jest przy tym podkreślenie zmienności w czasie błędów pomiaru, ponieważ permanentne, addytywne błędy pomiaru, o ile występują, są wchłaniane przez stały w czasie efekt grupowy α_i, a zatem zastosowanie modeli panelowych z założenia daje możliwość ich kontrolowania. Poniżej uzasadnione zostaną możliwości zastosowania FDGMM i SGMM do estymacji modelu, w którym występują zmienne obarczone błędami pomiaru.

Rozważmy model autoregresyjny postaci (3.23), ale przy założeniu, że zmienna y obarczona jest zmiennym w czasie błędem pomiaru, tzn. że zamiast rzeczywistych wartości zmiennej y obserwowane są wartości:

$$\tilde{y}_{it} = y_{it} + m_{it}, \quad i = 1, \ldots, N, t = 1, \ldots, T \tag{4.23}$$

gdzie $\mathrm{E}(m_{it}m_{is}) = 0$ i $\mathrm{E}(m_{it}\varepsilon_{is}) = 0$ dla dowolnego i i $t \neq s$, co oznacza, że błąd pomiaru m_{it} nie wykazuje autokorelacji oraz możliwa jest jego korelacja ze składnikiem losowym w tym samym okresie. Model (3.23) w połączeniu z warunkiem (4.23) zapisać można w postaci:

$$\begin{aligned} \tilde{y}_{it} &= \gamma\tilde{y}_{i,t-1} + \alpha_i + \xi_{it}, \quad |\gamma| < 1, \\ \xi_{it} &= \varepsilon_{it} + m_{it} - \gamma m_{i,t-1}, \end{aligned} \qquad i = 1, \ldots, N, t = 2, \ldots, T \tag{4.24}$$

a jego pierwsze różnice jako:

$$\begin{aligned} \Delta\tilde{y}_{it} &= \gamma\Delta\tilde{y}_{i,t-1} + \Delta\xi_{it}, \quad |\gamma| < 1, \\ \xi_{it} &= \Delta\varepsilon_{it} + \Delta m_{it} - \gamma\Delta m_{i,t-1}, \end{aligned} \qquad i = 1, \ldots, N, t = 3, \ldots, T \tag{4.25}$$

Składnik losowy ξ_{it} modelu (4.24) wykazuje autokorelację, zatem obserwowalne wartości zmiennej y opóźnione o dwa okresy $\tilde{y}_{i,t-2}$ nie są właściwymi instrumentami w równaniu (4.25). Oznacza to, że nie istnieją zmienne instrumentalne dla tego równania, dla $t = 3$, a zatem w celu oszacowania parametru γ wymagane jest, aby $T \geq 4$. Jeśli jednak liczba obserwacji w czasie $T \geq 4$, to spełnione są warunki momentów:

$$\mathrm{E}(\tilde{y}_{i,t-s}\Delta\xi_{it}) = 0 \text{ dla } t = 4, \ldots, T, s \geq 3 \tag{4.26}$$

możliwe jest więc użycie opóźnionych poziomów zmiennej $\tilde{y}$ z okresu $t - 3$ i wcześniejszych jako zmiennych instrumentalnych w równaniu (4.25). Oznacza to, że możliwe jest zastosowanie FDGMM do estymacji modelu, w którym zmienna y obarczona jest zmiennymi w czasie błędami pomiaru.

W omawianym powyżej przypadku możliwe jest również zastosowanie SGMM, co zostanie pokazane poniżej. Załóżmy, że spełniony jest warunek (3.80) postaci: $E(\alpha_i \Delta y_{i2}) = 0$ dla $i = 1, ..., N$. Gdyby nie występowały błędy pomiaru zmiennej y, założone w (4.23), to spełnione byłyby dodatkowe warunki momentów postaci (3.82). Jednakże wynikająca z (4.23), autokorelacja ξ_{it} powoduje, że przyrosty $\Delta\tilde{y}_{it}$ nie są właściwymi instrumentami dla równania na poziomach (4.24). Zastosowanie SGMM jest jednak możliwe przy dodatkowym założeniu, że błąd pomiaru m_{it} nie powoduje korelacji między obserwowalnymi pierwszymi różnicami $\Delta\tilde{y}_{it}$ a efektami grupowymi α_i, tzn.: $E(\alpha_i \Delta m_{it}) = 0$ dla każdego i i $t > 2$. Wtedy spełnione są warunki momentów postaci:

$$E\left(\Delta\tilde{y}_{i,t-2}(\alpha_i + \xi_{it})\right) = 0 \quad \text{dla } i = 1, \ldots, N,\ t = 4, \ldots, T \tag{4.27}$$

co oznacza, że odpowiednio opóźnione przyrosty obserwowalnych wartości zmiennej objaśnianej mogą być wykorzystane jako zmienne instrumentalne w równaniu (4.24).

Warto w tym miejscu podkreślić, że właśnie możliwość uzyskania zgodnych estymatorów w przypadku występowania błędów pomiaru zmiennych, oraz endogeniczności zmiennych objaśniających[9], jest jedną z podstawowych zalet stosowania do estymacji dynamicznych modeli panelowych metod bazujących na GMM. W procesie konstrukcji i estymacji modeli ekonometrycznych bardzo często zachodzi konieczność wprowadzenia do modelu zmiennych, będących przybliżonymi wartościami zmiennych nieobserwowalnych. Równie często zmienne objaśniające są endogeniczne, tzn. skorelowane z bieżącymi i opóźnionymi wartościami składnika losowego. Możliwość wyznaczenia zgodnych estymatorów parametrów modeli, zawierających takie zmienne za pomocą GMM wydaje się przyczyną popularności tej metody w praktyce ekonometrycznej. Własności innych estymatorów, proponowanych dla dynamicznych modeli panelowych, w tym bazujących na metodzie zmiennych instrumentalnych (np. Anderson, Hsiao [1981]), metodzie największej wiarygodności (np. Bhargava, Sargan [1983]; Blundell, Smith [1991]; Hsiao, Pesaran, Tahmiscioglu [2002]), oraz skorygowanym estymatorze wewnątrzgrupowym (Kiviet [1995]), nie są bowiem znane w przypadkach występowania błędów pomiaru lub endogeniczności zmiennych objaśniających. Ponadto, jak zauważają Bond, Hoeffler i Temple [2001], założenia dotyczące błędów pomiaru oraz endogeniczności zmiennych objaśniających, mogące mieć konsekwencje dla właściwości konkretnych instrumentów,

[9] Dobór odpowiednich zmiennych instrumentalnych w przypadku endogeniczności zmiennych objaśniających jest omawiany w rozdziałach poświęconych poszczególnym metodom estymacji, opartych na GMM.

podlegają w przypadku stosowania GMM testowaniu np. za pomocą testu Sargana (por. podrozdz. 3.3).

Korzyści ze stosowania metod bazujących na GMM do estymacji modeli, w których obecne są zmienne obarczone błędami pomiaru i endogeniczne zmienne objaśniające nie ograniczają się do modeli mikroekonomicznych. Są one znaczne również w obszarze mezo- i makroekonomii, szczególnie w modelowaniu wzrostu gospodarczego. Jeden z przykładów zastosowania FDGMM i SGMM do estymacji modelu wzrostu, w przypadku obecności w modelu endogenicznych zmiennych objaśniających oraz błędów pomiaru, jest zawarty w pracy Bonda, Hoeffler i Temple'a [2001]. Jego omówienie znajduje się w podrozdz. 6.3. Wyraźne są też korzyści z zastosowania SGMM w procesie estymacji modeli dla Polski pokazane w podrozdz. 5.7 i 6.5.

ROZDZIAŁ 5

ZASTOSOWANIE DYNAMICZNYCH MODELI PANELOWYCH W ANALIZACH SEKTOROWYCH

Obiektami analiz ekonometrycznych prowadzonych na poziomie mezoekonomicznym często są jednostki klasyfikacji działalności gospodarczej – sekcje, działy *etc*. Badania empiryczne prowadzone na podstawie danych z takich jednostek koncentrują się na zagadnieniach związanych z produktywnością. Często przedmiotem zainteresowania jest łączna produktywność czynników produkcji (TFP). Jest to zmienna bezpośrednio niemierzalna, której oszacowanie jest istotne, ponieważ jej wzrost jest syntetycznym sposobem oceny zmian efektywności procesów produkcyjnych, zachodzących pod wpływem postępu technicznego. Wykorzystanie danych panelowych, a w szczególności szacowanych na ich podstawie modeli dynamicznych daje możliwości modelowania kształtowania się TFP z uwzględnieniem jej dynamiki oraz wyodrębnienia efektów długo- i krótkookresowych.

W niniejszym rozdziale przedstawione zostaną przykłady zastosowania dynamicznych modeli panelowych w analizach funkcjonowania jednostek klasyfikacji działalności gospodarczej. Są to przykłady zaczerpnięte z literatury światowej oraz własne. Ponieważ tematyka prezentowanych analiz oscyluje wokół zagadnień związanych z produktywnością, a w szczególności łączną produktywnością czynników produkcji, w podrozdz. 5.1 przedstawiono podstawy ekonomiczne dla analiz produktywności. Kolejne trzy podrozdz. dotyczą: modelowania TFP według działów przetwórstwa przemysłowego w Wielkiej Brytanii, wpływu czynników makroekonomicznych na TFP sektorach i wpływu technologii informacyjno-telekomunikacyjnych (ICT) na produkcję i produktywność. W podrozdz. 5.5 opisano produktywność w działach sekcji „przetwórstwo” w Polsce, a ostatnie dwa podrozdz. dotyczą możliwości oszacowania i modelowania TFP w tej sekcji.

Podobnie jak uczyniono to w rozdz. 4, poniżej również przedstawione zostanie zestawienie analiz prowadzonych przy zastosowaniu dynamicznych modeli panelowych na podstawie danych sektorowych. Zestawienie to zawarte jest w tab. 5.1.

5.1. Podstawy ekonomiczne dla analiz produktywności

Postęp techniczny jest we współczesnej gospodarce coraz bardziej istotnym czynnikiem produkcji, ponieważ to właśnie on sprawia, że gospodarka znajduje się na ścieżce rozwoju. Mianem postępu technicznego określa się zmiany w technologii i organizacji produkcji prowadzące do zwiększania efektywności procesów gospodarowania. Efektem działania postępu technicznego jest przyrost strumienia produkcji, który nie wynika ze wzrostu zasobów czynników produkcji; prowadzi on w efekcie ostatecznym do wzrostu poziomu życia społeczeństw. Syntetycznym sposobem oceny zmian efektywności procesów produkcyjnych, zachodzących pod wpływem postępu technicznego, jest wzrost łącznej produktywności czynników produkcji (ang. *total factor productivity*, TFP). Ponieważ niemal wszystkie analizy prezentowane w niniejszym rozdziale dotyczą TFP, poniżej przedstawione zostaną podstawy ekonomiczne dotyczące postępu technicznego i sposobów pomiaru łącznej produktywności czynników produkcji.

Tabela 5.1

Przykłady analiz sektorowych z zastosowaniem dynamicznych modeli panelowych

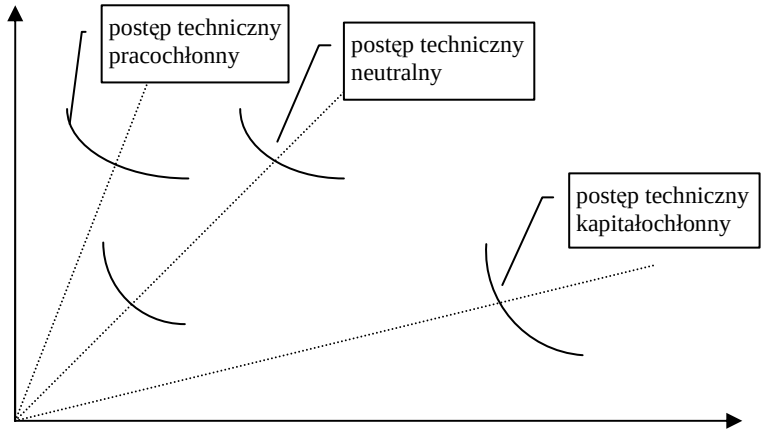
Problem badawczy	Opracowania
Modelowanie produktywności w sektorach	Aghion, Blundell, Griffith, Howitt, Prantl [2004]; **Cameron** [**2006**]
Wpływ czynników makroekonomicznych na TFP w sektorach	Thangavelu, Owyong [2003]; Quella [2006]; **Arizala**, **Cavallo**, **Galindo** [**2009**]
Zaangażowanie ICT a wyniki w sektorach	**O'Mahony**, **Vecchi** [2003]; Dahl, Kongsted, Sorensen [2010]
Analizy TFP według sekcji i działów w Polsce	Welfe [2000, 2001, 2003]; Tokarski [1996, 2002, 2008a]; Zienkowski [2003]; Dańska-Borsiak [2009c]; podrozdz. 5.6 i 5.7 w niniejszym rozdziale

Uwaga: opracowania wyróżnione czcionką półgrubą omawiane są w niniejszym rozdziale, natomiast wyróżnione podkreśleniem nie były prowadzone na podstawie danych panelowych.

Źródło: opracowanie własne.

Pełna definicja postępu technicznego w ujęciu makroekonomicznym, rozumianego jako postęp technologiczno-organizacyjny została sformułowana w pracy Solowa [1957], w kontekście neoklasycznej, dwuczynnikowej funkcji produkcji, postaci: $Y = f(K, L)$, gdzie: Y – wielkość produkcji, K – zasoby kapitału, L – nakłady pracy. Z definicji tej, cytowanej w pracy Świeczewskiej [2007] (s. 59) wynika, że efektem działania postępu technicznego jest

każdy wzrost strumienia produktu, który nie wynika ani z akumulacji kapitału, ani ze wzrostu nakładów pracy. Efekty postępu technicznego są zatem utożsamiane z takim przesunięciem funkcji produkcji, że ustalony poziom produkcji można otrzymać przy mniejszych nakładach. Przesunięciu temu może też towarzyszyć zmiana relacji między czynnikami produkcji, a więc w przypadku funkcji dwuczynnikowej, między pracą a kapitałem. Jeśli przy przesunięciu funkcji produkcji zmiana taka nie następuje, to postęp techniczny nazywa się neutralnym; jeśli stosunek *K*/*L* rośnie, to postęp jest kapitałochłonny (lub pracooszczędny); jeśli zaś stosunek *K*/*L* maleje, to postęp jest pracochłonny (lub kapitałooszczędny). Na rys. 5.1 przedstawiona jest graficzna ilustracja poszczególnych typów postępu technicznego.



Rys. 5.1. Typy postępu technicznego a położenie krzywych jednakowego produktu

Ź r ó d ł o: na podst.: Ś w i e c z e w s k a [2007], s. 61.

W dalszych rozważaniach, w pkt 5.6.1, wykorzystywane będzie pojęcie postępu technicznego w sensie Hicksa. Jest to taki typ postępu neutralnego, w którym nie zmienia się krańcowa stopa substytucji pracy kapitałem. Dwuczynnikowa funkcja produkcji z postępem technicznym w sensie Hicksa ma więc postać:

$$Y = A(t)\, f(K, L) \tag{5.1}$$

gdzie $A(t)$ to czynnik opisujący technologię produkcji, przy czym $\frac{\mathrm{d}A}{\mathrm{d}t} \geq 0$.

Oprócz wyżej wspomnianych typów postępu technicznego wyróżnia się też postęp techniczny nieucieleśniony (niezmaterializowany, ang. *disembodied*) i ucieleśniony (zmaterializowany, ang. *embodied*). Postęp nieucieleśniony to taki, który dotyczy całego zasobu czynników produkcji, przy czym czynniki te

są traktowane jako jednorodne. Wewnętrzna struktura czynników produkcji pozostaje przy tym niezmieniona. Natomiast postęp ucieleśniony dotyczy tylko pewnych grup urządzeń, łącznie z ich obsługą. Następujący pod jego wpływem wzrost produkcji wynika więc ze wzrostu efektywności jednego lub kilku czynników produkcji przy jednoczesnej zmianie ich wewnętrznej struktury. Założenie neoklasyczne o jednorodności czynników produkcji zostaje więc tu uchylone. Więcej na temat tych dwóch typów postępu technicznego, w tym tzw. modeli rocznikowych, zakładających postęp zmaterializowany, znaleźć można m. in. w pracy J u s z c z a k-S z u m a c h e r [1996] (s. 63–89).

Przy założeniu, że postęp jest nieucieleśniony można mówić o postępie zwiększającym efektywność kapitału (ang. *capital-augmenting*), zwiększającym efektywność pracy (ang. *labour-augmenting*) lub zwiększającym jednocześnie efektywność pracy i kapitału (ang. *factor-augmenting*). Kwalifikacja do jednego z tych typów zależy od tego, czy zmiany efektywności odnoszą się do jednego czy do wszystkich czynników produkcji; jednak w żadnym przypadku zmiany te nie dotyczą wewnętrznej struktury czynników.

Efektem postępu technicznego są zmiany efektywności procesu produkcji. Sposób oceny tych efektów jest zależny od skali analizy. W ujęciu mikroekonomicznym ocenia się pojedyncze przedsięwzięcia innowacyjne, a w ujęciu makroekonomicznym oceniane są zmiany efektywności, wyrażające się poprzez zmiany w: strukturze czynników produkcji i relacji między nimi (mierzone najczęściej wzrostem technicznego uzbrojenia pracy); strukturze produkcji; produktywności pojedynczych czynników (np. wydajności, kapitałochłonności *etc*.) lub produktywności wszystkich czynników (tzn. łącznej produktywności czynników produkcji). Produktywność pojedynczego czynnika mierzona jest jako wielkość produkcji przypadająca na jednostkę tego czynnika. Łączna produktywność czynników produkcji (TFP) jest natomiast zmienną nieobserwowalną. Definiowana jest ona jako wielkość produkcji, przypadająca na jednostkę kombinacji czynników produkcji. TFP jest, w odróżnieniu od produktywności poszczególnych czynników, syntetycznym miernikiem efektywności procesu produkcyjnego, wynikających z szeroko rozumianego postępu technicznego.

Wyznaczenia TFP dokonać można za pomocą jednej z dwóch alternatywnych metod. Jedna z nich bazuje na metodach indeksowych[1], a druga wykorzystuje funkcję produkcji. Jeśli funkcja produkcji ma postać Cobba–Douglasa, to poziom TFP (określany jako produkcja wytworzona z jednostki kombinacji czynników produkcji) w okresie *t* wyraża się jako:

$$TFP_t = \frac{Y_t}{f(K_t, L_t)} \tag{5.2}$$

[1] Bliższe informacje na temat metod indeksowych zawarte są np. w pracy Ś w i e c z e w s k i e j [2007] (s. 68–71).

Przekształcenie wyrażenia (5.2) do postaci:

$$Y_t = TFP_t \cdot f(K_t, L_t) \tag{5.3}$$

i porównanie (5.3) z (5.1) prowadzi do wniosku, że poziom TFP odpowiada czynnikowi $A(t)$, opisującemu technologię produkcji w dwuczynnikowej funkcji produkcji z neutralnym postępem technologicznym w sensie Hicksa. Poziom TFP określa więc wielkość przesunięcia funkcji produkcji w kierunku początku układu współrzędnych, wynikającą ze zwiększenia efektywności czynników produkcji (rys. 5.1).

Zróżniczkowanie (5.3) względem czasu i przyjęcie założenia, że każdy czynnik produkcji jest opłacany według jego produktu krańcowego prowadzi do równania:

$$\frac{T\dot{F}P_t}{TFP_t} = \frac{\dot{Y}_t}{Y_t} - \alpha_t^K \frac{\dot{K}_t}{K_t} - \left(1 - \alpha_t^K\right)\frac{\dot{L}_t}{L_t} \tag{5.4}$$

gdzie symbole $\dot{Y}_t$, $T\dot{F}P_t$, $\dot{K}_t$, $\dot{L}_t$ oznaczają pochodne po czasie odpowiednich zmiennych, α_t^K oznacza udział nakładu kapitału w produkcie, a $(\alpha_t^K - 1) = \alpha_t^L$ – udział nakładu pracy w produkcie. Równanie (5.4) jest nazywane równaniem reszt Solowa. Taki sposób pomiaru stopy postępu technicznego (utożsamianego ze stopą wzrostu TFP) oznacza, że stopę wzrostu TFP interpretuje się jako różnicę między stopą wzrostu produktu a ważoną sumą stóp wzrostu podstawowych czynników produkcji, z wagami, które określają udział tych czynników w produkcie. Równane (5.4) można w naturalny sposób uogólnić na przypadek *n*-czynnikowej funkcji produkcji z neutralnym postępem technicznym w sensie Hicksa.

Z równania (5.4) wynika w szczególności, że do oszacowania stopy wzrostu TFP wystarcza znajomość elastyczności produkcji względem poszczególnych jej czynników. Oszacowanie elastyczności wymaga zaś sprecyzowania postaci funkcji produkcji. Zazwyczaj przyjmuje się dwuczynnikową funkcję Cobba–Douglasa w postaci potęgowej:

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha} \tag{5.5}$$

gdzie: Y_t – wielkość produkcji w cenach stałych; A_t – łączna produktywność czynników produkcji ($A_t = TFP_t$); K_t – wartość kapitału w cenach stałych; L_t – nakłady pracy; α, $(1 - \alpha)$ – elastyczności produkcji odpowiednio względem nakładów kapitału i pracy.

Wyznaczona na podstawie funkcji produkcji wielkość stopy wzrostu TFP zależy od postaci funkcji produkcji, sposobu pomiaru czynników produkcji i uwzględnienia dodatkowych, oprócz pracy i kapitału, czynników produkcji (por. Ś w i e c z e w s k a [2007], s. 74–77).

Postać funkcji produkcji wpływa na wielkość elastyczności, a zatem również na wyznaczoną wartość TFP. Pomiar czynników produkcji może być dokonywany przy założeniu homogeniczności nakładów i ich nieograniczonej substytucji, co związane jest z koncepcją nieucieleśnionego postępu technicznego, lub bez tych założeń, i wówczas przyjmuje się, że postęp techniczny jest ucieleśniony. W tym drugim przypadku w odniesieniu do nakładów kapitału stosuje się metodę rocznikowania, a w odniesieniu do nakładów pracy wprowadza się pojęcie kapitału ludzkiego. Pod pojęciem kapitału ludzkiego w szerokim sensie rozumie się zbiór wszystkich cech psychofizycznych danej jednostki, które mogą wpływać na wydajność jej pracy, a w wąskim znaczeniu utożsamia się go z poziomem wykształcenia danej jednostki. Istnieje przy tym kilka alternatywnych mierników kapitału ludzkiego, a wybór jednego z nich podyktowany jest zazwyczaj dostępnością danych statystycznych. Więcej na temat możliwości uwzględnienia kapitału ludzkiego w funkcji produkcji znaleźć można np. w: F l o r c z a k [2007]; J u s z c z a k-S z u m a c h e r [2000].

Łączną produktywność czynników produkcji (A_t we wzorze (5.5)) dekomponować można na: efekty związane z wiedzą ogólnotechniczną (A_t^*), efekty związane ze wzrostem jakości kapitału (A_t^K) oraz efekty związane ze wzrostem jakości pracy (A_t^L). Więcej na temat możliwości dekompozycji łącznej produktywności czynników produkcji, jak również różnic między produkcją efektywną a produkcją potencjalną i wynikających stąd różnic w pomiarze TFP znaleźć można np. w pracach: W e l f e [2001], s. 112–188; W e l f e [2003].

W powyższych rozważaniach postęp techniczny, a zatem również łączna produktywność czynników produkcji, traktowane były egzogenicznie. Próby wyjaśnienia źródeł postępu technicznego (wzrostu TFP) są podejmowane na gruncie teorii wzrostu endogenicznego. W teorii tej przyjmuje się, że postęp techniczny jest efektem akumulacji kapitału ludzkiego oraz innowacji, powstających w wyniku działalności badawczo-rozwojowej. Wzrost TFP jest zatem efektem celowych decyzji inwestycyjnych, podejmowanych przez podmioty gospodarcze. Szersze omówienie teorii wzrostu endogenicznego, endogenizacji postępu technicznego i światowych badań empirycznych prowadzonych na gruncie tych teorii znaleźć można w pracy D. Ś w i e c z e w s k i e j [2007]; wyniki badań empirycznych dotyczących Polski przedstawiono m. in. w pracach: W e l f e [2001]; Z i e n k o w s k i [2003].

5.2. Model TFP według sektorów przemysłu w Wielkiej Brytanii

G. C a m e r o n [2006] skonstruował dynamiczny model panelowy opisujący kształtowanie się TFP w 19 sektorach działalności produkcyjnej (ang. *manufacturing*) w Wielkiej Brytanii w latach 1972–1992. Szczególny nacisk położony został na wyodrębnienie efektów działalności badawczo-rozwojowej (BiR) w każdym z tych sektorów. Postawiono przy tym tezę, że długookresowe współczynniki przy tej zmiennej są heterogeniczne względem sektorów i zależne od pewnych, słabo egzogenicznych charakterystyk działalności przemysłu. Przykładem sektora, który ma wyższą od przeciętnej elastyczność BiR, może być sektor o wysokim technicznym uzbrojeniu pracy lub sektor działający w warunkach silnej konkurencji zagranicznej. Rolę wspomnianych charakterystyk można interpretować przyczynowo, co oznaczałoby, że lepsze techniczne uzbrojenie pracy prowadzi do lepszych wyników działalności BiR albo alternatywnie jako wynik korelacji z pewnymi czynnikami nieobserwowalnymi – wysokie techniczne uzbrojenie pracy może być cechą sektorów, posiadających różnorodne możliwości technologiczne.

Punktem wyjścia analizy G. C a m e r o n a [2006] jest neoklasyczny model Solowa postaci (5.1). Stopa wzrostu TFP $\Delta A/A$ odpowiada tej składowej stopy wzrostu produkcji, która nie wynika z akumulacji kapitału, ani ze wzrostu nakładów pracy. Przyrost TFP w *i*-tym sektorze w okresie *t* aproksymowano za pomocą następującego indeksu Thörnquista–Theila:

$$\log\left(\frac{A_{i,t+1}}{A_{it}}\right) = \log\left(\frac{Y_{i,t+1}}{Y_{it}}\right) - (1-\overline{\alpha}_{it})\log\left(\frac{K_{i,t+1}}{K_{it}}\right) - \overline{\alpha}_{it}\log\left(\frac{L_{i,t+1}}{L_{it}}\right) \qquad (5.6)$$

gdzie $\overline{a}_{it} = (a_{it} + a_{i,t+1})/2$, zaś a_{it} oznacza udział kosztów pracy w wartości dodanej w sektorze *i* w okresie *t*. Roczne procentowe stopy wzrostu TFP, wyznaczone według wzoru (5.6), różnią się znacząco pomiędzy sektorami. Autor podkreśla, że prezentowana tu i wykorzystana w dalszych badaniach koncepcja szacowania TFP jest wrażliwa na błędy pomiaru wynikające z użycia pojedynczego deflatora (ang. *single deflation bias*)[2]. W zastosowanym, w dalszej analizie, modelu uwzględniono dwie zmienne, które mają za zadanie skorygowanie tych błędów. Są to: stosunek cen zaopatrzenia do cen sprzedaży i stosunek cen importu do cen sprzedaży.

[2] Wartość dodana jest deflowana przy użyciu deflatora produkcji globalnej, podczas gdy produkcja globalna powinna być deflowana przy użyciu deflatora produkcji globalnej, a nakłady – przy użyciu deflatora nakładów.

Dynamiczny model kształtowania się TFP ma postać:

$$\log TFP_{it} = \alpha_i + \alpha_1 \log TFP_{i,t-1} + \alpha_2 \log RD_{i,t-2} + \boldsymbol{\psi}\mathbf{X}_{i,t-1} \log RD_{i,t-2} + \boldsymbol{\varphi}\,\mathbf{Z}_{i,t-1} + \varepsilon_{it} \tag{5.7}$$

gdzie: *RD* – stosunek nakładów na BiR do nakładów na środki trwałe. Elementami wektora **Z** są: wskaźnik wykorzystania mocy, wspomniane powyżej zmienne korygujące błąd użycia pojedynczego deflatora, kapitał ludzki (logarytm stosunku liczby pracowników średnio i wysoko wykwalifikowanych do liczby wszystkich pracujących), wskaźnik unionizacji (logarytm odsetka pracowników fizycznych objętych umowami zbiorowymi pomnożonego przez odsetek wszystkich pracowników fizycznych), efekty czasowe. Elementami wektora **X** jest osiem charakterystyk działalności produkcyjnej poszczególnych sektorów, wyrażonych jako logarytmy wybranych wskaźników, mierzone względem przeciętnej dla całej działalności produkcyjnej[3]. Lista tych zmiennych jest zawarta w tekście oryginalnym w tab. 2 (s. 128). Zmienna $\mathbf{X}_{i,t-1} \log RD_{i,t-2}$ odzwierciedla interakcję zmiennej *RD* z sektorowymi charakterystykami działalności, a zatem pozwala uwzględnić zróżnicowanie efektów działalności BiR między sektorami. Jej wprowadzenie ma na celu uchylenie założenia konstrukcyjnego modelu o stałości parametrów strukturalnych względem obiektów i czasu. G. Cameron podkreśla również, że analiza wpływu efektów działalności BiR na TFP jest możliwa dzięki zastosowaniu danych panelowych, gdyż tylko wówczas możliwe jest odseparowanie tych efektów od trendu czasowego.

Oszacowanych zostało 9 alternatywnych wersji modelu (5.7). Różnią się one zastosowaną metodą estymacji i zbiorem zmiennych objaśniających. Zastosowane metody estymacji są być może nieco dyskusyjne. Ponieważ wymiar czasowy próby jest długi (21 lat), a liczba obiektów wynosi 19, autor stwierdza, że wykorzystanie specyficznych metod, proponowanych w literaturze do estymacji dynamicznych modeli panelowych (metoda Andersona–Hsiao, GMM) nie jest konieczne. Prawdopodobne jest bowiem, że liczba obserwacji w czasie wystarczy, by estymator wewnątrzgrupowy (WG) był zgodny[4]. W związku z tym, do estymacji modelu zastosowano ważoną MNK (WMNK) oraz KMNK[5]. Ponieważ zastosowanie WMNK nie usunęło problemu heteroske-

---

[3] Np. jednym z elementów wektora **X** jest $\log(K_i/L_i) - \log(K/L)$.

[4] G. Cameron powołuje się na wyniki S. N i c k e l l a [1981], przytoczone w podrozdz. 3.1 niniejszej publikacji. Nickell stwierdził, że zgodność estymatora WG dla panelowych modeli dynamicznych zależy od tego, czy wymiar czasowy *T* jest wystarczająco duży; obciążenie tego estymatora jest rzędu $O(1/T)$.

[5] W tekście oryginalnym metoda nazwana jest OLS, dlatego w niniejszym opracowaniu stosowany jest polski odpowiednik – KMNK. Jednakże zawarte w odpowiednich tabelach wyniki estymacji oraz ich opis w tekście, w szczególności uwzględnienie efektów grupowych, sugerują,

dastyczności, zatem dalszą analizę oparto na wynikach KMNK z błędami szacunku typu White'a, odpornymi na heteroskedastyczność.

W procesie estymacji część zmiennych, uwzględnionych w wyjściowym modelu (5.7), okazała się nieistotna i została usunięta. W szczególności elementy wektora **Z**: kapitał ludzki i wskaźnik unionizacji były łącznie nieistotne, zaś ich interakcje z *RD* – łącznie istotne, co doprowadziło do wniosku, że zmienne te nie wywierają bezpośredniego wpływu na wzrost, oddziałując raczej przez wpływ na elastyczność *RD*. Zmiennymi nieistotnymi okazały się też trzy składowe wektora **X**. Ponadto w pierwotnej wersji model (5.7) zawierał efekty czasowe, które w wersji ostatecznej zostały zastąpione innymi zmiennymi czasowymi, stałymi względem obiektów – trendem, miarą konkurencyjności i zmienną odzwierciedlającą długookresowy efekt zmian stosunku nakładów do wyników. Celem wprowadzenia tej ostatniej było skorygowanie systematycznych błędów pomiaru wartości dodanej (wspomniany problem pojedynczego deflatora). Najważniejsze wnioski, wynikające z oszacowania ostatecznej wersji modelu (5.7), przytoczone są poniżej.

Długookresowy efekt zmiennej *RD* wynosi 0,24, ocena parametru dla współczynnika konkurencyjności ma wartość 0,265, a długookresowy efekt zmian stosunku nakładów do wyników – 0,645. Zmienna czasowa okazała się nieistotna. Spośród elementów wektora **X** dodatni wpływ na efektywność *RD*, a zatem na TFP, wywierają: techniczne uzbrojenie pracy $\log(K_i/L_i) - \log(K/L)$, $\log RD_i - \log RD$, stosunek nakładów do wyników. Ujemny wpływ mają stosunek *RD* do kapitału i współczynnik unionizacji. Świadczy to o sektorowym zróżnicowaniu elastyczności TFP względem *RD*. Sektory charakteryzujące się lepszym technicznym uzbrojeniem pracy i większym wykorzystaniem dóbr pośrednich, dostarczanych przez sektory technologicznie zaawansowane, mają wyższą elastyczność TFP względem *RD*. To samo dotyczy sektorów, dla których stosunek importu do sprzedaży wewnętrznej jest wyższy, co świadczy o istotności otwarcia na zagranicę. Do zmniejszenia elastyczności TFP względem *RD* przyczynia się wyższy stosunek *RD* do kapitału i, pośrednio, współczynnik unionizacji. Na zakończenie, na podstawie uzyskanych wcześniej wyników G. Cameron wyznaczył elastyczności zmiennej *RD* dla wszystkich sektorów. Najwyższą elastyczność wykazały sektory: działalność wydawnicza (*paper & printing*) oraz produkcja wyrobów z gumy i tworzyw sztucznych (*rubber & plastics*), zaś niskie elastyczności obserwowano dla sektorów: produkcja wyrobów chemicznych (*chemicals*), produkcja maszyn i urządzeń (*machinery*) i pozostała działalność (*other manufacturing*).

że w rzeczywistości szacowano parametry modelu FE. Stosowanie nazwy OLS jest, jak można się domyślać, uzasadnione tym, że stosunkowo niewielki wymiar przekrojowy próby pozwolił na bezpośrednie oszacowanie parametrów przy zmiennych zero-jedynkowych, bez konieczności stosowania przekształcenia wewnątrzgrupowego (por. podrozdz. 2.1).

5.3. Wpływ rynków finansowych na TFP w sektorach

Pozytywny wpływ rozwoju na rynkach finansowych na wzrost gospodarczy jest odnotowywany w wielu opracowaniach. Często wskazywanym kanałem, przez który udzielone kredyty wpływają na wzrost gospodarczy jest produktywność. Zasadnicza idea mechanizmu oddziaływania kredytu jest taka, że rynki finansowe mogą zwiększać produktywność przez efektywną alokację kapitału, a zwłaszcza przesuwanie kapitału z sektorów upadających do rozwijających się. Celem badania A r i z a l a, C a v a l l o i G a l i n d o [2009] jest analiza związków między wzrostem TFP w sektorach a sytuacją na rynkach finansowych. Autorzy przeprowadzili badanie na podstawie próby z lat 1963–2003 dla 26 sekcji przemysłu przetwórczego w 77 krajach. Próba taka umożliwiła analizę zróżnicowania zmian TFP pod wpływem decyzji finansowych w przekroju zarówno krajowym, jak i sektorowym.

Na podstawie danych dotyczących liczby zatrudnionych, wartości dodanej i nakładów brutto na środki trwałe skonstruowano trzy alternatywne miary TFP na poziomie sekcji i krajów, spośród których preferowaną jest miara wyznaczona ze wzoru:

$$ltfp_{ict} = y_{ict} - \alpha_{Ki} k_{ict} - (1 - \alpha_{Ki}) l_{ict} \tag{5.8}$$

gdzie: y_{ict} oznacza logarytm realnej wartości dodanej w sekcji i, kraju c, okresie t, k_{ict} – logarytm nakładów brutto na środki trwałe; l_{ict} – logarytm liczby zatrudnionych. Parametr α_{Ki} jest elastycznością produkcji względem kapitału w sekcji i, a określono go na podstawie funkcji Cobba–Douglasa jako:

$$\alpha_{Ki} = \frac{K_i}{VA_i} r \tag{5.9}$$

gdzie $r = (\alpha_K VA) / K$ wyznaczono na podstawie danych o wartości dodanej VA i nakładach brutto na środki trwałe K w USA, oszacowanej przez B e r n a n k e i G u r k a y n a k a [2001] elastyczności produkcji względem pracy α_L w tym kraju i założeniu $\alpha_K = 1 - \alpha_L$.

Pozostałe dwie miary są wysoko skorelowane z miarą (5.8) i zostały również wykorzystane w pracy do sprawdzenia odporności wyników.

W celu porównania efektów decyzji finansowych na wzrost TFP, na poziomie sekcji i krajów, oszacowano model postaci:

$$\Delta ltfp_{ict} = \alpha Share_{ict} + \beta (RZ_i \cdot FD_{ct}) + \eta_{it} + \lambda_{ct} + \varepsilon_{ict} \tag{5.10}$$

gdzie z wszystkich danych wyznaczono średnie pięcioletnie w celu wyeliminowania fluktuacji cyklicznych i skoncentrowania się na efektach długookresowych. Zatem $\Delta ltfp_{ict}$ oznacza przeciętną stopę wzrostu TFP w sekcji *i*, kraju *c* w okresie 5 lat; $Share_{ict}$ – udział wartości dodanej wytworzonej w sekcji *i* w wartości wytworzonej w przemyśle wytwórczym kraju *c* na początki danej pięciolatki; RZ_i – miarę zależności sekcji *i* od finansowania zewnętrznego; FD_{ci} – miarę rozwoju finansowego kraju *c*. Wprowadzenie zmiennej $(RZ_i \cdot FD_{ct})$ umożliwia zbadanie, czy rozwój finansowy, mierzony jako dostępność kredytu dla firm prywatnych w stosunku do PKB w gospodarce, silniej wpływa na te sekcje, które z powodów technicznych są bardziej uzależnione od kredytu. Znak parametru β świadczy o średnim wpływie rozwoju finansowego kraju *c* na wzrost poziomu TFP sekcji *i*. η_{it} i λ_{ct} są zmiennymi w czasie efektami specyficznymi odpowiednio dla sekcji i kraju, traktowanymi jako nielosowe.

W hipotezie roboczej przypuszcza się, że sekcje bardziej zależne od funduszy zewnętrznych będą wykazywały wyższy wzrost TFP w tych krajach, w których system finansowy jest bardziej rozwinięty. Najważniejsze wnioski płynące z estymacji modelu (5.10) to:

– dostępność kredytu nie tylko prowadzi do wzrostu inwestycji, ale również zachęca do wdrażania nowych technologii;

– wniosek powyższy jest prawdziwy dla całej próby, jak również dla podpróby obejmującej kraje rozwijające się. Zdaniem autorów oznacza to, że w krajach tych, w których dostępność kredytu jest z reguły bardziej ograniczona niż w krajach rozwiniętych, znaczącą rolę mogą odegrać bezpośrednie inwestycje zagraniczne;

– ocena parametru przy zmiennej $(RZ_i \cdot FD_{ct})$ jest dodatnia, co oznacza, że rozwój finansowy wpływa na wartość dodaną również poprzez wzrost TFP;

– im wyższy poziom zmiennej RZ_i tym silniejszy krańcowy wpływ rynków finansowych na produktywność. Wpływ wzrostu dostępności kredytu o wartość 1 odchylenia standardowego waha się od 0 do 0,6% rocznie w zależności od sekcji. W świetle tego, że przeciętny wzrost TFP w próbie wynosi 0,3%, rocznie jest to rezultat bardzo znaczący.

Autorzy wykazali ponadto, że przedstawione powyżej wnioski są niezależne od doboru próby i specyfikacji modelu.

5.4. Wpływ technologii informacyjno-telekomunikacyjnych (ICT) na wyniki w sektorach

We współczesnej ekonomii wzrost produktywności pracy dzieli się na wzrost wynikający ze wzrostu kapitału na pracującego i wzrost wynikający ze wzrostu TFP. Nie ma też wątpliwości, że inwestycje w technologie informacyj-

no-telekomunikacyjne (ang. *information and communications technology*, ICT) miały znaczący wpływ na wzrost produktywności pracy zarówno w USA w latach 90., jak i w innych krajach OECD. Nie ma natomiast zgody, czy wpływ kapitału ICT miał wpływ na stopę postępu technicznego lub TFP. W opracowaniu M. O'M a h o n y i M. V e c c h i [2003] wykazano, że taki wpływ występuje i że jest on dodatni. Autorki dysponowały danymi ze wszystkich nierolniczych sektorów gospodarki: 31 z USA i 24 z Wielkiej Brytanii z lat 1976–2000. Podkreślają one też konieczność zastosowania adekwatnych metod estymacji, uwzględniających heterogeniczność poszczególnych sekcji i zbadania stacjonarności danych. W rozważanym kontekście heterogeniczność jest oczywista: np. w komunikacji 80% kapitału jest kapitałem ICT, a w przemyśle odzieżowym – tylko 10%. Konieczność zbadania stacjonarności wynika zaś z długości szeregu – 25 lat.

Wcześniejsze badania bazujące na danych panelowych, nieuwzględniające heterogeniczności i stacjonarności, skutkowały ujemnymi ocenami wpływu kapitału ICT na TFP. O'Mahony i Vecchi [2003] pokazały, że wnioski takie były efektem zastosowania niepoprawnej metody estymacji. W pierwszym etapie badania autorki oszacowały model statyczny postaci:

$$\Delta q_{it} = \alpha_{0i} + \alpha_{1i}\Delta l_{it} + \beta_{1i}\Delta k_{it}^{\mathrm{N}} + \beta_{2i}\Delta k_{it}^{\mathrm{I}} + \varepsilon_{it} \tag{5.11}$$

gdzie: q_{it} oznacza logarytm produkcji realnej, l_{it} – nakład pracy (liczba przepracowanych godzin), zaś całkowity nakład kapitału zdekomponowano na dwa typy kapitału: kapitał ICT (k_{it}^{I}) i kapitał spoza ICT (k_{it}^{N}).

Model (5.11) został oszacowany na poziomach i przyrostach za pomocą standardowych metod estymacji panelowej, traktując go jako model FE. Uzyskana ocena parametru przy zmiennej k_{it}^{I} jest ujemna i/lub nieistotna w zależności od tego, czy jest ona wyznaczona na podstawie całej próby czy dla każdego kraju osobno.

Do zbadania stacjonarności[6] wykorzystano test Ima, Pesarana i Shina (IPS) oraz test Hadri'ego. Wynik testu IPS wskazuje na odrzucenie H_0, że wszystkie obiekty mają pierwiastek jednostkowy na poziomie istotności 0,05 dla zatrudnienia i na poziomie 0,1 dla obu typów kapitału. Pamiętać jednak trzeba, że hipoteza alternatywna testu IPS dopuszcza możliwość, aby część obiektów miała pierwiastek jednostkowy, a pozostałe nie (wzór (D.I.4)). Zastosowany następnie test Hadri'ego pozwala na odrzucenie H_0 o stacjonarności dla wszystkich zmiennych.

[6] Na temat testowania stacjonarności dla danych panelowych czytaj w dodatku I.

Do estymacji modelu zastosowano zatem estymator *Pooled Mean Group*, zaproponowany przez P e s a r a n a, S h i n a i S m i t h a [1999]. Metoda ta jest rozszerzeniem modelu korekty błędem na modele panelowe. Na parametry długookresowe nakłada się założenie o homogeniczności, a współczynnik korekty błędem i inne parametry krótkookresowe określa się jako średnie grupowe. Założenie homogeniczności parametrów długookresowych oznacza, że w długim okresie elastyczności produkcji względem jej czynników będą takie same dla wszystkich sektorów. Heterogeniczność sektorową uwzględnia się dopuszczając różnice w dynamice krótkookresowej między sektorami. Szacowany model dynamiczny ma postać:

$$\Delta q_{it} = \phi_i (q_{i,t-1}) - \theta_{01} - \theta_{1i} l_{it} + \theta_{2i} k_{it}^{\mathrm{N}} + \theta_{3i} k_{it}^{\mathrm{I}}) - \delta_{11} \Delta l_{it} - \delta_{21} \Delta k_{it}^{\mathrm{N}} - \delta_{31} \Delta k_{it}^{\mathrm{I}} + \varepsilon_{it} \tag{5.12}$$

Model (5.12) estymowany był zarówno na podstawie danych indywidualnych, jak i danych przekształconych poprzez odjęcie od każdej obserwacji średniej dla danego okresu (ang. *cross-sectionally demeaned data*). Przekształcenie takie odpowiada wprowadzeniu do modelu czasowych zmiennych zero-jedynkowych. Ponadto estymacje przeprowadzono na podstawie całego zbioru danych oraz osobno dla USA i Wielkiej Brytanii. Dla całej próby uzyskano znaczący statystycznie i silny długookresowy wpływ kapitału ICT na produkcję. Ocena parametru przy zmiennej k_{it}^{I} na podstawie danych uśrednionych jest większa niż ocena uzyskana na podstawie danych indywidualnych. Wydaje się to potwierdzeniem nieadekwatności standardowych metod panelowych w rozważanym przypadku. Parametr korekty błędem jest statystycznie istotny i w przypadku danych uśrednionych sugeruje szybsze roczne tempo dostosowania. Krańcowe korzyści z kapitału wynoszą ok. 18% zarówno dla kapitału ICT, jak i spoza ICT na podstawie danych indywidualnych, a na podstawie danych uśrednionych odpowiednio 22 i 12%. Wyniki estymacji modelu (5.12) osobno dla USA i Wielkiej Brytanii są generalnie zbliżone do opisanych powyżej, jednak korzyści z kapitału ICT są wyższe w USA niż w Wielkiej Brytanii.

5.5. Produktywność czynników produkcji w sekcji D w Polsce

Empiryczne analizy własne, prezentowane w podrozdz. 5.6 i 5.7 przeprowadzone zostały na podstawie danych z lat 1998–2007. Obowiązującą wówczas klasyfikacją działalności gospodarczej była polska klasyfikacja działalności

(PKD 1997)[7], zatem objęte badaniem działy wyodrębnione zostały zgodnie z tą klasyfikacją. Analizą objęto działy podsekcji D „przetwórstwo przemysłowe", które zgodnie z PKD 1997 oznaczone są numerami 15–37.

Tabela 5.2

Klasyfikacja PKD 1997 (sekcje C, D, E)

Sekcja	Nr działu	Nazwa
C + **D** + **E**		**przemysł**
C		**górnictwo**
CA	10	w tym górnictwo węgla kamiennego i brunatnego; wydobywanie torfu
D		**przetwórstwo przemysłowe**
DA	15	produkcja artykułów spożywczych i napojów
DA	16	produkcja wyrobów tytoniowych
DB	17	włókiennictwo
DB	18	produkcja odzieży i wyrobów futrzarskich
DC	19	produkcja skór wyprawionych i wyrobów ze skór wyprawionych
DD	20	produkcja drewna i wyrobów z drewna oraz z korka (z wyłączeniem mebli), wyrobów ze słomy i materiałów używanych do wyplatania
DE	21	produkcja masy włóknistej, papieru oraz wyrobów z papieru
DE	22	działalność wydawnicza; poligrafia i reprodukcja zapisanych nośników informacji
DF	23	wytwarzanie koksu, produktów rafinacji ropy naftowej i paliw jądrowych
DG	24	produkcja wyrobów chemicznych
DH	25	produkcja wyrobów gumowych i z tworzyw sztucznych
DI	26	produkcja wyrobów z pozostałych surowców niemetalicznych
DJ	27	produkcja metali
DJ	28	produkcja metalowych wyrobów gotowych, z wyłączeniem maszyn i urządzeń
DK	29	produkcja maszyn i urządzeń, gdzie indziej niesklasyfikowana
DL	30	produkcja maszyn biurowych i komputerów
DL	31	produkcja maszyn i aparatury elektrycznej, gdzie indziej niesklasyfikowana

[7] Obowiązującą obecnie klasyfikacją działalności gospodarczej jest polska klasyfikacja działalności gospodarczej (PKD 2007).

DL	32	produkcja sprzętu i urządzeń radiowych, telewizyjnych i telekomunikacyjnych
DL	33	produkcja instrumentów medycznych, precyzyjnych i optycznych, zegarów i zegarków
DM	34	produkcja pojazdów samochodowych, przyczep i naczep
DM	35	produkcja pozostałego sprzętu transportowego
DN	36	produkcja mebli; działalność produkcyjna, gdzie indziej niesklasyfikowana
DN	37	przetwarzanie odpadów
E		**wytwarzanie i zaopatrywanie w energię elektryczną, gaz, wodę**
E	40	wytwarzanie i zaopatrywanie w energię elektryczną gaz, parę wodną i gorącą wodę
E	41	pobór, uzdatnianie i rozprowadzanie wody

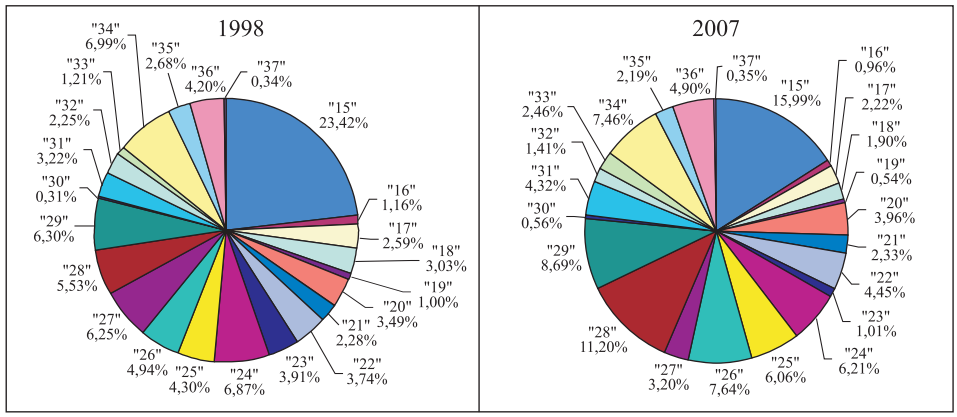
Ź r ó d ł o: opracowanie własne na podstawie informacji GUS.

Część struktury PKD 1997, odnosząca się do przemysłu, przedstawiona jest w tab. 5.2. Zaznaczyć trzeba, że chociaż z dniem wejścia w życie rozporządzenia dotyczącego PKD 2007 klasyfikacja PKD 1997 straciła aktualność, to może ona być stosowana do działalności oznaczonej zgodnie z tą klasyfikacją przed dniem wejścia w życie rozporządzenia RM z 24 grudnia 2007, jednak nie dłużej niż do 31 grudnia 2009 r.

Przed omówieniem zasadniczych badań własnych, dotyczących modelowania łącznej produktywności czynników produkcji (TFP) według działów, interesujące wydaje się dokonanie statystycznej analizy kształtowania się zmiennych powiązanych z TFP, takich jak wartość produkcji i produktywności jej czynników.

W okresie objętym analizą nastąpił wzrost produkcji sprzedanej sekcji D z ok. 334 887 mln zł w roku 1998 do ok. 543 243 mln zł w roku 2007 (ceny stałe z 1998 r.). Zmieniła się również struktura tej produkcji według działów. Największy spadek udziału produkcji sprzedanej (o prawie 7,5 pkt proc.) miał miejsce w dziale 15: „produkcja artykułów spożywczych i napojów". Spadek udziału o ok. 3 pkt proc. odnotowano też w działach 27: „produkcja metali" i 23: „wytwarzanie koksu, produktów rafinacji ropy naftowej i paliw jądrowych". Działy, które zwiększyły swój udział w produkcji sprzedanej sekcji D, to: 28: „produkcja metalowych wyrobów gotowych" (o 5,7 pkt proc.), 26: „produkcja wyrobów z pozostałych surowców niemetalicznych" i 29: „produkcja maszyn i urządzeń" (po ok. 2,5 pkt proc.), 25: „produkcja wyrobów gumowych i z tworzyw sztucznych" (1,8 pkt proc.). W innych działach wahania

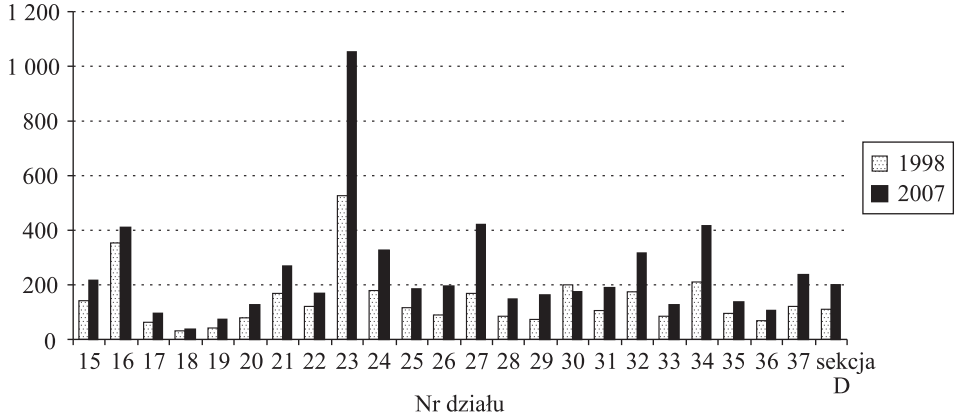
udziałów były mniejsze, nieprzekraczające 2 pkt proc. Tendencje te zilustrowane są na rys. 5.2.



Rys. 5.2. Struktura produkcji sprzedanej według działów sekcji „D”

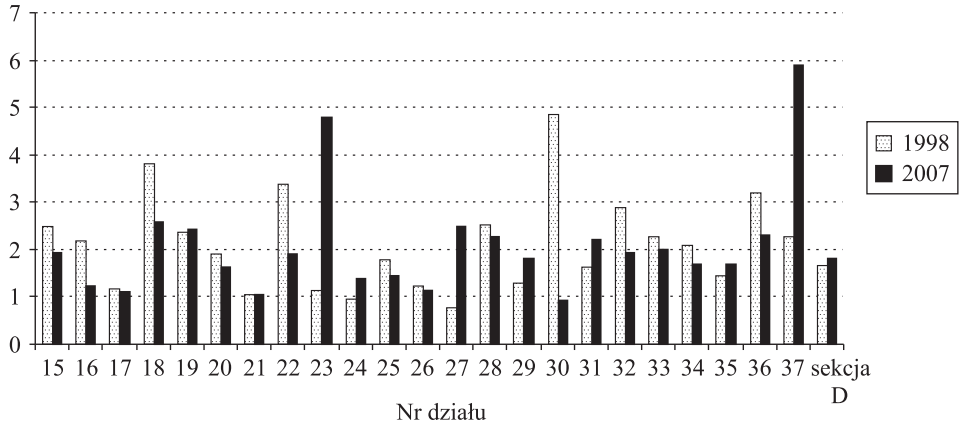
Ź r ó d ł o: opracowanie własne.

Produktywność głównych czynników produkcji: pracy i kapitału, liczonych jako stosunek wielkości produkcji do zasobu danego czynnika również zmieniała się w okresie analizy. Na rys. 5.3 i 5.4 przedstawiono porównanie wydajności pracy (produkcji sprzedanej w tys. zł na jednego pracującego) oraz produktywności kapitału (stosunku produkcji sprzedanej do wartości środków trwałych brutto) w pierwszym i ostatnim roku analizowanego okresu.



Rys. 5.3. Produkcja sprzedana na pracującego (w tys. zł, ceny stałe)

Ź r ó d ł o: opracowanie własne.



Rys. 5.4. Stosunek produkcji sprzedanej do wartości środków trwałych brutto

Ź r ó d ł o: opracowanie własne.

Wzrokowa analiza rys. 5.3 pozwala na stwierdzenie, że we wszystkich działach, za wyjątkiem działu 30: „produkcja maszyn biurowych i komputerów", miał miejsce wzrost wydajności pracy na przestrzeni lat 1998–2007. Skala tego wzrostu była zróżnicowana. Stopa wzrostu wydajności pracy w skali całej sekcji D wyniosła 0,83. Wyższymi stopami wzrostu charakteryzowały się działy o numerach: 27, 26, 29, 23, 34, 37, 32, 24, przy czym najwyższą stopą wzrostu (równą 1,55) charakteryzował się dział 27: „produkcja metali". Najniższe stopy wzrostu odnotowano zaś w działach: 30 (wartość –0,13, zatem spadek wydajności nie był znaczący), 18 (wartość 0,14) i 16 (wartość 0,17).

Zmiany produktywności kapitału między rokiem 1998 a 2006, liczonej jako stosunek produkcji sprzedanej do wartości środków trwałych brutto były, jak wynika z rys. 5.3, różnokierunkowe. Aż w piętnastu działach nastąpił spadek produktywności, największy w dziale 30: „produkcja maszyn biurowych i komputerów". Stopa wzrostu produktywności kapitału w tym dziale wyniosła –0,81, co w połączeniu ze spadkiem wydajności pracy w tym dziale jest wysoce zastanawiające, zwłaszcza że zjawisko to dotyczy działu nowoczesnego, o dużym potencjale – procentowy udział produkcji sprzedanej tego działu wzrósł w analizowanym okresie o 0,24 pkt proc. (patrz rys. 5.1). Działami o relatywnie dużych co do modułu, ujemnych wartościach stopy wzrostu (od 0,43 do 0,27) były też działy o numerach: 16, 22, 32, 18, 36. W sekcji D jako całości odnotowano jednak nieznaczny wzrost produktywności kapitału – stopa wzrostu wyniosła 0,10. Przyczyniły się do tego wzrosty produktywności w działach: 23, 27, 37, 24, 29, 31, 35. Szczególnie korzystna była sytuacja w trzech pierwszych wymienionych działach, gdzie stopy wzrostu wyniosły odpowiednio: 3,16, 2,22 oraz 1,61. O efektywnym funkcjonowaniu działów: 23, 27, 37, 24, 29 świadczy również to, że charakteryzowały się one stopami wzrostu wydajności pracy wyższymi niż właściwa dla sekcji D jako całości.

5.6. Szacowanie łącznej produktywności czynników produkcji (TFP)

W podrozdz. 5.1 omówiono możliwości oszacowania łącznej produktywności czynników produkcji (TFP), będącej zmienną niemierzalną. W podrozdz. 5.2–5.4 zaprezentowano przykłady światowych badań empirycznych, podejmujących próby oszacowania i modelowania TFP. Badania takie prowadzone są też w Polsce. Wśród znaczących opracowań o charakterze empirycznym i teoretycznym wymienić należy następujące prace: Welfe [2000, 2001, 2003]; Tokarski [1996, 2002, 2008a]; Zienkowski [2003]. Poniżej przedstawiona zostanie własna próba oszacowania TFP według działów sekcji „przetwórstwo", a następnie zastosowania dynamicznego modelu panelowego do wskazania czynników wpływających na kształtowanie się tej wielkości.

Na potrzeby prezentowanych tu badań TFP oszacowano na podstawie modelu produkcji przy zastosowaniu dwóch alternatywnych podejść. Pierwsze z nich opiera się na propozycji T. Tokarskiego [2008b], i wymaga oszacowania parametrów funkcji wydajności. Drugie podejście opiera się na założeniu o doskonałej konkurencji na rynku czynników produkcji i założeniu, że elastyczność produkcji względem nakładu pracy jest stała w czasie, a może się różnić między działami. Zostało ono zaproponowane przez Ascari, Di Cosmo [2004] i jest szerzej omówione w podrozdz. 6.4.

5.6.1. Szacowanie TFP na podstawie modelu wydajności

W niniejszym punktcie wykorzystano metodę, zaproponowaną przez T. Tokarskiego [2008b] do wyznaczenia wartości TFP według województw. Polega ona na wyznaczeniu wartości TFP charakterystycznych dla poszczególnych obiektów (w niniejszym badaniu – działów) na podstawie funkcji produkcji Cobba–Douglasa postaci (5.5). Po dokonaniu prostych przekształceń otrzymuje się następującą zależność:

$$\ln\left(\frac{Y_{it}}{L_{it}}\right) = \ln A + gt + \alpha \ln\left(\frac{K_{it}}{L_{it}}\right) \tag{5.13}$$

Jest to zlogarytmowana funkcja wydajności pracy Cobba–Douglasa, gdzie: Y – wartość produkcji sprzedanej w mln zł[8]; L – nakłady pracy (wyrażone w tys. pracujących); K – nakłady kapitału rzeczowego (mierzone wartością środków trwałych brutto w mln zł); $Ae^{gt} > 0$ – łączna produktywność czynników

[8] Oszacowano również model, w którym za Y przyjęto wartość dodaną brutto w mln zł, ale wyniki były znacznie gorsze, dlatego nie są tu prezentowane.

produkcji (TFP); g – stopa postępu technicznego w sensie Hicksa; α – elastyczność zmiennej Y względem zmiennej K. Wartości zmiennych obserwowane są dla i-tego działu (i = 15, ..., 37) w roku t (t = 1, ..., 9). Dane statystyczne, wykorzystane w badaniu, pochodzą z „Roczników Statystycznych Przemysłu" z lat 1998–2007. Dotyczą one 22 działów sekcji D – „przetwórstwo", oznaczonych w klasyfikacji PKD numerami 15–37.

Z zależności (5.13) wynika, że wydajność Y_{it}/L_{it} zależy od technicznego uzbrojenia pracy K_{it}/L_{it} oraz od TFP, przy czym TFP (równe Ae^{gt}) jest jednakowe dla wszystkich działów i lat. Ponieważ założenie takie jest trudne do przyjęcia, do modelu (5.13) wprowadzono zmienne zero-jedynkowe dla każdego działu (jest to tzw. procedura dywersyfikacji stałej (por. Pindyck, Rubinfeld [1991])). Ostateczny model, na podstawie którego estymowano wartości TFP w działach, ma postać:

$$\ln\left(\frac{Y_{it}}{L_{it}}\right) = \sum_{i=15}^{37} \beta_i d_i + gt + \alpha \ln\left(\frac{K_{it}}{L_{it}}\right) + \varepsilon_{it} \tag{5.14}$$

gdzie d_i (i = 15, ..., 37) są zmiennymi zero-jedynkowymi, przyjmującymi wartość 1 dla i-tego działu.

Tabela 5.3

Wyniki estymacji modelu wydajności

Zmienna	Ocena parametru	Średni błąd szacunku	Istotność zmiennej (*p-value*)
ln(K/L)	0,26321	0,03787	0,0000
t	0,04899	0,00276	0,0000
d_{15}	3,72854	0,15603	0,0000
d_{16}	4,14847	0,21008	0,0000
d_{17}	3,12663	0,15126	0,0000
d_{18}	2,6696	0,09342	0,0000
d_{19}	2,94711	0,11283	0,0000
d_{20}	3,32181	0,14289	0,0000
d_{21}	3,78712	0,19199	0,0000
d_{22}	3,67515	0,15046	0,0000
d_{23}	4,60832	0,22596	0,0000
d_{24}	3,82393	0,19363	0,0000
d_{25}	3,62051	0,16181	0,0000
d_{26}	3,38837	0,17256	0,0000
d_{27}	3,77649	0,19518	0,0000

Tabela 5.3 (cd.)

Zmienna	Ocena parametru	Średni błąd szacunku	Istotność zmiennej (*p-value*)
d_{29}	3,38387	0,15818	0,0000
d_{28}	3,5099	0,13516	0,0000
d_{30}	4,316	0,16016	0,0000
d_{31}	3,63075	0,15624	0,0000
d_{32}	4,25593	0,16883	0,0000
d_{33}	3,39338	0,13642	0,0000
d_{34}	4,19379	0,19182	0,0000
d_{35}	3,39602	0,15159	0,0000
d_{36}	3,36845	0,12117	0,0000
d_{37}	3,55972	0,1508	0,0000
$R^2 = 0,9744$			

Ź r ó d ł o: opracowanie własne.

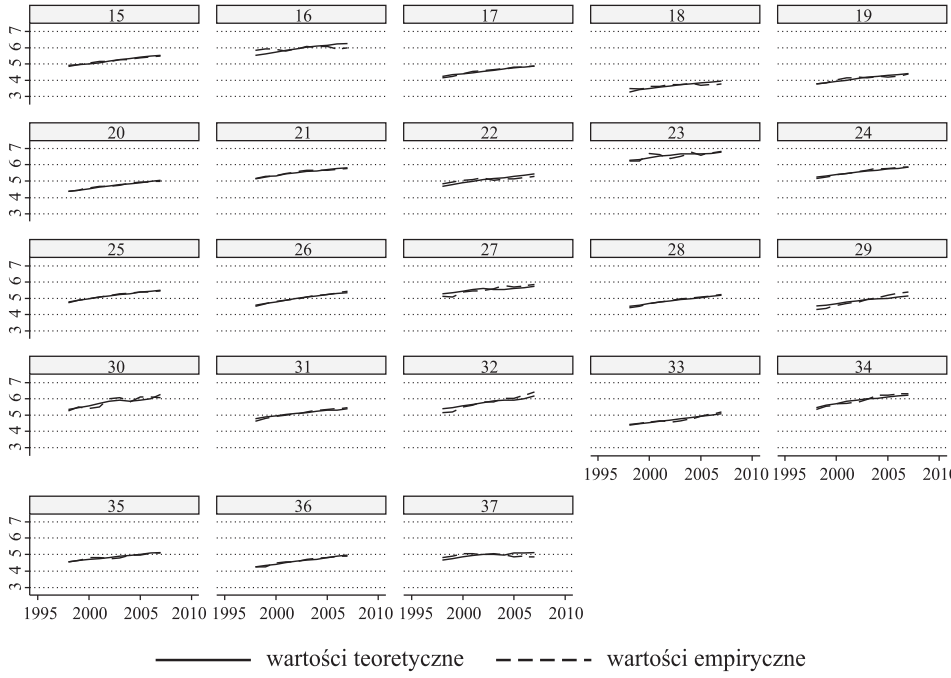
Do estymacji parametrów modelu (5.14) zastosowano uogólnioną metodę najmniejszych kwadratów (UMNK), gdyż ze względu na duże zróżnicowanie wydajności i technicznego uzbrojenia pracy między działami założono występowanie heteroskedastyczności grupowej. Uzyskane wyniki, prezentowane w tab. 5.3, wydają się zadowalające. Oszacowana stopa postępu technicznego w sensie Hicksa wynosi ok. 5%, a elastyczność wydajności pracy względem technicznego uzbrojenia pracy jest równa 0,26. Ponadto wszystkie zmienne są statystycznie istotne, a bardzo dobre dopasowanie modelu, mierzone współczynnikiem determinacji, pozwala uznać, że wartości TFP, które zostaną wyznaczone na podstawie modelu (5.14), będą wiarygodne. Dopasowanie wartości teoretycznych wydajności prezentowane jest na rys. 5.5.

Na podstawie wyników estymacji modelu (5.14) obliczono następnie wartości TFP w dziale *i* w roku *t* według wzoru:

$$TFP_{it}^{\text{tok}} = \frac{(Y_{it} / L_{it})}{(K_{it} / L_{it})^{0,26321}} \tag{5.15}$$

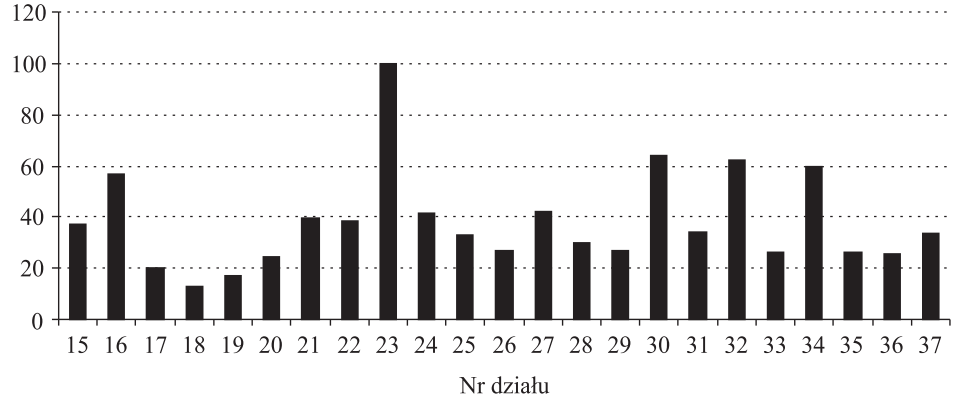
gdzie wartość wykładnika w mianowniku jest oceną parametru α modelu (5.14).

Zróżnicowanie łącznej produktywności czynników produkcji według działów, obliczonej według wzoru (5.15), przedstawione jest na rys. 5.6. Pokazane są tam przeciętne wartości zmiennej TFP^{tok} z lat 1998–2007 w poszczególnych działach. Ponieważ wartości tej cechy są niemianowane, przyjęto że przeciętne *TFP* dla działu 23: „produkcja koksu i produktów rafinacji ropy naftowej”, w którym poziom tej zmiennej jest najwyższy, równe jest 100.



Rys. 5.5. Wydajność pracy według działów sekcji D (przetwórstwo) – wartości empiryczne i teoretyczne na podstawie modelu (5.14)

Ź r ó d ł o: opracowanie własne.

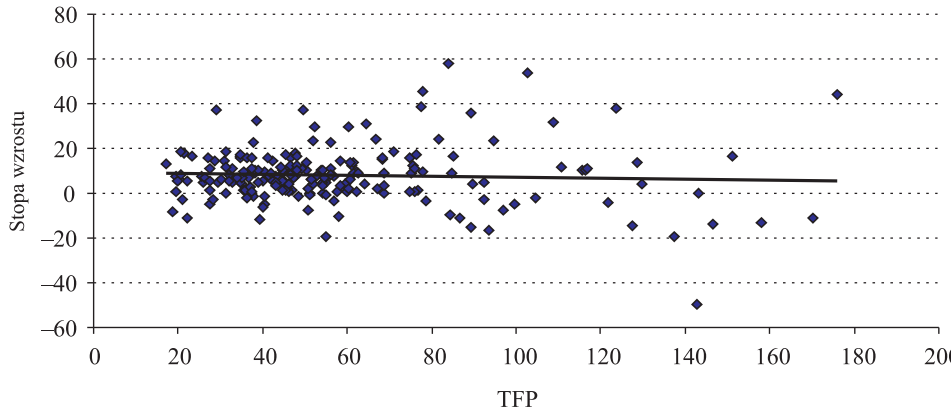


Rys. 5.6. Zróżnicowanie przeciętnych wartości TFP_{it}^{tok} według działów w latach 1998–2007 (dział 23 = 100)

Ź r ó d ł o: opracowanie własne.

Na podstawie rys. 5.6 stwierdzić można, że wysokimi wartościami TFP^{tok} (powyżej 55 w stosunku do działu 23) charakteryzują się działy: 30: „produkcja maszyn biurowych i komputerów", 32: „produkcja sprzętu i urządzeń radiowych, telewizyjnych i telekomunikacyjnych", 34: „produkcja pojazdów samochodowych, przyczep i naczep", 16: „produkcja wyrobów tytoniowych". Szczególnie niskie wartości TFP^{tok}, nieprzekraczające 20, stwierdzono zaś dla działów: 18: „produkcja odzieży i wyrobów futrzarskich", 19: „produkcja skór wyprawionych i wyrobów z nich", 17: „włókiennictwo".

W celu określenia, czy łączne produktywności charakterystyczne dla poszczególnych działów wykazują tendencję do wyrównywania się, obliczono współczynnik korelacji między roczną stopą wzrostu TFP (s*TEP*) a opóźnionym poziomem wartości tej cechy w działach (por. Tokarski [2008b]). Współczynnik ten ma wartość $r(TFP^{tok}_{i,t-1}, sTFP^{tok}_{it}) = -0{,}0519$. Ujemny znak wskazywałby na istnienie tendencji zbieżnej, a więc na zmniejszanie się różnic w łącznej produktywności między działami. Jednak wartość bezwzględna jest na tyle niska, że wnioskowanie na jej podstawie wydaje się pochopne. Z drugiej strony, w analogicznym badaniu, opisanym w pracy Dańskiej-Borsiak [2009], przeprowadzonym na podstawie próby o rok krótszej (z lat 1998–2006), uzyskano wartość $r(TFP_{i,t-1}, sTFP_{it}) = -0{,}049$. Istnieją więc być może przesłanki, by stwierdzić, że zjawisko konwergencji łącznej produktywności czynników produkcji na poziomie działów zaczyna się pojawiać. Byłaby to tendencja korzystna, gdyż oznaczałaby widoczny efekt działania postępu technicznego w tych działach, w których początkowo proces produkcji był mniej efektywny. Zależność między rocznymi stopami wzrostu TFP^{tok} a opóźnionym poziomem wartości tej zmiennej w działach w latach 1998–2007 przedstawiono na rys. 5.7.



Rys. 5.7. Stopy wzrostu TFP^{tok} (w %) w roku *t* a poziom TFP^{tok} w roku *t* – 1 w latach 1998–2007

Źródło: opracowanie własne.

5.6.2. Szacowanie TFP przy założeniu doskonałych rynków czynników produkcji

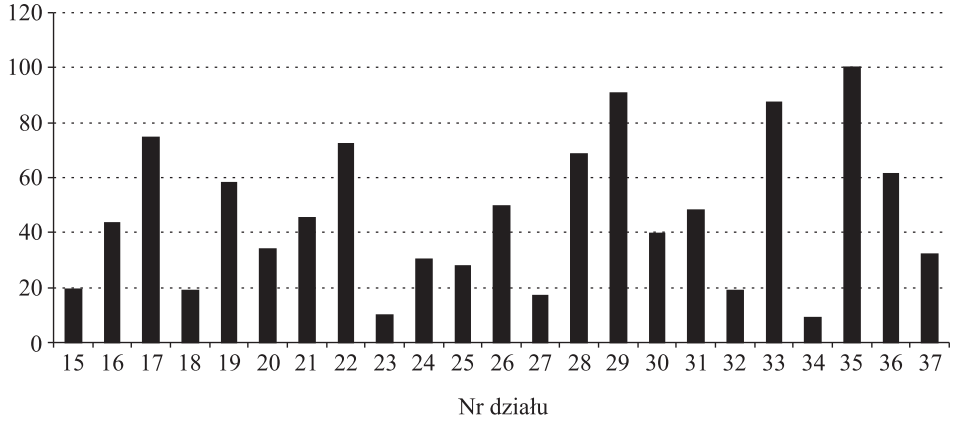
Kolejnym etapem badań jest próba alternatywnego oszacowania wartości łącznej produktywności czynników produkcji, przy zastosowaniu podejścia Ascari, Di Cosmo [2004]. Bazuje ono na funkcji produkcji Cobba–Douglasa i polega na wyznaczeniu oceny parametru α_i, będącego elastycznością produkcji względem nakładu pracy na podstawie zależności $\alpha_i = w_i \cdot L_i / Y_i$, gdzie w_i jest przeciętnym wynagrodzeniem, a L_i – przeciętnym zatrudnieniem w dziale i, zaś Y_i – wartością produkcji sprzedanej w dziale i. Wszystkie wartości są specyficzne dla działów oraz stałe w czasie i reprezentują średnie dla poszczególnych działów po czasie. Stałość w czasie α_i wynika z faktu, że jest to parametr technologiczny, który nie powinien się zmienić w okresie objętym próbą[9]. Po oszacowaniu wartości parametru α_i wyznaczono łączną produktywność czynników produkcji, korzystając z zależności:

$$TFP_{it}^{\mathrm{ac}} = \frac{Y_{it}}{(K_{it}^{1-\alpha_i} L_{it}^{\alpha_i})} \tag{5.16}$$

Zależność (5.16) powstaje z przekształcenia funkcji produkcji Cobba–Douglasa ze stałymi efektami skali. Metoda Ascari i Di Cosmo [2004] wykorzystuje więc tę samą funkcję, co metoda Tokarskiego [2008b]. Różnica polega na przyjętym na wstępie sposobie oszacowania elastyczności produkcji względem nakładu pracy.

Przeciętne wartości zmiennej TFP^{ac} z lat 1998–2007 w poszczególnych działach przedstawione są na rys. 5.8. Ponieważ wartości tej cechy są niemianowane, przyjęto, że przeciętne TFP dla działu 35: „koks, ropa naftowa”, w którym poziom tej zmiennej jest najwyższy, równy jest 100.

[9] Próba wykorzystana przez Ascari i Di Cosmo [2004] liczyła 15 lat, zatem założenie o stałości w czasie parametru α_i w okresie 8 lat objętych niniejszym badaniem wydaje się tym bardziej uzasadnione.



Rys. 5.8. Zróżnicowanie średnich wartości TFP^{ac} według działów w latach 1998–2007 (dział 35 = 100)

Ź r ó d ł o: opracowanie własne.

Porównanie wzrokowe rys. 5.8 i 5.6 pozwala zauważyć zasadnicze różnice między wartościami TFP oszacowanymi na podstawie wzoru (5.16) a wartościami oszacowanymi na podstawie wzoru (5.15). Najwyższymi wartościami zmiennej TFP^{ac} charakteryzują się działy: 35: „produkcja pozostałego sprzętu transportowego", 29: „produkcja maszyn i urządzeń", 33: „produkcja instrumentów medycznych, precyzyjnych i optycznych, zegarów i zegarków", 17: „włókiennictwo", a najniższymi – działy 34: „produkcja pojazdów samochodowych, przyczep i naczep", 23: „produkcja koksu i produktów rafinacji ropy naftowej", 27: „produkcja metali", i 32: „produkcja sprzętu i urządzeń radiowych, telewizyjnych i telekomunikacyjnych".

Wydaje się, że oszacowania TFP w poszczególnych działach uzyskane na podstawie wzoru (5.16) nie są poprawne. Znacznie bardziej prawdopodobne jest, że średnie wartości TFP w działach rozkładają się tak, jak na rys. 5.6. Efekty postępu technicznego powinny być bardziej widoczne w dynamicznie rozwijających się działach związanych z produkcją sprzętu komputerowego (30), telekomunikacyjnego (32) czy samochodowego (34), niż w działach związanych z włókiennictwem (17), optyką (33) czy pozostałym sprzętem transportowym (35). Przypuszczać można, że podstawową przyczyną niewiarygodności oszacowań TFP uzyskanych na podstawie wzoru (5.16) jest niespełnienie założenia o doskonałości rynków czynników produkcji. Jest ono poddawane w wątpliwość przez samych autorów, którzy określają je mianem „*a heroic assumption*" (A s c a r i, D i C o s m o [2004], s. 4).

Zasadnicze różnice między wartościami zmiennych TFP^{tok} i TFP^{ac} pozwalają przypuszczać, że wyniki dalszych badań, których celem będzie konstrukcja i estymacja modelu kształtowania się łącznej produktywności

czynników produkcji będą rozbieżne, w zależności od tego, czy zmienną objaśnianą modelu będzie TFP^{tok} czy TFP^{ac}. Porównanie wyników estymacji tych dwóch alternatywnych modeli umożliwi weryfikację postawionej powyżej hipotezy, że bardziej wiarygodne są oszacowania TFP na podstawie modelu wydajności.

5.7. Modelowanie kształtowania się łącznej produktywności czynników produkcji

Po oszacowaniu wartości TFP w działach podjęto próbę skonstruowania i oszacowania modelu ekonometrycznego, opisującego kształtowanie się tej zmiennej według działów. Mimo przedstawionych w pkt 5.6.2 zastrzeżeń do poprawności oszacowań TFP, uzyskanych na podstawie wzoru (5.16), skonstruowane i oszacowane zostały dwa alternatywne modele, w których zmienną objaśnianą było TFP^{tok} albo TFP^{ac}. Uzyskanie dobrego modelu, w którym zmienną objaśnianą było TFP^{ac}, okazało się jednak niemożliwe, mimo wielu prób modyfikacji zbioru zmiennych objaśniających, postaci funkcyjnej i założeń co do endogeniczności zmiennych objaśniających. Dlatego poniżej przedstawione zostaną wyniki estymacji modelu, w którym zmienną objaśnianą jest TFP^{tok}. W celu uproszczenia symboliki będzie ona oznaczana przez *TFP*.

Na podstawie opracowań o charakterze teoretycznym i empirycznym, znanych z literatury światowej (por. Cameron [2006]; Griffith, Redding, Van Reen [2003]; Coe, Helpman [1995]; Acharya, Keller [2007]), wytypowano czynniki, mogące wpływać na poziom TFP. Są to: kapitał ludzki, kapitał fizyczny, nakłady na działalność badawczo-rozwojową (BiR), wskaźnik unionizacji (odsetek pracowników fizycznych objętych umowami zbiorowymi pomnożony przez odsetek wszystkich pracowników fizycznych), stosunek eksportu do produkcji i stosunek importu do produkcji. Dwie ostatnie zmienne uwzględnione zostały w celu zbadania wpływu, wywieranego na produktywność przez transfer technologii, dokonujący się poprzez handel międzynarodowy. Oszacowanych zostało kilkanaście modeli, w których w roli zmiennych objaśniających uwzględniano alternatywne miary wymienionych powyżej czynników, mogących mieć wpływ na produktywność. Testowano następujące zmienne objaśniające:

- dla określenia kapitału ludzkiego:

hk – stosunek liczby zatrudnionych na stanowiskach nierobotniczych do liczby zatrudnionych ogółem,

zbrlp – stosunek liczby zatrudnionych w działalności BiR do liczby pracujących ogółem;

- dla określenia kapitału fizycznego:

stbh – wartość środków trwałych brutto na jedna przepracowaną godzinę (w zł/godz.);

- dla określenia wpływu transferu technologii:

imps – stosunek wartości importu do wartości produkcji sprzedanej,

ekps – stosunek wartości eksportu do wartości produkcji sprzedanej;

- dla określenia wpływu działalności badawczo-rozwojowej:

brinw – stosunek nakładów na BiR do nakładów inwestycyjnych.

Konstrukcja wskaźnika unionizacji okazała się niemożliwa, ponieważ brak jest niezbędnych danych statystycznych. Wymienione powyżej zmienne wprowadzane były do modelu albo jako wartości bieżące, albo opóźnione o 1 lub 2 okresy. Dodatkowo konieczne okazało się wprowadzenie do modelu zmiennych zero-jedynkowych dla wybranych lat. Zmienne zero-jedynkowe dla lat 2000, 2001, ..., 2007 oznaczone zostały symbolami *r*00, *r*01, ..., *r*07.

W związku z niedostępnością części danych, dotyczących działalności badawczo-rozwojowej (brak danych o nakładach i zatrudnieniu dla działów 16, 18, 19, 21, 37) oraz danych o imporcie i eksporcie dla wszystkich działów w latach 1998–1999, zdecydowano się zawęzić próbę. Analiza kształtowania się TFP przeprowadzona została zatem na podstawie danych dla 18 działów (*i* = 15, 17, 20, 22, ..., 36) z lat 2000–2007.

Ostateczna postać modelu, który charakteryzował się najlepszymi własnościami merytoryczno-statystycznymi jest następująca:

$$\ln(TFP_{it}) = \alpha_0 + \gamma \ln(TFP_{i,t-1}) + \beta_1 \ln(hk_{it}) + \beta_2 \ln(brinw_{i,t-2}) + \beta_3 \ln(stbh_{it}) + \\ + \beta_4 \ln(imps_{i,t-1}) + \beta_5 r02 + \beta_6 r03 + \beta_7 r05 + (\alpha_i + \varepsilon_{it}) \qquad (5.17)$$

Do estymacji modelu (5.17) stosowane były FDGMM i SGMM. Zmienne objaśniające (wszystkie lub część z nich) traktowane były alternatywnie jako zmienne ściśle egzogeniczne, słabo egzogeniczne lub endogeniczne. W wyniku wielu prób uznano ostatecznie, że zmiennymi nieegzogenicznymi są zmienne *stbh* (wartość środków trwałych brutto na jedną przepracowaną godzinę) i *hk* (stosunek liczby zatrudnionych na stanowiskach nierobotniczych do liczby zatrudnionych ogółem). W tab. 5.4 przedstawione są wybrane wyniki estymacji modelu (5.17) uzyskane przy zastosowaniu wspomnianych powyżej metod. Oznaczenia „endo", „pre" i „egzo" przy nazwach metod SGMM i FDGMM określają, jak traktowane były zmienne *stbh* i *hk*: endogenicznie, słabo egzogenicznie lub ściśle egzogenicznie. Prezentowane wyniki uzyskano na podstawie estymatorów jednostopniowych. Zastosowanie metod dwustopniowych nie dało zadowalających wyników. Symbolem WG oznaczono estymator wewnątrzgrupowy.

Tabela 5.4

Wyniki estymacji modeli kształtowania się TFP

Zmienna objaśniająca/ test statystyczny	Metoda estymacji					
	SGMMendo	SGMMpre	SGMMegzo	FDGMMendo	WG	KMNK
	1	2	3	4	5	6
$\ln(TFP_{i,t-1})$	0,6725*** (0,0849)	0,6699*** (0,0685)	0,7101*** (0,1030)	0,4749*** (0,1046)	0,5045*** (0,0839)	0,9342*** (0,0351)
$\ln(hk_{it})$	0,2235*** (0,1052)	0,1926*** (0,0744)	0,0705 (0,1494)	0,3459** (0,2163)	0,3363*** (0,1609)	0,0112 (0,0293)
$\ln(brinw_{i,t-2})$	0,0308*** (0,0106)	0,0101 (0,0249)	0,0208 (0,0275)	0,0314 (0,0271)	0,02788 (0,0181)	0,0104 (0,0085)
$\ln(stbh_{it})$	0,0530** (0,0297)	0,0372 (0,0438)	0,0131 (0,0506)	–0,0119 (0,0648)	–0,0122 (0,0565)	0,0377** (0,0213)
$\ln(imps_{i,t-1})$	0,1164*** (0,0462)	0,0675** (0,0445)	0,1135 (0,0798)	0,0319 (0,0337)	0,0156 (0,0739)	0,0011 (0,0201)
*r*02	–0,1366*** (0,0361)	–0,1397*** (0,0298)	–0,1413*** (0,0331)	–0,2102*** (0,0378)	–0,2036*** (0,0328)	–0,0914*** (0,0264)
*r*03	–0,0889*** (0,0253)	–0,1025*** (0,0262)	–0,0806*** (0,0336)	–0,1367*** (0,0310)	–0,1299*** (0,0299)	–0,0333 (0,0262)
*r*05	–0,0609** (0,0327)	–0,0814*** (0,0285)	–0,0855*** (0,0297)	–0,0675*** (0,0225)	–0,0676*** (0,0231)	–0,0710*** (0,0254)
const.	1,8583*** (0,5178)	1,7213*** (0,3641)	1,5738*** (0,5416)	2,9839*** (0,4545)	2,8153*** (0,5085)	0,2768** (0,1535)
m_1 [*p-value*]	–2,247 [0,025]	–2,503 [0,012]	–2,093 [0,036]	–2,497 [0,013]	–	–

Tabela 5.4 (cd.)

Zmienna objaśniająca/ test statystyczny	Metoda estymacji					
	SGMM endo	SGMM pre	SGMM egzo	FDGMM endo	WG	KMNK
	1	2	3	4	5	6
m_2 [*p-value*]	–0,889 [0,374]	–0,138 [0,890]	–0,143 [0,886]	0,664 [0,506]	–	–
Sargan [*p-value*]	92,368 [0,112]	102,331 [0,097]	36,753 [0,061]	70,227 [0,150]	–	–
dif-Sargan [*p-value*]	22,141 [0,226]	24,072 [0,153]	1,527 [0,958]	–	–	–
R^2	–	–	–	–	0,824	0,958

Symbol *** przy ocenie parametru oznacza, że wartość $p < 0,05$ dla testu *t*-Studenta; symbol ** oznacza $p \leq 0,1$; pozostałych zmiennych nie można uznać za istotne;

w nawiasach, pod ocenami parametrów zamieszczono błędy szacunku parametrów, wyznaczane na podstawie odpornych estymatorów wariancji składnika losowego;

m_1 i m_2 oznaczają wartości empiryczne testu autokorelacji Arellano–Bonda, odpowiednio AR(1) i AR(2); w nawiasach kwadratowych – wartości *p* dla H_0: autokorelacja pierwszego (drugiego) rzędu nie występuje;

Sargan i *dif*-Sargan oznaczają wartości empiryczne testu Sargana i różnicowego testu Sargana; w nawiasach kwadratowych – wartości *p* dla H_0: warunki ponadidentyfikujące są poprawne (instrumenty są właściwe).

Ź r ó d ł o: obliczenia własne.

Wyniki wszystkich testów statystycznych[10] zawartych w tab. 5.4 jednoznacznie wskazują na poprawność instrumentów, wykorzystanych w GMM. Wniosek taki dotyczy obu zastosowanych metod estymacji: FDGMM i SGMM i jest niezależny od tego, czy zmienne objaśniające są endogeniczne, słabo egzogeniczne czy ściśle egzogeniczne. Wartości statystyk m_1 i m_2 testu Arellano–Bonda wskazują na występowanie w modelu pierwszych różnic autokorelacji pierwszego rzędu oraz brak autokorelacji drugiego rzędu. Występowanie w tym modelu autokorelacji pierwszego rzędu jest zjawiskiem spodziewanym, gdyż jeśli ε_{it} są niezależne, to ich pierwsze różnice są skorelowane rzędu 1. Występowanie autokorelacji rzędu wyższego niż 1 oznaczałoby jednak, że warunki momentów są niespełnione, lub inaczej, że instrumenty użyte podczas estymacji GMM nie są właściwe.

Drugą możliwością weryfikacji jakości modelu jest zastosowanie testu Sargana. Za jego pomocą bada się prawdziwość warunków ponadidentyfikujących, niewykorzystanych w procesie estymacji. W hipotezie zerowej zakłada się, że wykorzystane instrumenty są właściwe, tzn. nieskorelowane ze składnikami losowymi modelu pierwszych różnic. Błędy szacunku parametrów omawianego modelu wyznaczane były na podstawie odpornych estymatorów macierzy wariancji-kowariancji składnika losowego. W takim przypadku rozkład statystyki empirycznej testu Sargana jest nieznany. Dlatego w tab. 5.4 prezentowane są wartości statystyk testu Sargana, wyznaczone dla modeli, w których błędy szacunku liczone były standardowo. Wydaje się jednak, że wartości te są miarodajne dla omawianych modeli, a postępowanie takie stosowane jest w wielu publikacjach światowych, np.: Coviello, Islam [2006]; Di Liberto, Mura, Pigliaru [2004]; Casu, Girardone [2007]. Wartości statystyk testu Sargana dla wszystkich wariantów GMM, prezentowanych w kolumnach 1–4 tab. 5.4 nie dają podstaw do odrzucenia H_0 na poziomie istotności 0,05. Oznacza to, że warunki ponadidentyfikujące są poprawne. W kolumnach 1–3 prezentowane są ponadto statystyki różnicowego testu Sargana, który pozwala rozstrzygnąć, czy zastosowanie SGMM prowadzi do bardziej poprawnych wyników niż FDGMM. Dokładniej, test ten bada, czy dodatkowe warunki momentów, wykorzystane w SGMM, są spełnione. Uzyskane wartości jednoznacznie wskazują, że stosowanie SGMM jest w pełni uzasadnione.

Dodatkową możliwość sprawdzenia, czy oceny parametrów uzyskane na podstawie FDGMM lub SGMM są nieobciążone, daje porównanie ich z ocenami wyznaczonymi na podstawie estymatora wewnątrzgrupowego (WG) i estymatora KMNK modelu *pooled*[11]. W podrozdz. 3.1 przytoczone są wyniki uzyskane

---

[10] Konstrukcja stosowanych w prezentowanym badaniu testów statystycznych została omówiona w podrozdz. 3.3.

[11] Estymator wewnątrzgrupowy jest stosowany do szacowania modeli statycznych, w których efekty grupowe α_i są nielosowe (modeli FEM). Model *pooled* to model szacowany na podstawie danych panelowych, w którym jednak nie wyróżnia się efektów grupowych ani czasowych i zakłada się, że macierz wariancji-kowariancji składnika losowego jest sferyczna.

przez N i c k e l l a [1981], który wykazał, że estymator WG parametru autoregresyjnego γ jest, przy ustalonym T, obciążony w dół, a estymator KMNK jest obciążony w górę (np. H s i a o [2003], s. 73–74). Wartość zgodnego estymatora parametru γ powinna zawierać się zatem pomiędzy wartościami tych dwóch estymatorów. Z porównania ocen parametru autoregresyjnego, prezentowanych w kolumnach 1–6 tab. 5.4, wynika, że ocena FDGMM jest obciążona, ponieważ jej wartość (0,4749) jest niższa niż ocena WG (0,5045). Oceny SGMM, niezależnie od wariantu, zawierają się pomiędzy oceną WG a oceną KMNK (0,9342), co świadczy o tym, że są one zgodne.

Oceny SGMM parametru γ są zbliżone, jednak biorąc pod uwagę efektywność estymatorów parametrów uznać należy, że najlepszą metodą estymacji modelu (5.17) jest SGMM z założeniem endogeniczności zmiennych *stbh* i *hk* (kolumna 1). Dalsza analiza dotyczyć będzie wyników uzyskanych tą metodą.

Najsilniejszy wpływ na wzrost TFP wywiera wzrost kapitału ludzkiego (elastyczność 0,22%). Bardzo znaczący jest również wpływ transferu technologii na TFP, dokonujący się poprzez wymianę międzynarodową. Elastyczność zmiennej *imps* względem TFP wynosi 0,12%. Nieco mniejsza jest siła wpływu kapitału fizycznego (elastyczność 0,05%) i nakładów na BiR (elastyczność 0,03%). Oddziaływanie transferu technologii i nakładów na BiR jest odroczone w czasie, przy czym długość tego opóźnienia wynosi jeden rok w przypadku zmiennej *imps*, a w przypadku zmiennej *brinw* – nawet 2 lata.

Mnożniki długookresowe, obliczone według wzoru $n_k = \frac{\hat{\beta}_k}{1-\hat{\gamma}}$, gdzie $\hat{\beta}_k$ jest oceną parametru przy zmiennej x_k a $\hat{\gamma}$ – oceną parametru autoregresyjnego, wynoszą: 0,6824 dla zmiennej *hk*, 0,3554 dla zmiennej *imps*, 0,1618 dla zmiennej *stbh* i 0,0940 dla zmiennej *brinw*. Określają one zmianę TFP pod wpływem jednostkowego przyrostu wszystkich wartości bieżących i opóźnionych zmiennej objaśniającej.

Wysoka wartość oceny parametru autoregresyjnego modelu (5.17), równa 0,673, świadczy o stabilności łącznej produktywności czynników produkcji w działach. Wniosek taki jest zgodny z konkluzją, wynikającą z przeprowadzonej we wcześniejszej części rozdziału analizy współczynnika korelacji między roczną stopą wzrostu TFP a opóźnionym poziomem wartości tej zmiennej. Należy uznać, że wartości TFP w poszczególnych działach są silnie zróżnicowane i nie wykazują widocznej tendencji do wyrównywania się.

Ostatecznie uznać należy, że metoda zaproponowana przez T. T o k a r s k i e g o [2008b] pozwoliła na wiarygodne oszacowanie łącznej produktywności czynników produkcji. W konsekwencji możliwe było skonstruowanie i poprawne oszacowanie modelu kształtowania się TFP, a co za tym idzie, wskazanie i zmierzenie siły wpływu poszczególnych czynników na łączną produktywność czynników produkcji.

ROZDZIAŁ 6

ZASTOSOWANIE DYNAMICZNYCH MODELI PANELOWYCH W BADANIACH REGIONALNYCH

Jednym z podstawowych problemów współczesnej gospodarki staje się przestrzenne zróżnicowanie rozwoju i tempa wzrostu gospodarek. W kontekście rozważań prowadzonych poniżej, podkreślić trzeba różnice między tymi dwoma pojęciami. Wzrost gospodarczy mierzony jest jako wzrost produktu na jednego mieszkańca. Rozwój gospodarczy jest pojęciem szerszym i oznacza poprawę różnych, w tym niemierzalnych aspektów poziomu życia (por. Bartkowiak [2003]. Wzrost jest więc zmianą ilościową, zaś rozwój obejmuje zarówno zmiany ilościowe, jak i jakościowe (por. Chojnicki, Czyż [2004]). Ze względu na występowanie wzajemnych zależności między wzrostem a rozwojem, przyjmuje się, że wzrost jest podstawowym warunkiem wystąpienia rozwoju gospodarczego.

Proces integracji europejskiej i realizacja unijnych programów spójności napotyka na wiele problemów wynikających z dysproporcji rozwoju gospodarczego integrujących się regionów. Głównym celem działań podejmowanych w ramach polityki regionalnej jest zatem wyrównywanie tych dysproporcji ze szczególnym uwzględnieniem stymulowania rozwoju w regionach opóźnionych. W realizacji tego celu dużą rolę odgrywają ekonometryczne analizy gospodarek regionalnych, których liczba w ostatnim okresie znacząco rośnie.

Ekonometryczne modele regionalne, w tym modele wzrostu i konwergencji, coraz częściej są konstruowane i estymowane na podstawie danych panelowych (lub szerzej – przekrojowo-czasowych) albo metodami ekonometrii przestrzennej, z zastosowaniem macierzy wag, odzwierciedlającej siłę powiązań między regionami.

W niniejszym rozdziale prezentowane są przykłady zastosowania dynamicznych modeli panelowych do badania wzrostu, konwergencji i produktywności w regionach. Są to zarówno przykłady zaczerpnięte z literatury światowej (podrozdz. 6.1–6.4 oraz 6.6), jak i analizy własne (podrozdz. 6.5). Tabela 6.1 zawiera zestawienie światowych badań dotyczących gospodarek regionalnych prowadzonych na podstawie dynamicznych modeli panelowych w trzech głównych nurtach: wzrostu, konwergencji i produktywności. Przykłady badań przeprowadzonych przez polskich autorów wymienione są we wstępie do podrozdz. 6.5.

Tabela 6.1

Przykłady światowych analiz regionalnych z zastosowaniem dynamicznych modeli panelowych

Problem badawczy	Opracowania
Modelowanie wzrostu	**Islam [1995]**; De la Fuente, Doménech [2000]; Bassanini, Scarpetta, Hemmings [2001]; Lee [2009]
Konwergencja regionalna	**Caselli, Esquivel, Lefort [1996]**; **Bond, Hoeffler, Temple [2001]**; Soukiazis, Castro [2004]
Analizy produkcji i produktywności	**Ascari, Di Cosmo [2004]**; **Byrne, Fazio, Piacentino [2005]**; Guellec, van Pottelsberghe de la Potterie [2001]; Koop, Osiewalski, Steel [2000]

Uwaga: opracowania wyróżnione czcionką półgrubą omawiane są w niniejszym rozdziale.

Źródło: opracowanie własne.

6.1. Podstawy ekonomiczne teorii wzrostu i konwergencji

Według neoklasycznej teorii wzrostu zakłada się, że długookresowy wzrost produkcji możliwy jest tylko dzięki zwiększeniu produkcji potencjalnej. Determinantami wzrostu długookresowego są: akumulacja kapitału rzeczowego, wzrost liczby ludności i związana z nim wielkość zatrudnienia, oraz postęp techniczny (o charakterze egzogenicznym).

W empirycznych badaniach wzrostu gospodarczego wykorzystywana jest stosunkowo często funkcja produkcji Cobba–Douglasa postaci: $Y = A \cdot F(K, L)$. Symbolem Y oznacza się ogólnie wartość produkcji; w modelach wzrostu i konwergencji przyjmuje się zazwyczaj, że jest to poziom PKB *per capita*. A jest łączną produktywnością czynników produkcji (TFP), K jest miarą nakładu kapitału, L – miarą nakładu pracy. Przyjmuje się zazwyczaj, że model ekonometryczny, bazujący na funkcji Cobba–Douglasa ma postać potęgową:

$$Y = AK^{\alpha_1} L^{\alpha_2} \mathrm{e}^{\varepsilon_t} \tag{6.1}$$

Uwzględnia się przy tym założenia tzw. neoklasycznego modelu Solowa, przedstawione w pracy: R. Solow [1957]. Zgodnie z nimi funkcja produkcji charakteryzuje się:

– stałymi korzyściami skali: $F(\lambda K, \lambda L) = \lambda F(K, L)$,

– dodatnimi, malejącymi krańcowymi produktywnościami kapitału i pracy dla $K > 0$ i $L > 0$,

– spełnieniem tzw. warunków Inady: krańcowy produkt nakładu zbliża się do nieskończoności (zera), gdy jego wielkość zbliża się do zera (nieskończoności),

– ograniczoną substytucyjnością czynników produkcji.

R. Solow, wykorzystując powyższe założenia, stwierdził, że jeśli klasyczne czynniki produkcji są opłacane według ich produktów krańcowych, a funkcja produkcji charakteryzuje się stałymi efektami skali, to stopa postępu technicznego (nazwana później resztą Solowa) jest różnicą miedzy stopą wzrostu wydajności pracy a stopą wzrostu technicznego uzbrojenia pracy ważoną udziałem nakładów kapitału w produkcie. Model Solowa podlegał wielu modyfikacjom, wśród których szczególnie istotne w kontekście przedstawianych tu analiz są modele wzrostu endogenicznego. Ideę endogenizacji postępu technicznego wysunął w 1966 r. K. Shell. Z modelu Shella wynika, że ponieważ proces akumulacji wiedzy naukowo-technicznej ma charakter endogeniczny, to gospodarka może skutecznie podnieść stopy wzrostu gospodarczego poprzez podnoszenie stóp inwestycji w zasób kapitału i wiedzę naukowo-techniczną (por. Welfe [2001], s. 17). Dynamiczny wzrost liczby prac, dotyczących modelowania wzrostu datuje się od połowy lat 80. XX w., a pionierskimi pracami z tej dziedziny są artykuły Abramovitza [1986] i Baumola [1986]. Jednak podstawowymi modelami teorii wzrostu endogenicznego są modele Romera [1986, 1990], Lucasa [1988, 1990], Barro [1990, 1991]. Szczególnie istotna jest w nich endogenizacja postępu technicznego, którego źródeł szuka się po stronie akumulacji wiedzy naukowo-technicznej lub kapitału ludzkiego. Więcej informacji na temat modeli wzrostu endogenicznego znaleźć można np. w pracach: Welfe [2001, 2007]; Barro [1997].

Opierając się na neoklasycznym modelu wzrostu gospodarczego Solowa, R. J. Barro i X. Sala-i-Martin [1992] wysunęli koncepcję konwergencji warunkowej, czyli wyrównywania poziomu PKB między krajami lub regionami o zróżnicowanych początkowych wartościach tej zmiennej. Warunkiem koniecznym zaistnienia konwergencji między regionami są malejące krańcowe produktywności kapitału. Jeśli jest on spełniony, to kapitał napływa do regionów, w których jego zasób na pracującego (techniczne uzbrojenie pracy) jest niższy, a zatem produkt krańcowy – wyższy. Prowadzi to do wyrównywania się poziomu kapitału na pracującego. O ile poziom zaawansowania technologicznego regionów jest zbliżony, to konwergencja technicznego uzbrojenia pracy przekłada się na konwergencję wydajności. Gdyby założenie malejących krańcowych produktywności kapitału nie było spełnione, to kapitał koncentrowałby się w regionach, w których techniczne uzbrojenie pracy, a zatem i produkt krańcowy, są wyższe. Regiony te rozwijałyby się szybciej od pozostałych, co prowadziłoby do zjawiska dywergencji gospodarczej. Czynnikami, wpływającymi na występowanie konwergencji lub dywergencji są: regionalne różnice w generowaniu i adaptowaniu nowych technologii, migracje, przekształcenia strukturalne, poziom infrastruktury. Dokładniejsza analiza tych czynników zawarta jest w pracy P. Gajewskiego [2006].

W literaturze formułowane są hipotezy o konwergencji warunkowej i bezwarunkowej typu β lub typu σ[1]. Pojęcia te definiuje się następująco:

- konwergencja typu σ oznacza postępujące w czasie zmniejszanie się przestrzennego zróżnicowania dochodów *per capita*, mierzonego najczęściej odchyleniem standardowym;
- konwergencja typu β oznacza, że kraje o niższym początkowym poziomie dochodu *per capita* rozwijają się szybciej niż kraje początkowo wyżej rozwinięte. Występuje wówczas tzw. efekt doganiania, co prowadzi do wyrównywania się poziomów dochodu *per capita*.

Konwergencja typu σ oznacza więc zmianę rozkładu dochodu w czasie, a konwergencja typu β dotyczy mobilności dochodu między regionami w ramach tego samego rozkładu. Ponadto konwergencja typu β jest warunkiem koniecznym, ale niewystarczającym, do zaistnienia konwergencji typu σ. Rozróżnia się przy tym β-konwergencję bezwarunkową i warunkową:

– konwergencja bezwarunkowa (absolutna) oznacza, że regiony dążą do jednakowego poziomu PKB *per capita* niezależnie od warunków początkowych. Zakłada się więc podobieństwo podstawowych parametrów opisujących gospodarki różnych regionów;

– konwergencja warunkowa oznacza, że każdy z badanych obiektów zdąża do właściwego dla siebie stanu wzrostu zrównoważonego[2], zależnego od wyjściowych cech jego gospodarki. Hipoteza ta zakłada więc zróżnicowanie podstawowych wskaźników opisujących gospodarki różnych regionów. Jak już wspomniano, hipoteza konwergencji warunkowej jest istotnym wnioskiem, wynikającym z neoklasycznych modeli wzrostu.

Znaczna część współcześnie prowadzonych badań empirycznych nad wzrostem i konwergencją wykorzystuje dynamiczne modele panelowe. Obszerne uzasadnienie przyczyn tego stanu rzeczy zawarte jest w pracy Bonda, Hoeffler i Temple'a [2001]. Jak stwierdzają autorzy, zmienne objaśniające w modelach wzrostu są zazwyczaj zmiennymi endogenicznymi (np. stopa inwestycji), podlegającymi błędom pomiaru. Istotne znaczenie ma też problem pominiętych zmiennych[3]. Taką zmienną w modelach konwergencji jest początkowy poziom efektywności, który jest nieobserwowalny i skorelowany z początkowym poziomem dochodu, będącym jedną ze zmiennych objaśniających. Skutkiem pominięcia tej zmiennej w modelu konwergencji jest zatem obciążenie estymatora KMNK. Zastosowanie dynamicznych modeli panelowych, estymowanych metodami bazującymi na GMM, daje możliwość rozwiązania tych

---

[1] Więcej informacji na temat konwergencji znaleźć można m. in. w: Barro, Sala-i-Martin [2003]; De la Fuente [2000]; Malaga, Kliber [2007].

[2] Stan wzrostu zrównoważonego (ang. *steady state*) to sytuacja, w której wszystkie zmienne mają identyczną stopę wzrostu: $\Delta Y/Y = \Delta L/L = \Delta K/K = const.$

[3] O możliwościach rozwiązania problemów błędów pomiaru i pominiętych zmiennych, jakie daje zastosowanie panelowych modeli dynamicznych, napisano w podrozdz. 3.5.

problemów. Po pierwsze, obliczanie pierwszych różnic eliminuje z modelu zmienne stałe w czasie, m. in. początkowy poziom efektywności. Po drugie, wykorzystanie zmiennych instrumentalnych daje możliwość uzyskania zgodnych estymatorów parametrów przy endogenicznych zmiennych objaśniających, oraz rozwiązania problemów wynikających z błędów pomiaru. Zastosowanie systemowej GMM (SGMM) R. Blundella i S. Bonda daje dodatkowo szansę uzyskania estymatorów nieobciążonych w przypadku, gdy szeregi czasowe charakteryzują się wysokim stopniem trwałości (ang. *highly persistent time series*) (por. podrozdz. 3.2.2).

6.2. Modelowanie wzrostu gospodarczego w latach 90. XX w.

Wzmożone zainteresowanie konstrukcją i wynikami estymacji modeli wzrostu i konwergencji datuje się od początku lat 90. XX w., a jako przykłady znaczących prac z tego zakresu wymienić można prace: Barro, Sala-i-Martin [1992]; Mankiw, Romer, Weil [1992]. Klasyczne modele wzrostu, oparte na pracy Solowa, zostały w późniejszych latach podważone przez modele bazujące na teorii wzrostu endogenicznego. Generalnie uważa się, że stwierdzenie występowania zjawiska konwergencji jest potwierdzeniem słuszności założeń Solowa, a stwierdzenie, że konwergencja nie występuje przemawia za poprawnością teorii wzrostu endogenicznego. Wspólną cechą badań empirycznych nad wzrostem gospodarczym, które bazują na danych przekrojowych jest założenie jednakowej zagregowanej funkcji produkcji dla wszystkich krajów. Do przyjęcia takiego założenia zmusza brak możliwości uwzględnienia w modelach danych przekrojowych niemierzalnych różnic między krajami. Jedną z pierwszych prac, które wykorzystują dane panelowe do estymacji dynamicznego modelu wzrostu jest praca N. Islama [1995]. Jak zauważa autor, wykorzystanie danych panelowych daje możliwość uwzględnienia wspomnianych różnic w konstrukcji funkcji produkcji, poprzez wyodrębnienie efektów grupowych. Jednym z celów Islama [1995] było porównanie wyników uzyskanych na podstawie danych panelowych z wynikami uzyskanymi na podstawie regresji przekrojowej przez N. Mankiwa, D. Romera i D. Weila [1992]. Dlatego wykorzystał on analogiczne trzy zbiory krajów (98, 75 i 22 kraje). Ponieważ konstrukcja panelu na podstawie danych rocznych nie eliminowałaby wpływu zakłóceń krótkookresowych, autor wykorzystał dane z okresów 5-letnich z lat 1960–1985. Estymacji dokonano czterema metodami: KMNK na podstawie danych przekrojowych, KMNK na podstawie danych panelowych (ang. *pooled regression*), metodą minimalnej odległości[4] (ang.

[4] Metoda minimalnej odległości została opracowana przez Chamberlaina [1982]. Zamiast eliminowania efektów grupowych α_i poprzez wyznaczenie pierwszych różnic, specyfikuje się α_i jako funkcję tych zmiennych, z którymi może ona być skorelowana.

minimum distance) Chamberlaina [1982], na podstawie modelu FE (estymator WG). Islam stwierdził, że tempo konwergencji wyznaczone na podstawie modeli panelowych jest znacznie szybsze niż wyznaczone na podstawie regresji przekrojowej, a oszacowane wartości elastyczności produkcji względem kapitału – niższe i bardziej akceptowalne. Z punktu widzenia teorii wzrostu, wykorzystanie modeli panelowych daje możliwość wyodrębnienia wpływu różnic technologicznych i instytucjonalnych między krajami na wzrost gospodarczy i na proces konwergencji.

W świetle późniejszej wiedzy, a szczególnie prac Arellano i Bonda [1991] oraz Blundella i Bonda [1998], stwierdzić trzeba, że wyniki uzyskane przez Islama [1995] są obciążone w górę, co jest skutkiem zastosowania estymatora WG do estymacji modelu dynamicznego. Istotnie, tempo konwergencji, wyznaczone dla najmniejszej grupy krajów na podstawie estymatora WG, wynoszące aż 9%, wydaje się zbyt szybkie. Ogólnie stwierdzić trzeba, że estymacja dynamicznych modeli wzrostu wymaga zastosowania właściwych metod, ze względu na występowanie problemów związanych z endogenicznością zmiennych objaśniających, błędami pomiaru zmiennych lub pominiętymi zmiennymi.

Autorami, którzy pierwsi zastosowali FDGMM do estymacji modelu wzrostu na podstawie panelowego modelu dynamicznego, byli F. Caselli, G. Esquivel i F. Lefort [1996]. Wykorzystali oni dane dotyczące 97 krajów z lat 1960–1985 (średnie z okresów 5-letnich). Kryterium doboru obiektów była dostępność danych. Jak zauważyli autorzy, wykorzystanie do estymacji modeli wzrostu danych panelowych i zastosowanie GMM ma przewagę nad regresją przekrojową i innymi metodami estymacji. Po pierwsze, dzięki wykorzystaniu danych panelowych można oszacować model na przyrostach, co pozwala uniknąć obciążenia estymatorów wynikającego z pominięcia zmiennych objaśniających stałych w czasie (stanowią one składową efektu grupowego i są eliminowane podczas obliczania pierwszych różnic). W szczególności, w warunkowych równaniach konwergencji pominiętą zmienną może być początkowy poziom efektywności, który jest nieobserwowalny, a dodatkowo skorelowany z początkowym poziomem dochodu, będącym zazwyczaj jedną ze zmiennych objaśniających. Po drugie, wykorzystanie w FDGMM i SGMM zmiennych instrumentalnych pozwala na wyznaczenie zgodnych estymatorów parametrów w modelach zawierających endogeniczne zmienne objaśniające takie, jak np. stopa inwestycji, a także w przypadku obecności błędów pomiaru zmiennych.

6.3. Współczesne modelowanie wzrostu gospodarczego

S. Bond, A. Hoeffler i J. Temple [2001] zwrócili uwagę na to, że zastosowanie FDGMM do estymacji modeli wzrostu może dać niepoprawne wyniki, ze względu na problem słabych instrumentów. Dane dotyczące produk-

cji mają często cechy procesu o wysokim stopniu trwałości, a ponadto panelowe modele wzrostu estymowane są często na podstawie kilku obserwacji w czasie, które dodatkowo mogą być średnimi z kilku lat. Te cechy powodują, że opóźnione poziomy zmiennych są słabymi instrumentami dla równań na przyrostach, co skutkuje obciążeniem estymatorów.

Bond, Hoeffler i Temple [2001] reestymowali za pomocą SGMM model, oszacowany oryginalnie przez Caselli'ego, Esquivela i Leforta [1996] za pomocą FDGMM. Porównanie wyników uzyskanych na podstawie FDGMM i SGMM pozwala, zdaniem autorów na stwierdzenie, że SGMM jest najbardziej odpowiednią metodą estymacji panelowych modeli wzrostu. Model wzrostu Solowa, zastosowany przez Bonda, Hoeffler i Temple'a [2001] ma postać:

$$\Delta y_{it} = \alpha_{0t} + (\gamma - 1) y_{i,t-1} + \mathbf{x}_{it}^{\mathrm{T}} \boldsymbol{\beta} + \alpha_i + \varepsilon_{it} \quad \text{dla } i = 1, ..., N,\ t = 2, ..., T \tag{6.2}$$

gdzie: Δy_{it} jest przyrostem logarytmu PKB *per capita* w okresie 5 lat; $y_{i,t-1}$ – logarytmem PKB *per capita* na początku tego okresu; a $\mathbf{x}_{it}$ – wektorem zmiennych objaśniających, mierzonych na początku lub podczas tego okresu, którymi mogą być: logarytm stopy inwestycji (s_{it}), logarytm stopy wzrostu populacji (n_{it}) plus 0,05, gdzie 0,05 reprezentuje sumę wspólnej dla wszystkich obiektów egzogenicznej stopy postępu technicznego (g) i wspólnej stopy deprecjacji (δ). W rozszerzonym modelu Solowa regresorami mogą też być miary jakości kapitału ludzkiego, np. logarytm wskaźnika skolaryzacji (enr_{it}) dla szkół średnich (stosunku liczby uczniów tych szkół do liczby ludności w wieku odpowiadającym ustawowemu okresowi pobierania nauki na tym poziomie edukacji). Nieobserwowalny efekt grupowy α_i odzwierciedla m. in. różnice początkowego poziomu efektywności, zaś zmieniający się w czasie wyraz wolny (efekt czasowy α_{0t}) odzwierciedla jednakowe dla wszystkich obiektów zmiany produktywności.

Model (6.2) można zapisać równoważnie w postaci:

$$y_{it} = \alpha_{0t} + \gamma y_{i,t-1} + \mathbf{x}_{it}^{\mathrm{T}} \boldsymbol{\beta} + \alpha_i + \varepsilon_{it} \quad \text{dla } i = 1, ..., N,\ t = 2, ..., T \tag{6.3}$$

Spełnienie dodatkowych warunków momentów, wymaganych dla zastosowania SGMM, jest zagwarantowane, jeśli średnie po czasie wszystkich zmiennych (objaśnianej i objaśniających) są stałe dla każdego obiektu. Zostało to pokazane w pracy R. Blundella i S. Bonda [2000] dla funkcji produkcji. Bond, Hoeffler i Temple [2001] zauważyli, że o ile stałość średnich dla stopy inwestycji i stopy wzrostu populacji jest zgodna z postulatami modelu wzrostu Solowa, to stałość średniej wartości PKB *per capita* jest problematyczna. Jednakże włączenie do modelu efektów czasowych α_{0t} pozwala na założenie, że długookresowy wzrost PKB *per capita* oraz postęp techniczny są jednakowe

dla wszystkich krajów. Założenie, dotyczące jednakowego postępu technicznego, zostało przyjęte w pracy N. Mankiwa, D. Romera i D. Weila [1992] i jest odtąd standardowo przyjmowane w modelach wzrostu.

Bond, Hoeffler i Temple [2001] podkreślają, że włączenie efektów czasowych do modelu (6.2) jest równoważne przekształceniu zmiennych w odchylenia od średnich grupowych, czyli ich transformacji za pomocą operatora wewnątrzgrupowego. Jednakże w omawianym modelu, tak samo jak w pracy F. Caselli'ego, G. Esquivela i F. Leforta [1996], zmienne zostały przekształcone w odchylenia od średnich, co umożliwia porównywalność wyników oraz pominięcie efektów czasowych. Wyniki, prezentowane przez autorów, uzyskane są na podstawie estymatorów jednostopniowych FDGMM i SGMM ze średnimi błędami szacunku odpornymi na heteroskedastyczność.

Podstawowy model, postaci (6.2), w którym elementami wektora $\mathbf{x}_{it}$ są $\ln(s_{it})$ i $\ln(n_{it} + g + \delta)$, szacowany jest czterema metodami: MNK, WG, FDGMM, SGMM. Stopa inwestycji i stopa wzrostu populacji zostały uznane za zmienne endogeniczne, co przy zastosowaniu dwóch ostatnich metod znalazło odzwierciedlenie w zastąpieniu ich zmiennymi instrumentalnymi. Szczegółowe wyniki zawarte są w pracy: Bond, Hoeffler i Temple [2001] (tab. 1, s. 31). Wśród najważniejszych wniosków wymienić należy to, że ocena FDGMM parametru przy dochodzie początkowym $y_{i,t-1}$ jest niższa niż ocena WG, co wskazuje na jej obciążenie w dół, właściwe dla małej próby. Ponadto odpowiadająca jej ocena tempa konwergencji ($\hat{\beta}$ = 0,6) jest znacznie wyższa niż uzyskana innymi metodami. SGMM daje ocenę parametru przy dochodzie początkowym równą –0,101, przy czym jest to wartość znajdująca się pomiędzy oceną WG i MNK. Wynikająca stąd ocena tempa konwergencji $\hat{\beta}$ = 0,021 wydaje się znacznie sensowniejsza, oznacza bowiem, że stopa konwergencji wynosi 2%. Wyniki testu Sargana i różnicowego testu Sargana wskazują na właściwość użytych instrumentów i poprawność dodatkowych warunków momentów wykorzystanych przez SGMM. Zauważalne jest też zmniejszenie błędów szacunku estymatorów SGMM w stosunku do FDGMM. Instrumentami w modelu pierwszych różnic, wykorzystanymi w FDGMM, są dla odpowiednich zmiennych: $\ln(y_{i,t-2})$, $\ln(s_{i,t-2})$, $\ln(n_{i,t-2} + g + \delta)$ oraz wcześniejsze opóźnienia. Dodatkowymi instrumentami w równaniach na poziomach, wykorzystanymi w SGMM, są dla odpowiednich zmiennych: $\Delta\ln(y_{i,t-1})$, $\Delta\ln(s_{i,t-1})$, $\Delta\ln(n_{i,t-1} + g + \delta)$. Traktowanie stopy inwestycji i stopy wzrostu populacji jako potencjalnych zmiennych endogenicznych umożliwia uwzględnienie zmiennego w czasie, tzn. niesystematycznego błędu pomiaru każdej z nich (założenie (4.1)). Przy założeniu niesystematycznego, błędu pomiaru PKB *per capita*, zarówno poziom tej zmiennej z okresu $t - 2$ w równaniu na przyrostach, jak i przyrost z okresu $t - 1$ w równaniu na poziomach nie są właściwymi instrumentami. Mimo że testy Sargana nie wykazały niewłaściwości instrumen-

tów użytych w SGMM, metodę tę zastosowano ponownie, z wyłączeniem wymienionych zmiennych instrumentalnych. Dokładniej, $\ln(y_{i,t-2})$ została pominięta w zbiorze instrumentów dla modelu pierwszych różnic, a $\Delta\ln(y_{i,t-1})$ – zastąpiona przez $\Delta\ln(y_{i,t-2})$ w równaniach na poziomach. Wyniki, również prezentowane przez autorów w tab. 1, różnią się bardzo nieznacznie od uzyskanych w podstawowej wersji SGMM. Wydaje się to wskazywać na brak problemów z błędami pomiaru PKB *per capita*. Oprócz wspomnianego już określenia stopy konwergencji na poziomie 2%, oceny SGMM wskazują na istotny, pozytywny wpływ stopy inwestycji (ok. 0,19) na poziom PKB *per capita* właściwy dla stanu wzrostu zrównoważonego (ang. *steady state*) i to, co podkreślają autorzy, przy uwzględnieniu efektów grupowych i możliwej endogeniczności inwestycji.

S. Bond, A. Hoeffler i J. Temple [2001] estymowali też model, w którym do zbioru zmiennych objaśniających dołączono wskaźnik skolaryzacji dla szkół średnich. Wyniki są prezentowane przez autorów w tab. 2 (s. 32). Dodanie tej zmiennej poprawiło nieco jakość estymatora FDGMM – ocena parametru przy dochodzie początkowym jest porównywalna z oceną WG. Wiadomo jednak, że w przypadku modeli dynamicznych, estymator WG w małych próbach jest obciążony, ponadto tempo konwergencji obliczone na podstawie ocen tego estymatora wynosiłoby aż 0,08. Bardziej poprawne wyniki uzyskano na podstawie SGMM – stopa konwergencji wyznaczona z modelu uwzględniającego wpływ wskaźnika skolaryzacji wynosi 1,7%, jednakże błąd szacunku parametru przy tej zmiennej przekracza wartość oceny tego parametru. Na tej podstawie zasugerowano nieuwzględnianie wskaźnika skolaryzacji bezpośrednio w modelu, ale raczej dołączenie opóźnionych wartości tej zmiennej do zbioru instrumentów, wykorzystywanych w FDGMM w celu rozwiązania problemu słabych instrumentów. Rezultaty przedstawione są w tab. 3 (s. 33), w tekście oryginalnym. Dokładniej, zbiór instrumentów dla równań na przyrostach podstawowego modelu Solowa, został rozszerzony o $\ln(enr_{i,t-1})$ oraz wcześniejsze opóźnienia. W ten sposób uzyskane oceny FDGMM są bliskie ocenom SGMM. Opóźniony wskaźnik skolaryzacji pomaga zatem przewidywać wzrost PKB *per capita* w równaniach postaci zredukowanej, podczas gdy bieżące wartości tej zmiennej nie wpływają na właściwy dla stanu wzrostu zrównoważonego poziom PKB *per capita* w rozważanym modelu. Oznacza to, zdaniem autorów, że wskaźnik skolaryzacji nie oddziałuje na wzrost gospodarczy bezpośrednio, ale za pośrednictwem stopy inwestycji.

6.4. Determinanty TFP w regionach

Prace, w których badane są różnice w rozwoju regionalnym we Włoszech, należą w literaturze do licznych i często cytowanych. Wydaje się to wynikać ze znacznego rozwarstwienia między północnymi a południowymi regionami tego

kraju. Poniżej przedstawione są szerzej dwa opracowania analizujące kształtowanie się TFP w regionach Włoch.

G. A s c a r i i V. D i C o s m o [2004] zaprezentowali przykład zastosowania dynamicznego modelu panelowego do wyjaśnienia różnic łącznej produktywności czynników produkcji (TFP) między regionami administracyjnymi Włoch. Autorzy podkreślają, że TFP oraz kapitał ludzki są najbardziej istotnymi czynnikami, wyjaśniającymi różnice w dochodach między krajami, regionami *etc*. Wśród potencjalnych determinant TFP wymienia się: poziom działalności badawczo-rozwojowej, kapitał ludzki, transfer technologii, rolę związków zawodowych, stabilność polityczną i poziom demokracji.

Poziom TFP wyznaczony został na podstawie funkcji Cobba–Douglasa z zależności:

$$TFP_{it} = \frac{Y_{it}}{K_{it}^{1-\alpha_i} L_{it}^{\alpha_i}} \tag{6.4}$$

gdzie Y_{it} – wartość dodana w regionie i, w roku t, K_{it} – miara kapitału, L_{it} – miara nakładu pracy. Wartość parametru α_i oszacowano przyjmując założenie doskonałych rynków czynników produkcji jako $\alpha_i = w_i \cdot L_i / Y_i$, gdzie w_i jest przeciętnym wynagrodzeniem w cenach stałych, w regionie i. Parametr ten jest stały w czasie, gdyż jest, zdaniem autorów, parametrem technologicznym, który nie powinien się zmienić w okresie 15 lat objętych próbą. Z drugiej strony, udział pracy w poszczególnych regionach wykazuje dużą zmienność w czasie, co autorzy uzasadniają działaniem czynników innych niż technologiczne. Zatem wartość α_i obliczono jako średnią po czasie dla i-tego regionu.

Na podstawie danych pochodzących z 20 regionów administracyjnych Włoch z lat 1985–2000 skonstruowano statyczny i dynamiczny model kształtowania się TFP. Model statyczny ma postać:

$$\ln TFP_{it} = \beta_0 + \beta_1 rd_{it} + \beta_2 ricpop_{it} + \beta_3 import_{it} + \beta_4 edu_{it} + \beta_5 socialk_{it} + \alpha_i + \varepsilon_{it} \tag{6.5}$$

a model dynamiczny:

$$\ln TFP_{it} = \gamma \ln TFP_{i,t-1} + \beta_1 rd_{it} + \beta_2 ricpop_{it} + \beta_3 import_{it} + \beta_4 edu_{it} + \beta_5 socialk_{it} + \\ + \alpha_i + \varepsilon_{it} \tag{6.6}$$

gdzie: *TFP* – łączna produktywność czynników produkcji, wyznaczona według wzoru (6.4); *rd* – stosunek nakładów na badania i rozwój do inwestycji w środki

trwałe ogółem; *ricpop* – stosunek liczby pracujących naukowców do liczby pracujących ogółem; *import* – stosunek importu do PKB; *edu* – przeciętna dla pracujących liczba lat nauki; *socialk* – kapitał społeczny, zmienna wyznaczona metodą głównych składowych, jako pierwsza główna składowa trzech wybranych wskaźników przestępczości. Zmienną objaśnianą modelu jest ln*TFP*, gdyż zależność między TFP a zmiennymi objaśniającymi nie ma charakteru liniowego.

Model (6.5) potraktowano jako model FE i oszacowano za pomocą estymatora wewnątrzgrupowego. Zmienne, opisujące działalność badawczo-rozwojową i kwalifikacje pracowników są statystycznie istotne, a oceny parametrów są dodatnie (0,65 dla *rd*, 2,49 dla *ricpop* i 0,099 dla *edu*). Wyciągnięto stąd wniosek, że postęp technologiczny jest istotną determinantą reszty Solowa oraz że kapitał ludzki i wykształcenie pracowników przekłada się na wzrost TFP. Wnioski takie są potwierdzeniem tez stawianych w publikacjach o charakterze teoretycznym na temat tego, że postęp technologiczny i kapitał ludzki są najbardziej istotnymi czynnikami, wpływającymi na wzrost TFP. Zmienne *import* i *socialk* są nieistotne w modelu (6.5). Może to, zdaniem autorów, wynikać z braku danych, odzwierciedlających wpływ wymiany handlowej na TFP na poziomie regionalnym. Zmienna *import* nie uwzględnia tego, czy importowane dobra pozostają w regionie, czy są przesyłane dalej, np. z regionów przygranicznych. Ponadto brak jest danych regionalnych opisujących strukturę sprowadzanych dóbr, co uniemożliwia wyodrębnienie wpływu transferu technologii na TFP. G. Ascari i V. Di Cosmo stwierdzili też, że wysokie wartości statystyk *t*-Studenta dla efektów grupowych (rzędu $t_{emp} = -75$) mogą wskazywać na problem pominiętych zmiennych. Wprowadzili zatem do modelu opóźnioną zmienną endogeniczną, modyfikując go do postaci (6.6). Model ten oszacowano dwiema metodami: za pomocą estymatora wewnątrzgrupowego (WG) z korektą Kivieta (patrz pkt 3.1.7) oraz FDGMM. Wyniki wskazują na wysoki stopień trwałości zmiennej objaśnianej – ocena parametru γ modelu (6.6) wynosi od 0,977 dla estymatora WG do 0,794 dla estymatora FDGMM. Jednocześnie, w przypadku estymatora WG wszystkie pozostałe zmienne są nieistotne, co może, zdaniem autorów, wynikać ze zbyt małego wymiaru przekrojowego próby. W wyniku zastosowania FDGMM z macierzą instrumentów postaci (3.41), a więc z uwzględnieniem endogeniczności zmiennych objaśniających, uzyskano wyniki zbliżone do wyników modelu statycznego, w sensie istotności zmiennych objaśniających. Zmiennymi statystycznie istotnymi okazały się *rd* i *ricpop*, zaś pozostałe zmienne (*import*, *edu*, *socialk*) są nieistotne. Nieistotność zmiennej *edu* w modelu dynamicznym[5] tłumaczona jest jako efekt jej skorelowania z opóźnioną wartością TFP; w modelu statycznym

[5] Sytuacja ta jest podobna do opisanej w podrozdz. 6.3 próby włączenia współczynnika solaryzacji do modelu Bonda, Hoeffler i Temple'a.

zmienna *edu* odzwierciedlać mogła wpływ opóźnionej wartości TFP na bieżącą wartość tej zmiennej. Jako alternatywne wyjaśnienie G. Ascari i V. Di Cosmo wskazują na problem regresji pozornej – obie zmienne: *TFP* i *edu* wykazują trend rosnący, nieobecny w pozostałych zmiennych modelu.

G. Ascari i V. Di Cosmo podjęli też próbę opisu rozwarstwienia rozwoju ekonomicznego między uprzemysłowioną północą Włoch a słabo rozwiniętym południem poprzez oszacowanie oddzielnych modeli dla tych dwóch części kraju. Estymacja modelu dynamicznego nie powiodła się ze względu na zbyt małą liczbą obserwacji przestrzennych (odpowiednio 12 i 8 regionów). Model statyczny postaci (6.5) dla regionów północnych można uznać za zadowalający. Wszystkie zmienne objaśniające, łącznie z *socialk* są istotne, a oceny parametrów dodatnie. Jakość modelu dla regionów południowych jest gorsza: zmienne *rd* i *ricpop* są nieistotne, a ocena parametru przy *socialk* jest ujemna. Autorzy interpretują te wyniki jako odzwierciedlenie prawidłowo funkcjonującej gospodarki północnych Włoch i nieprawidłowości charakteryzujących południe, w szczególności skutków przestępczości zorganizowanej. Szczególnie podkreślany jest wpływ zmiennej *socialk*, który w przypadku północy oznacza pozytywny wpływ kapitału społecznego, traktowanego jako przybliżona miara poczucia bezpieczeństwa społecznego, na wzrost produktywności. W zakończeniu podkreślono nierównomierność rozwoju gospodarczego Włoch i uznano ją za główną przyczynę niedoskonałości modelu postaci (6.6) dla całego kraju.

Nieco inne podejście do badania konwergencji zastosowali J. Byrne, G. Fazio i D. Piacentino [2005]. Podczas gdy większość badań empirycznych nad konwergencją koncentruje się na analizach PKB *per capita* lub produktywności pracy, autorzy badali proces konwergencji między regionami Włoch w odniesieniu do TFP w długim okresie. Uzasadnieniem takiego podejścia jest teoria neoklasyczna, zgodnie z którą różnice strukturalne między regionami wpływają na TFP i w konsekwencji na długookresowy wzrost. Ponadto wykazać można, że zróżnicowanie wydajności pracy zależy w większym stopniu od TFP niż od technicznego uzbrojenia pracy. TFP odzwierciedla zaś szerokie spektrum materialnych i niematerialnych czynników, które wpływają na efektywność gospodarki. J. Byrne, G. Fazio i D. Piacentino [2005] odeszli ponadto od tradycyjnych metod badania konwergencji typu β lub σ na rzecz zastosowania panelowych testów stacjonarności. Występowanie stacjonarności interpretują oni jako występowanie konwergencji. Wartości TFP dla regionów Włoch dla lat 1970–2001 oszacowano stosując metodologię rachunku wzrostu bazującą na funkcji Cobba–Douglasa z założeniem postępu technicznego zwiększającego efektywność pracy (patrz podrozdz. 5.1). Ocenę parametru α wyrażającego elastyczność produkcji względem pracy oszacowano jako $\alpha_{it} = w_{it} \cdot L_{it} / Y_{it}$, a więc podobnie jak w pracy: Ascari, Di Cosmo [2004]; dopuszczono jednakże zróżnicowanie tego parametru zarówno po obiektach, jak

i w czasie. Taka konstrukcja odzwierciedla dynamikę regionalnych różnic w strukturach gospodarek. Następnie na podstawie funkcji Cobba–Douglasa wyznaczono wartości TFP specyficzne dla regionów i czasu. Zaobserwowano przy tym znaczne różnice w produktywności pracy: w regionach północno-zachodnich przekracza ona średnią krajową o 11%, a w regionach południowych jest od niej niższa o 14%. Podobne dysproporcje zauważono w odniesieniu do TFP; zróżnicowanie kapitałochłonności jest zaś znacznie słabsze. Ponadto produktywność pracy jest silnie skorelowana z TFP (r = 0,8), a słabo (r = 0,26) z kapitałochłonnością. Wyciągnięto stąd wniosek, że dla wyjaśnienia różnic w produktywności pracy najbardziej istotna jest analiza TFP.

Jako ostatni etap badania autorzy zastosowali panelowe testy stacjonarności[6] w celu zbadania konwergencji regionalnej w odniesieniu do TFP. Zastosowano dwie alternatywne miary odległości TFP między regionami. Dla całej próby i dla obu tych miar dwa testy: LLC Levina, Lina i Chu [2002] i IPS Ima, Pesarana i Shina [2003], dały jednakowy wynik: brak podstaw do odrzucenia hipotezy zerowej o braku stacjonarności. Oznacza to niewystępowanie konwergencji między regionami. Test MADF Sarno i Taylora [1998] wskazał zaś na występowanie konwergencji. Ze względu na różnice konstrukcji testów, wyniki te zinterpretowano jako wskazanie na występowanie konwergencji klubowej. Dokładniej, po wyodrębnieniu odpowiednich podprób stwierdzono występowanie długookresowej tendencji zbieżnej między regionami południowymi (do wartości poniżej średniej krajowej). Regiony północne rozwijają się natomiast według własnych ścieżek, przypuszczalnie ze względu na różne endogeniczne źródła wzrostu.

6.5. Analiza β-konwergencji województw

Badania dotyczące rozwoju i wzrostu regionalnego, obejmujące regiony Polski, prezentowane były m. in. w pracach: Bal-Domańska [2009]; Strahl i Bal-Domańska [2009]; Ciołek [2004, 2005]; Malaga [2004]; Malaga, Kliber [2007]; Michałek, Siwiński, Socha [2003, 2007]; Próchniak, Witkowski [2006]; Rapacki [2009]; Strahl [2006]; Sztaudynger [2005]; Tokarski [2008b]; Welfe [2001, 2007]; Wójcik [2008] i in. Jednakże w pracach tych albo posługiwano się metodami innymi niż panelowe (modelami ekonometrycznymi szacowanymi na podstawie danych przekrojowych lub czasowych, modelami geografii ekonomicznej, metodami

[6] Na temat konstrukcji alternatywnych panelowych testów stacjonarności czytaj w dodatku I.

kalibracji lub metodami wielowymiarowej analizy porównawczej), albo zasięg terytorialny badań był szerszy niż obszar Polski.

W niniejszym punkcie przedstawiona zostanie próba skonstruowania i oszacowania własnego modelu badającego występowanie β-konwergencji województw w Polsce. Zasadniczy model bazuje na modelu Bonda, Hoeffler i Temple'a [2001] (patrz podrozdz. 6.3). Podjęto przy tym próby zastosowania alternatywnych miar zmiennych objaśniających oraz próbę uwzględnienia w modelu miar kapitału ludzkiego.

6.5.1. Postać ogólna modelu β-konwergencji i specyfikacja zmiennych

Punktem wyjścia neoklasycznego modelu wzrostu, zakładającego występowanie konwergencji, jest równanie:

$$\ln\left(\frac{y_{it}}{y_{i,t-1}}\right) = \alpha_0 - \gamma \ln(y_{i,t-1}) + \boldsymbol{\theta}^{\mathrm{T}}\mathbf{x}_{it} + (\alpha_i + \varepsilon_{it}) \tag{6.7}$$

Model (6.7) zapisać można alternatywnie, w postaci dogodniejszej do estymacji:

$$\ln(y_{it}) = \alpha_0 + (1-\gamma)\ln(y_{i,t-1}) + \boldsymbol{\theta}^{\mathrm{T}}\mathbf{x}_{it} + (\alpha_i + \varepsilon_{it}) \tag{6.8}$$

gdzie:

y_{it} – poziom PKB *per capita* w i-tym regionie w okresie t,

$\mathbf{x}_{it}$ – wektor (logarytmów) zmiennych objaśniających,

α_0, γ, $\boldsymbol{\theta}$ – parametry strukturalne,

α_i – efekty grupowe,

$\varepsilon_{\mathrm{it}}$ – składnik losowy.

Parametr γ w modelach (6.7) i (6.8) jest, zgodnie z teorią konwergencji warunkowej, określany jako:

$$\gamma = \frac{1}{T}\left(1 - \mathrm{e}^{-\beta T}\right) \tag{6.9}$$

Wartość β decyduje o istnieniu i sile konwergencji. Jeśli w wyniku estymacji otrzyma się dodatnią ocenę tego parametru, to potwierdza się hipoteza o występowaniu konwergencji typu β, a zatem, że gospodarki biedniejszych regionów wykazują wyższą stopę wzrostu niż gospodarki bogatsze. Wartości ujemne oznaczałyby występowanie dywergencji, czyli przestrzennego zwiększania zróżnicowania zmiennej y.

Podstawowy model wzrostu i konwergencji (6.8) może być dla celów badań empirycznych modyfikowany. Różnice dotyczą zbioru zmiennych objaśniających, czyli elementów wektora $\mathbf{x}_{it}$. Polegają one na wprowadzaniu alternatywnych miar nakładów kapitału i pracy lub dodatkowych zmiennych objaśniających. Wspomnieć tu warto o propozycji G. Juszczak-Szumacher [1996] rozszerzenia zbioru zmiennych objaśniających produkcję przemysłową w Polsce w ujęciu gałęziowym o (skonstruowany na potrzeby badania) współczynnik kosztów materiałowych i miernik postępu technicznego. Inną interesującą propozycją, którą traktować można jako próbę endogenizacji postępu technicznego, jest poszerzenie zbioru zmiennych objaśniających o nakłady na badania i rozwój (BiR). Przedstawiona jest ona w pracy T. Tokarskiego [1995]. Podejmowane są też próby konstrukcji modeli wzrostu uwzględniających kapitał ludzki. Pod pojęciem kapitału ludzkiego rozumie się zazwyczaj poziom uzyskanego wykształcenia, lub szerzej, kondycję zdrowotną pracujących. Obszerne rozważania na temat kapitału ludzkiego zawarte są też w pracy J. J. Sztaudyngera [2005]. Omówienie alternatywnych metod pomiaru kapitału ludzkiego i przykłady empirycznych modeli, które uwzględniają tę zmienną, znaleźć można np. w pracy W. Welfe [2007] (rozdz. 5). Jedna z tych miar opisana jest wzorem:

$$hk = \sum_{i=0}^{4} \frac{h_i}{h_0} lprac_i \tag{6.10}$$

gdzie:

hk – rozmiar kapitału ludzkiego,

h_i – liczba lat kształcenia dla *i*-tego poziomu wykształcenia (indeks 0 oznacza wykształcenie podstawowe),

$lprac_i$ – liczba pracujących z *i*-tym poziomem wykształcenia.

Również w prezentowanym tu badaniu podejmowano próby alternatywnego doboru zmiennych objaśniających, z wykorzystaniem sugestii przedstawionych powyżej. Zmienną objaśnianą estymowanych modeli był produkt krajowy brutto na 1 mieszkańca (*PKBpc*; w zł). Lista potencjalnych zmiennych objaśniających, którymi dysponowano podczas przeprowadzania badania, jest następująca:

• dla określenia wielkości kapitału fizycznego:

gludpr – stopa wzrostu liczby ludności w wieku produkcyjnym (powyżej 17 lat) powiększona o stopę deprecjacji kapitału i stopę postępu technicznego $(g + \delta)$[7],

glp – stopa wzrostu liczby pracujących powiększona o $(g + \delta)$,

[7] Wartość 0,05, o którą powiększane są stopy wzrostu, jest stopą deprecjacji kapitału i stopą postępu technicznego. Jest to wartość standardowo przyjmowana w większości badań, a przedstawiona po raz pierwszy w: Caselli, Esquivel, Lefort [1996].

glud – stopa wzrostu liczby ludności powiększona o ($g + \delta$),

tup – techniczne uzbrojenie pracy, mierzone wartością środków trwałych brutto na pracującego (w tys. zł);

• dla określenia wielkości inwestycji:

sinw – stopa inwestycji, liczona jako udział nakładów inwestycyjnych w PKB,

inwpr – nakłady inwestycyjne na pracującego (w zł);

• dla określenia wielkości nakładów na działalność badawczo-rozwojową (BiR):

snbr – nakłady na BiR w relacji do PKB (w %),

nbrpc – nakłady na BiR na 1 mieszkańca (w zł),

nbrinw – nakłady na BiR w relacji do nakładów inwestycyjnych;

• dla określenia wielkości kapitału ludzkiego:

hk – kapitał ludzki, obliczony według wzoru (6.10),

edu – kapitał ludzki, obliczony jako suma udziałów pracujących z kolejnymi poziomami wykształcenia, mnożonych przez liczbę lat nauki. Miarę tę obliczono zgodnie z sugestią G. Ascari i V. Di Cosmo [2004] (s. 22).

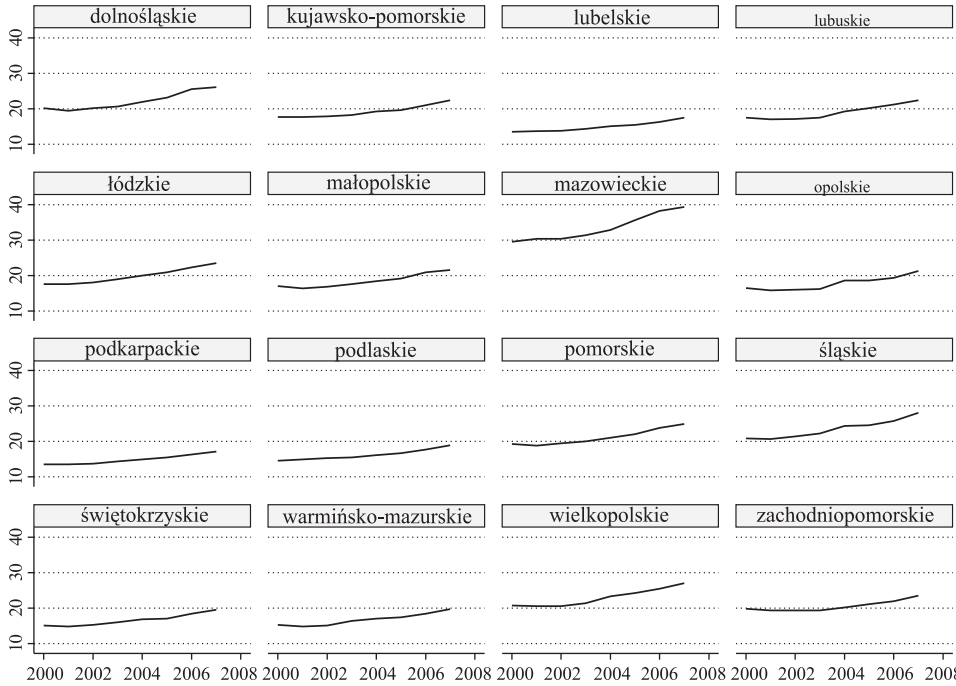
Wszystkie zmienne wyrażone są w cenach stałych z 2000 r.

Dane statystyczne, wykorzystane w badaniu, pochodzą z witryny internetowej GUS oraz z roczników statystycznych województw. Obejmują one okres od 2000 do 2007 r. Zebrano je dla 16 województw naszego kraju.

Dane liczby pracujących według poziomu wykształcenia, według województw, niezbędne do konstrukcji miar wielkości kapitału ludzkiego dostępne są dopiero od roku 2002. Z tego powodu na podstawie pełnej próby (z lat 2000–2007) oszacowano modele, w których kapitał ludzki nie był uwzględniany. Następnie próba została zawężona do okresu 2002–2007 i oszacowano modele uwzględniające wpływ kapitału ludzkiego na PKB *per capita*. Wyniki estymacji przedstawione są poniżej.

6.5.2. Wyniki estymacji modelu β-konwergencji

Zmienną objaśnianą w prezentowanych modelach jest PKB *per capita* (*PKBpc*). Kształtowanie się tej zmiennej w poszczególnych województwach w okresie objętym badaniem przedstawione jest na rys. 6.1. Analiza wzrokowa tego wykresu pokazuje, że poziom PKB *per capita* we wszystkich województwach w latach 2000–2007 wykazywał tendencję rosnącą. Województwa różnią się jednak dość znacznie, zarówno pod względem wyjściowego poziomu dochodu, jak i pod względem tempa jego wzrostu. Dlatego zasadne wydaje się pytanie, czy tempo wzrostu poziomu PKB *per capita* jest wystarczające, by następowało wyrównywanie poziomu *PKBpc* w województwach, a zatem obserwowalna była ich konwergencja.



Rys. 6.1. Poziom PKB *per capita* w tys. zł według województw, w latach 2000–2007

Ź r ó d ł o: opracowanie własne.

Badanie występowania konwergencji województw prowadzone było w oparciu o model (6.8). Oszacowanych zostało kilkanaście modeli, w których, w roli zmiennych objaśniających, uwzględniano wymienione powyżej alternatywne miary czynników, mogących mieć wpływ na PKB *per capita*. Wymienione powyżej zmienne wprowadzane były do modelu albo jako wartości bieżące, albo opóźnione. Dodatkowo próbowano wprowadzać do modelu zmienne zero-jedynkowe dla wybranych lat.

Najlepsze wyniki statystyczno-merytoryczne uzyskano dla najprostszego modelu, dość szeroko wykorzystywanego w badaniach empirycznych. Stosowany on był m. in. w: C a s e l l i, E s q u i v e l, L e f o r t [1996]; B o n d, H o f f l e r, T e m p l e [2001]; C i o ł e k [2004]. Wymienione badania dotyczyły konwergencji regionów Unii Europejskiej, natomiast w prezentowanym tu badaniu postawiono za cel weryfikację hipotezy o występowaniu konwergencji typu β województw w Polsce. Postać modelu jest następująca:

$$\ln(PKBpc_{it}) = \alpha_0 + (1-\gamma)\ln(PKBpc_{i,t-1}) + \theta_1 \ln(glud) + \theta_2 \ln(inwpr_{i,t-1}) + \alpha_i + \varepsilon_{it} \tag{6.11}$$

Tabela 6.2

Wyniki estymacji modeli konwergencji typu β nieuwzględniających kapitału ludzkiego

Zmienna objaśniająca/ test statystyczny	Metoda estymacji					
	SGMM endo	SGMM pre	SGMM egzo	FDGMM endo	WG	KMNK
	1	2	3	4	5	6
$\ln(PKBpc_{i,t-1})$	1,0355*** (0,0472)	1,0092*** (0,0408)	1,0932*** (0,0644)	1,1755*** (0,0617)	1,0813*** (0,0695)	0,9781*** (0,0256)
$\hat{\gamma}$	–0,0355	–0,0092	–0,0932	–0,1755	–0,0813	0,0219
$\hat{\beta}$	–0,0313	–0,0089	–0,0696	–0,1096	–0,0626	0,0241
$\ln(glud_{it})$	–0,3380*** (0,1441)	–0,4201** (0,1637)	–0,4420*** (0,7933)	–0,5553*** (0,1676)	–0,3838** (0,2122)	–0,1489*** (0,0649)
$\ln(inwpr_{i,t-1})$	0,0671*** (0,0263)	0,0821*** (0,0252)	0,0478 (0,0343)	0,0135 (0,0316)	0,0539** (0,0331)	0,0542*** (0,0191)
const.	–1,4738** (0,7964)	–1,4939*** (0,7374)	–2,3750*** (0,9763)	–3,3943*** (0,8793)	–2,0352 (0,9656)	–0,3110 0,3070
m_1 [*p-value*]	–2,016 [0,014]	–1,982 [0,017]	–1,998 [0,046]	–2,145 [0,032]	–	–
m_2 [*p-value*]	1,109 [0,268]	1,084 [0,278]	0,662 [0,508]	0,288 [0,773]	–	–
Sargan [*p-value*]	64,215 [0,139]	86,744 [0,014]	66,904 [0,000]	66,464 [0,007]	–	–
dif-Sargan [*p-value*]	15,751 [0,203]	15,571 [0,273]	10,343 [0,111]	–	–	–
R^2	–	–	–	–	0,980	0,982

Wartość β określa szybkość konwergencji i jest obliczona ze wzoru (6.9) na podstawie oszacowanej wartości parametru $\hat{\gamma}$ przy zmiennej $\ln(PKBpc_{i,t-1})$;

symbol *** przy ocenie parametru oznacza, że wartość $p < 0{,}05$ dla testu *t*-Studenta; symbol ** oznacza $p \leq 0{,}1$; pozostałych zmiennych nie można uznać za istotne;

w nawiasach, pod ocenami parametrów, zamieszczono błędy szacunku parametrów, wyznaczane na podstawie odpornych estymatorów wariancji składnika losowego;

m_1 i m_2 oznaczają wartości empiryczne testu autokorelacji Arellano–Bonda, odpowiednio AR(1) i AR(2); w nawiasach kwadratowych – wartości *p* dla H_0: autokorelacja pierwszego (drugiego) rzędu nie występuje;

Sargan i *dif*-Sargan oznaczają wartości empiryczne testu Sargana i różnicowego testu Sargana; w nawiasach kwadratowych – wartości *p* dla H_0: warunki ponadidentyfikujące są spełnione, odczytane dla odpowiedniej liczby stopni swobody, zależnej od liczby wykorzystanych instrumentów.

Ź r ó d ł o: obliczenia własne.

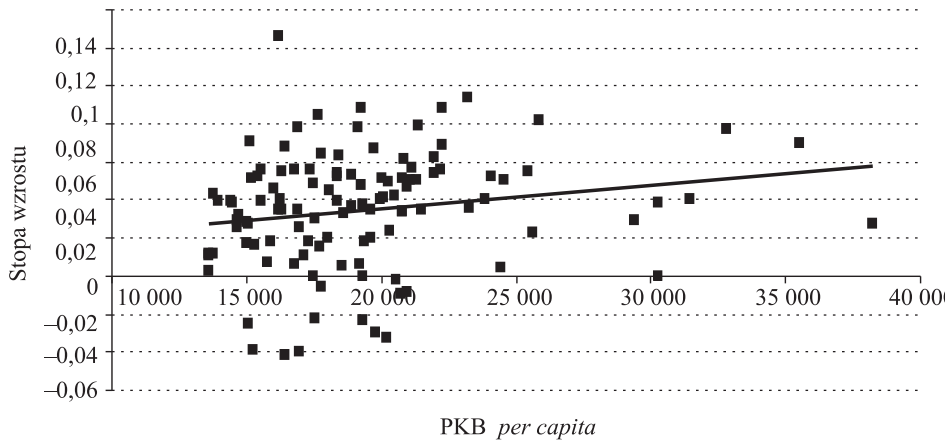
Do estymacji modelu (6.11) stosowane były FDGMM i SGMM. Zmienne objaśniające (wszystkie lub część z nich) traktowane były alternatywnie jako zmienne ściśle egzogeniczne, słabo egzogeniczne lub endogeniczne.

W tab. 6.2 przedstawione są wybrane wyniki estymacji modelu (6.11) uzyskane przy zastosowaniu wspomnianych powyżej metod. Prezentowane wyniki uzyskano na podstawie estymatorów jednostopniowych. Zastosowanie metod dwustopniowych nie dało zadowalających wyników. W wyniku wielu prób, uznano ostatecznie, że zmienną nieegzogeniczną są nakłady inwestycyjne na pracującego (*inwpr*). Endogeniczność tej zmiennej jest zgodna z oczekiwaniami (por. pkt 6.1.1). Oznaczenia „endo", „pre" i „egzo" przy nazwach metod SGMM i FDGMM określają, jak traktowana była zmienna *inwpr*: endogenicznie, słabo egzogenicznie lub ściśle egzogenicznie.

W kolumnach 5 i 6 tab. 6.2 zawarte są wyniki estymacji modelu (6.11) uzyskane na podstawie estymatora wewnątrzgrupowego (WG) i estymatora KMNK modelu *pooled*. Celem ich zamieszczenia jest określenie przedziału, w którym powinna się zawierać zgodna ocena parametru autoregresyjnego γ. W przedziale tym: od –0,0813 do 0,0219 zawierają się zamieszczone w kolumnach 1 i 2 oceny $\hat{\gamma}$ uzyskane na podstawie SGMM, przy założeniu, że nakłady inwestycyjne są zmienną endogeniczną ($\hat{\gamma}$ = –0,0355) lub z góry ustaloną ($\hat{\gamma}$ = –0,0092). Potraktowanie *inwpr* jako zmiennej egzogenicznej w procesie estymacji SGMM (kolumna 3) lub zastosowanie FDGMM (kolumna 4) skutkuje tym, że estymatory parametru γ, a co za tym idzie oceny parametru konwergencji β są obciążone. Tempo rocznej konwergencji wyznaczone na podstawie FDGMM wynosiłoby prawie –11%, a na podstawie SGMM, przy założeniu egzogeniczności zmiennej *inwpr* – prawie –7%. Są to wartości co do modułu przeszacowane. Opisane wnioski są zgodne z przedstawianymi w literaturze tezami, że zastosowanie FDGMM do estymacji modeli wzrostu nie jest poprawne[8]. Wysoki stopień trwałości szeregów (ang. *high persistence*) powoduje bowiem, że liczba instrumentów możliwych do wykorzystania dla modelu pierwszych różnic maleje. Co ciekawe, wyniki testu Arellano–Bonda nie wskazują na problemy ze słabością instrumentów. Wartości statystyk m_2 wskazują na brak autokorelacji drugiego rzędu w modelu pierwszych różnic, niezależnie od metody estymacji. Jedynie porównanie *p-value* dla statystyki m_1, która w kolumnach 3 i 4 jest wyższa niż w kolumnach 1 i 2, może sygnalizować pojawiające się problemy z poprawnością metody estymacji. Znacznie wrażliwszy natomiast okazał się test Sargana. Zgodne z jego wynikami, na poziomie istotności 0,05, tylko SGMM z założeniem endogeniczności nakładów inwestycyjnych jest poprawną metodą estymacji modelu (6.11). Dlatego dalsza analiza dotyczyć będzie wyników uzyskanych tą metodą.

[8] Więcej na ten temat znaleźć można w podrozdz. 3.2 i 6.3.

Najważniejszym wnioskiem uzyskanym w wyniku oszacowania modelu (6.11) jest brak tendencji zbieżnej w kształtowaniu się PKB *per capita* według województw. Uzyskana ocena parametru β, decydującego o istnieniu i sile konwergencji, jest ujemna: $\hat{\beta} = -0{,}0322$. Świadczy to o występowaniu w Polsce zjawiska dywergencji, a więc rozbieżności w kształtowaniu się PKB *per capita* w województwach, przy czym tempo tej rozbieżności wynosi 3,2% rocznie. Jest to tendencja negatywna, ponieważ utrwala ona podział województw na bogatsze i biedniejsze. Niestety, potwierdzeniem uzyskanych wyników jest ocena współczynnika korelacji między roczną stopą wzrostu PKB *per capita* a opóźnionym poziomem wartości tej cechy w województwach (por. T o k a r s k i [2008b]). Współczynnik ten: $r(PKBpc_{i,t-1}, gPKBpc_{it})$ ma wartość 0,1821. Znak dodatni wskazuje na istnienie tendencji rozbieżnej, a więc na zwiększanie się różnic PKB *per capita* między województwami. Zależność między rocznymi stopami wzrostu BKB *per capita* a opóźnionym poziomem wartości tej zmiennej w województwach w latach 2000–2007 przedstawiono na rys. 6.2.



Rys. 6.2. Stopy wzrostu *PKBpc* w roku *t* a poziom *PKBpc* w zł, w roku *t* – 1, w latach 2000–2007

Ź r ó d ł o: opracowanie własne.

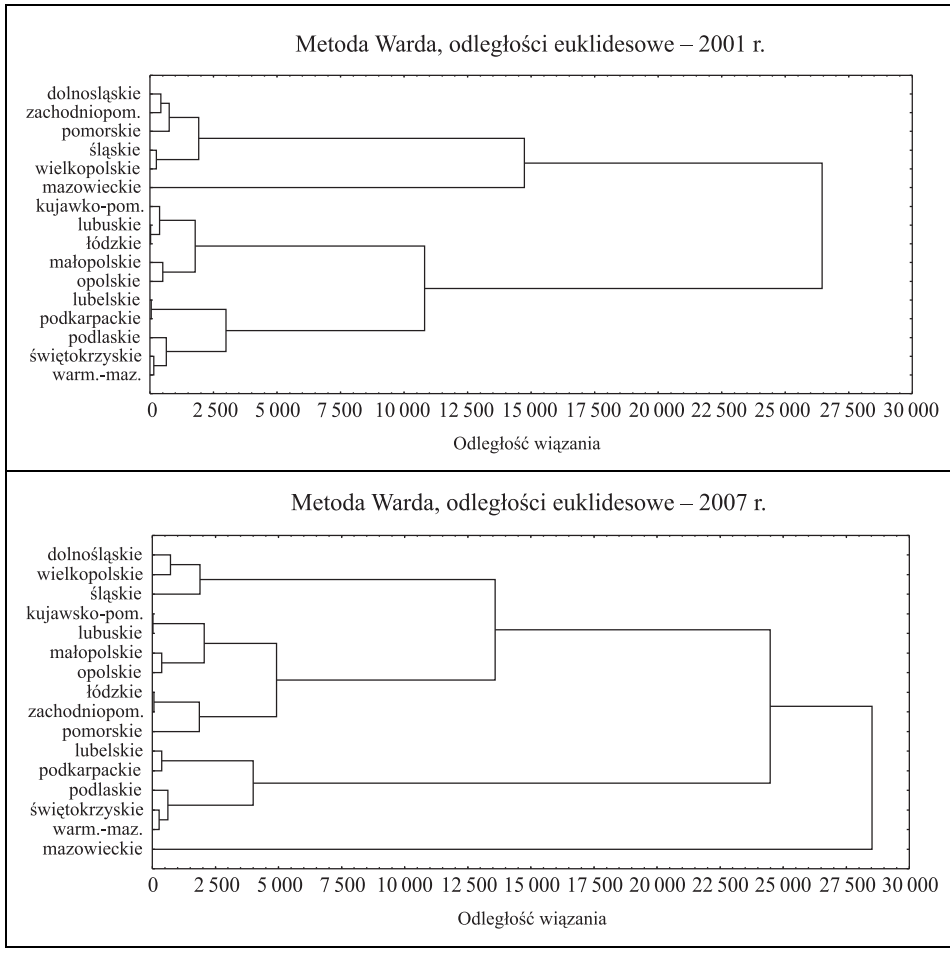
Wniosek o występowaniu dywergencji województw jest zbieżny z wnioskami innych opracowań, np. P. W ó j c i k a [2004] i cytowanych przez niego wcześniejszych prac. Autor prowadził swoje badanie na podstawie danych z lat 1990–2001, a więc bezpośrednio poprzedzających okres niniejszej analizy. Wyodrębnił on lata 1990–1995 jako okres zmniejszania dysproporcji miedzy województwami, a lata 1995–2001 jako okres rozwarstwiania się dochodu i pojawiania się tendencji dywergencyjnej. Być może prezentowane obecnie wyniki uchwyciły kontynuację tej tendencji. Uzyskane obecnie wyniki są ponadto zgodne z wynikami prezentowanymi w referacie B. D a ń s k i e j-

-B o r s i a k [2008], uzyskanymi na podstawie danych z lat 2001–2006. Otrzymana wówczas ocena parametru konwergencji wynosiła $\hat{\beta} = -0{,}05$.

W kontekście uzyskanych wyników wspomnieć trzeba o pojawiającej się wśród ekonomistów koncepcji konwergencji klubowej. Polega ona na zbliżaniu się do siebie regionów o podobnych charakterystykach strukturalnych i podobnych warunkach początkowych, które tworzą „kluby" coraz bardziej różniące się od innych „klubów". Zbadanie, czy proces konwergencji klubowej dokonuje się w Polsce, może być przedmiotem dalszej analizy. Próby jej przeprowadzenia na podstawie dynamicznych modeli panelowych zostały podjęte, niestety nie udało się uzyskać zadowalających rezultatów. Wydaje się, że weryfikacja hipotezy o konwergencji klubowej w Polsce wymagałaby zastosowania innych narzędzi niż modele ekonometryczne. Powodem jest zbyt mała liczba województw należących do wyodrębnionych „klubów", która powoduje problemy metodologiczne, związane z obciążeniem estymatorów charakterystycznym dla małych prób. Być może właściwą metodą byłaby stosowana w artykule P. W ó j c i k a [2004] metoda oparta na procesach Markowa i estymacji pełnej warunkowej funkcji gęstości. Jej zastosowanie wykracza jednak poza ramy niniejszej pracy.

W celu wstępnego rozpoznania, czy hipoteza konwergencji klubowej na szczeblu województw jest prawdopodobna, zastosowano metodę analizy skupień (ang. *cluster analysis*). Jest to narzędzie do eksploracyjnej analizy danych, której celem jest wyodrębnienie grup podobnych do siebie obiektów. Obiekty łączy się w grupy w taki sposób, aby stopień powiązania dowolnego obiektu z obiektami należącymi do tej samej grupy był jak największy, a z obiektami z pozostałych grup – jak najmniejszy. Początkowo każdy obiekt stanowi swoją własną klasę. Następnie, osłabiając kryteria przypisania obiektów do tego samego skupienia, tworzy się kolejne, większe skupienia, coraz bardziej zróżnicowanych elementów. Na ostatnim etapie, wszystkie obiekty zostają ze sobą połączone.

Na rys. 6.3 przedstawiono schemat łączenia 16 województw w grupy. Jest to tzw. wykres drzewkowy, w którym na osi poziomej odłożone są odległości aglomeracyjne. Przy każdym węźle, gdzie uformowało się nowe skupienie, można odczytać odległość, w której odpowiednie elementy zostały powiązane ze sobą, tworząc nowe skupienie. Grupowania dokonano na podstawie zmiennych, które uwzględnione były w modelu konwergencji (6.11). Przyjęto bieżące lub opóźnione wartości kolejnych zmiennych, uwzględniając to, jakie wartości występują w modelu. Metodę analizy skupień stosowano dwukrotnie: na podstawie danych z roku 2001 – pierwszego roku próby, dla którego znane były opóźnione wartości zmiennych, i na podstawie danych z roku 2007 – ostatniego roku próby.



Rys. 6.3. Grupowanie województw ze względu na zmienne: $PKBpc_{i,t-1}$, $glud_{it}$, $inwpr_{i,t-1}$
Ź r ó d ł o: opracowanie własne.

Porównanie efektów grupowania województw dla lat 2001 i 2007 pozwala stwierdzić kilka zmian. Podstawową jest oddalenie się województwa mazowieckiego, które w 2007 r. utworzyło odrębne, odległe skupienie – odległość wiązania wynosi ok. 28 000 jednostek. W roku 2001 województwo to było bardziej zbliżone do skupienia, utworzonego przez województwa: dolnośląskie, zachodniopomorskie, pomorskie, śląskie i wielkopolskie – jego odległość od tego skupienia wynosiła ok. 15 000 jednostek. Ponadto średnie odległości wiązania dla tzw. skupień trzeciego rzędu, powstających w trzecim kroku łączenia województw, zwiększyły się. W roku 2001 w odległości 2000 utworzyły się już dwa duże skupienia trzeciego rzędu (województwa: dolnośląskie, zachodniopomorskie, pomorskie połączone zostały z województwami: śląskim i wielkopol-

skim, a województwa: kujawsko-pomorskie, lubuskie i łódzkie – z województwami: małopolskim i opolskim). Natomiast w roku 2007 w odległości 2500 jednostek utworzone było dopiero jedno skupienie trzeciego rzędu, składające się zaledwie z czterech województw: kujawsko-pomorskie i lubelskie połączone zostało z małopolskim i opolskim. Dodatkowo, o ile w 2001 r. w odległości ok. 15000 widoczne były dwa zasadnicze skupienia, z których pierwsze tworzyły województwa dolnośląskie, zachodniopomorskie, pomorskie, śląskie, wielkopolskie i mazowieckie, a drugie – pozostałe województwa, to w 2007 r. w tej samej odległości można wyodrębnić aż cztery takie skupienia: 1) dolnośląskie, wielkopolskie, śląskie, 2) kujawsko-pomorskie, lubuskie małopolskie, opolskie, łódzkie, zachodniopomorskie, pomorskie, 3) lubelskie, podkarpackie, podlaskie, świętokrzyskie, warmińsko-mazurskie, 4) mazowieckie. Przeprowadzona powyżej analiza aglomeracyjna wydaje się wskazywać na duże prawdopodobieństwo tworzenia się w Polsce grup województw podobnych, oddalających się coraz bardziej od pozostałych grup, a więc procesu konwergencji klubowej.

Można przypuszczać, że przyczyną zjawiska dywergencji lub konwergencji klubowej mogą być różnice poziomu infrastruktury, efekty aglomeracji i różnice w zasobach kapitału ludzkiego, pogłębiane przez migracje między województwami i zagraniczne. W kolejnym punkcie prezentowana jest próba uwzględnienia w modelu kształtowania się PKB *per capita* dodatkowych czynników, które wpływają na wzrost gospodarczy województw.

6.5.3. Próba uwzględnienia kapitału ludzkiego w modelu wzrostu województw w Polsce

Jak wspomniano w pkt 6.5.1, konstrukcja miar kapitału ludzkiego możliwa była od roku 2002. Postać modelu, uwzględniającego kapitał ludzki, charakteryzującego się najlepszymi własnościami merytoryczno-statystycznymi, jest następująca:

$$\begin{aligned}\ln(PKBpc_{it}) = \alpha_0 + (1-\gamma)\ln(PKBpc_{i,t-1}) + \theta_1 \ln(glud) + \theta_2 \ln(hk_{it}) + \\ + \theta_3 \ln(tup_{i,t-1}) + \theta_4 \ln(sinw_{i,t-1}) + \alpha_i + \varepsilon_{it}\end{aligned} \tag{6.12}$$

Model (6.12) estymowany był, podobnie jak model (6.11), za pomocą jednostopniowych FDGMM i SGMM, przy czym przyjmowano alternatywne założenia o endogeniczności (lub egzogeniczności) zmiennych objaśniających. Ostatecznie uznano, że zmienną nieegzogeniczną jest jedynie zmienna *sinw*. W tab. 6.3 przedstawione są wyniki estymacji modelu (6.12) uzyskane przy zastosowaniu wspomnianych powyżej metod. Oznaczenia „endo”, „pre” i „egzo” są analogiczne, jak w tab. 6.2.

Tabela 6.3

Wyniki estymacji modelu (6.12)

Zmienna objaśniająca/ test statystyczny	Metoda estymacji					
	SGMMendo	SGMMpre	SGMMegzo	FDGMMendo	WG	KMNK
	1	2	3	4	5	6
$\ln(PKBpc_{i,t-1})$	0,8939*** (0,1018)	0,8262*** (0,1071)	0,7838*** (0,1461)	0,8923*** (0,1267)	0,8492*** (0,0865)	0,9897*** (0,0302)
$\hat{\gamma}$	0,1061	0,1738	0,2162	0,1077	0,1508	0,0103
$\hat{\beta}$	0,1687	–	–	0,1731	0,3917	0,0106
$\ln(glud_{it})$	–0,4462*** (0,2078)	–0,4513** (0,2611)	–0,5937*** (0,2053)	–0,3970** (0,2168)	–0,3374 (0,2408)	–0,0934 (0,0619)
$\ln(hk_{it})$	0,1539** (0,0845)	0,1149*** (0,0488)	0,1083** (0,0571)	0,1361 (0,1357)	0,1570 (0,0947)	0,0026 (0,0081)
$\ln(tup_{i,t-1})$	0,1232*** (0,0585)	0,0862 (0,0739)	0,1562** (0,0815)	0,1228** (0,0658)	0,1417*** (0,0562)	0,0175 (0,0228)
$\ln(sinw_{i,t-1})$	0,1208** (0,0674)	0,0823** (0,0476)	0,1040** (0,0405)	0,0993** (0,0618)	0,0915*** (0,0406)	0,0531*** (0,0258)
const.	–0,9809 (1,1408)	–0,5527 (1,4736)	–0,7589 (1,2075)	–1,4239 (1,1286)	–1,0716 (0,9579)	–0,1221 (0,2835)
m_1 [*p-value*]	–2,383 [0,017]	–2,137 [0,033]	–2,354 [0,019]	–2,429 [0,015]	–	–
m_2 [*p-value*]	–0,448 [0,654]	–0,467 [0,641]	–0,735 [0,462]	–0,523 [0,601]	–	–
Sargan [*p-value*]	32,450 [0,053]	33,008 [0,162]	24,936 [0,024]	24,746 [0,037]	–	–

Tabela 6.3 (cd.)

Zmienna objaśniająca/ test statystyczny	Metoda estymacji					
	SGMM endo	SGMM pre	SGMM egzo	FDGMM endo	WG	KMNK
	1	2	3	4	5	6
dif-Sargan [*p-value*]	7,704 [0,359]	7,148 [0,521]	5,465 [0,243]	–	–	–
R^2	–	–	–	–	0,946	0,989

Wartość β określa szybkość konwergencji i jest obliczona ze wzoru (6.9) na podstawie oszacowanej wartości parametru $1 - \gamma$ przy zmiennej $\ln(PKBpc_{i,t-1})$;

symbol *** przy ocenie parametru oznacza, że wartość $p < 0{,}05$ dla testu *t*-Studenta; symbol ** oznacza $p \leq 0{,}1$; pozostałych zmiennych nie można uznać za istotne;

w nawiasach, pod ocenami parametrów zamieszczono błędy szacunku parametrów, wyznaczane na podstawie odpornych estymatorów wariancji składnika losowego;

m_1 i m_2 oznaczają wartości empiryczne testu autokorelacji Arellano–Bonda, odpowiednio AR(1) i AR(2); w nawiasach kwadratowych – wartości *p* dla H_0: autokorelacja pierwszego (drugiego) rzędu nie występuje;

Sargan i *dif*-Sargan oznaczają wartości empiryczne testu Sargana i różnicowego testu Sargana; w nawiasach kwadratowych – wartości *p* dla H_0: warunki ponadidentyfikujące są poprawne (instrumenty są właściwe).

Ź r ó d ł o: obliczenia własne.

Podobnie jak w przypadku wyników prezentowanych w tab. 6.2, najbardziej poprawną metodą estymacji modelu kształtowania się PKB *per capita* jest SGMM z założeniem endogeniczności nakładów inwestycyjnych (miarą wielkości inwestycji jest w modelu (6.12) zmienna *sinw* – stopa inwestycji, liczona jako udział nakładów inwestycyjnych w PKB). Za poprawnością tej metody estymacji przemawiają dwa fakty: a) ocena parametru przy opóźnionej zmiennej objaśnianej zawiera się pomiędzy oceną WG i oceną KMNK oraz b) właściwość instrumentów, stwierdzona za pomocą testów Arellano–Bonda i Sargana. W przypadku pozostałych metod wyniki testów nie są tak jednoznacznie pozytywne. Test Arellano–Bonda stwierdza brak autokorelacji drugiego rzędu niezależnie od metody estymacji, ale test Sargana stwierdza poprawność warunków ponadidentyfikujących tylko w przypadku SGMM, o ile stopa inwestycji jest traktowana nie egzogenicznie (kolumny 1 i 2 w tab. 6.3). Z kolei ocena parametru przy opóźnionej zmiennej objaśnianej zawiera się pomiędzy ocenami WG i KMNK jedynie, gdy stopa inwestycji traktowana jest endogenicznie a metodą estymacji jest SGMM (kolumna 1) lub FDGMM (kolumna 4). Dalsze rozważania dotyczące modelu (6.12) ograniczone zatem zostaną do wyników uzyskanych na podstawie SGMM z endogeniczną zmienną *sinw*, prezentowanych w kolumnie 1 tab. 6.3.

Wszystkie zmienne uwzględnione w modelu wpływają istotnie na kształtowanie się PKB *per capita*. Zgodnie z oczekiwaniami, wpływ wszystkich zmiennych objaśniających, za wyjątkiem stopy wzrostu liczby ludności (*glud*), jest pozytywny. Relatywnie duże znaczenie dla wzrostu PKB *per capita* ma wzrost kapitału ludzkiego – elastyczność wynosi 0,15% przy innych czynnikach niezmienionych. Kapitał ludzki wpływa więc na wzrost *PKBpc* silniej niż wzrost stopy inwestycji i technicznego uzbrojenia pracy (elastyczności ok. 0,12%). Oszacowane wartości parametrów strukturalnych są ekonomicznie sensowne. Ponadto oceny parametrów przy zmiennych *glud* i *sinw* zbliżone są do ocen uzyskanych w publikacji: B o n d, H o e f f l e r, T e m p l e [2001], dla modelu wzrostu i konwergencji krajów OECD, który zawierał tylko te dwie zmienne objaśniające.

Wątpliwości natomiast może budzić ocena parametru autoregresyjnego. Jej wartość $1 - \hat{\gamma} = 0{,}8939$ jest niższa niż uzyskiwana w badaniach konwergencji. Wynikająca stąd ocena tempa konwergencji byłaby równa $\hat{\beta} = 0{,}1687$, co oznaczałoby, że wyrównywanie poziomu dochodów między województwami zachodzi w tempie prawie 17% rocznie. Po pierwsze, tak szybkie tempo konwergencji jest niemożliwe, po drugie, jest ono sprzeczne z wynikami uzyskanymi na podstawie modelu (6.11), oraz z wynikami uzyskanymi na podstawie analizy danych: współczynnik korelacji między roczną stopą wzrostu PKB *per capita* a opóźnionym poziomem tej cechy w województwach w latach 2002–2007 $r(PKBpc_{i,t-1}, gPKBpc_{it})$ ma wartość wynoszącą 0,1280. Dodatni znak tego współczynnika sugeruje występowanie rozbieżnych tendencji w kształto-

waniu się PKB *per capita* w województwach. Wydaje się więc, że model (6.12) służyć może raczej do opisu czynników, od których zależny jest wzrost PKB *per capita*, ale nie można go interpretować jako modelu konwergencji. Zauważyć warto, że próby wprowadzenia do zbioru zmiennych objaśniających w modelach dynamicznych miar kapitału ludzkiego podejmowane przez innych autorów również się nie powiodły. Dwie takie próby opisane są w niniejszym rozdziale – w podrozdz. 6.3 i 6.4. W celu wyjaśnienia przyczyn tych niepowodzeń, warto przytoczyć hipotezę sformułowaną w pracy B o n d a, H o e f f l e r i T e m p l e'a [2001], że oddziaływanie wskaźnika skolaryzacji (lub innej miary kapitału ludzkiego) na wzrost gospodarczy dokonuje się za pośrednictwem stopy inwestycji.

6.6. Poziom wykształcenia społeczeństw a wskaźnik demokracji

Nieco odmiennym przykładem zastosowania dynamicznych modeli panelowych w analizach regionalnych jest badanie wpływu poziomu wykształcenia społeczeństwa na rozwój demokracji. Jest to jeden z przykładów pokazujących, że potrzeba i możliwości zastosowania dynamicznych modeli panelowych wykraczają poza obszar ekonomii, wchodząc w obszar nauk politycznych i społecznych[9]. Wpływ wykształcenia na poziom demokracji badany był przez D. A c e m o g l u i in. [2005], oraz M. B o b b ę i D. C o v i e l l o [2007]. Autorzy obu prac posługują się tym samym zbiorem danych ze 108 krajów dla okresów 5-letnich z lat 1960–2000. Szacowany model ma postać:

$$d_{it} = \gamma d_{i,t-1} + \delta s_{i,t-1} + \boldsymbol{\beta}^{\mathrm{T}} \mathbf{x}_{i,t-1} + \alpha_i + \lambda_t + \varepsilon_{it} \tag{6.13}$$

gdzie d_{it} oznacza poziom demokracji w kraju i w okresie t. Jest to wartość *freedom house political rights index*, transformowana w taki sposób, aby wartości, oryginalnie od 1 do 7, znajdowały się w przedziale od 0 do 1, przy czym wartość 1 odpowiada krajowi najbardziej demokratycznemu. Zmienna $s_{i,t-1}$ oznacza opóźnioną wartość przeciętnej liczby lat pobierania nauki obliczoną dla populacji powyżej 25. roku życia, $\mathbf{x}_{i,t-1}$ jest wektorem pozostałych zmiennych objaśniających (opóźnionych), α_i, λ_t są efektami grupowymi i czasowymi.

W naukach społecznych przyjmuje się powszechnie, że rosnący poziom wykształcenia sprzyja rozwojowi demokracji, ponieważ podnosi tzw. „kulturę demokracji", zmniejsza przyzwolenie dla nurtów ekstremistycznych i prowadzi

[9] Obszerne uzasadnienie konieczności stosowania modeli panelowych, również dynamicznych w naukach socjologicznych przedstawia C. H a l a b y [2004]. Nieco ogólniejsze rozważania o potrzebie stosowania metod ekonometrycznych, w tym również modeli panelowych, w naukach społecznych zawarte są w książce R. Ta r l i n g a [2008].

do wzrostu poziomu życia, który przyczynia się do rozwoju społeczno-politycznego. Przykładem możliwości potwierdzenia tych tez za pomocą metod ekonometrycznych jest cytowana przez M. B o b b y'a i D. C o v i e l l o [2007] praca G l a e s e r a [2004]. Uzyskane przez niego na podstawie KMNK wyniki potwierdziły wpływ wykształcenia na rozwój demokracji. D. A c e m o g l u i in. [2005] pokazali, że po wprowadzeniu do modelu nielosowych efektów grupowych zmienna $s_{i,t-1}$ staje się nieistotna statystycznie. Do estymacji modelu (6.13) zastosowali oni dwustopniową FDGMM, traktując opóźniony poziom wykształcenia jako zmienną egzogeniczną. M. B o b b a i D. C o v i e l l o [2007] argumentują, że ze względu na długookresowy charakter inwestycji w kapitał ludzki poziom wykształcenia powinien być traktowany endogenicznie. Ponadto, ponieważ poziom edukacji i poziom wykształcenia charakteryzują się wysokim stopniem trwałości, poziomy tych zmiennych, wykorzystane przez D. A c e m o g l u i in. [2005] w FDGMM są słabymi instrumentami dla odpowiednich przyrostów. Właściwą metodą estymacji modelu (6.13) jest zatem SGMM. Jej zastosowanie wymaga, aby spełnione były dodatkowe warunki momentów postaci:

$$\mathrm{E}(\Delta d_{i,t-1}(\alpha_i + \varepsilon_{it})) = 0 \quad \text{dla } t = 4, 5, ..., T \tag{6.14}$$

Sprawdzenie, czy warunki te są spełnione, jest możliwe po oszacowaniu modelu za pomocą różnicowego testu Sargana.

W pracyM. B o b b y i D. C o v i e l l o [2007] oszacowano najpierw model, w którym na poziom demokracji wpływa tylko opóźniony poziom tej zmiennej i opóźniony poziom wykształcenia, a więc model (6.13), w którym nie występuje wektor $\mathbf{x}_{i,t-1}$. Szczegółowe wyniki przedstawiają w tab. 1 (s. 304). Zastosowano kilka alternatywnych metod estymacji: *pooled* KMNK, estymator WG, jedno- i dwustopniową FDGMM, oraz jedno- i dwustopniową SGMM, przy czym błędy szacunku dwustopniowej GMM liczone były na podstawie korekty Windmeijera (wzór (3.46)). Powodem zastosowania dwóch pierwszych wymienionych metod jest określenie przedziału, w którym powinna znajdować się nieobciążona wartość parametru autoregresyjnego[10]. Oceny parametru γ uzyskane FDGMM są w obu przypadkach niższe niż oceny WG, a zatem są one obciążone w dół. Ponadto błędy szacunku parametru przy zmiennej $s_{i,t-1}$ w obu przypadkach przekraczają jego wartość, co świadczyłoby o braku wpływu poziomu wykształcenia na poziom demokracji. Co więcej, dwustopniowa FDGMM daje ujemną ocenę tego parametru.

Zdecydowaną poprawę wyników estymacji uzyskano po zastosowaniu SGMM. Ocena parametru γ wynosi 0,45 dla jednostopniowej i 0,506 dla dwustopniowej SGMM i znajduje się pomiędzy ocenami WG (0,387) i KMNK (0,703). Do oceny istotności wpływu poszczególnych zmiennych na poziom

[10] Na ten temat pisano w końcowej części pkt 3.2.2.

demokracji autorzy posługują się statystykami *t*-Studenta. W ogólnym przypadku, zastosowanie GMM wyklucza posługiwanie się tym testem, gdyż GMM nie wymaga założenia o normalności rozkładu składnika losowego. Wartości empiryczne statystyk *t*-Studenta można wówczas traktować pomocniczo do określenia stosunku oceny parametru do jego błędu, bez odnoszenia ich do stablicowanych wartości krytycznych. Jeśli jednak normalność zostanie założona dodatkowo, klasyczne wnioskowanie jest uzasadnione. Z tym zastrzeżeniem, zarówno $d_{i,t-1}$ jak i $s_{i,t-1}$ wpływają na poziom demokracji, przy czym ocena parametru przy $s_{i,t-1}$ uzyskana jednostopniową SGMM wynosi 0,099 i oznacza, że dodatkowy rok nauki podnosi wartość wskaźnika demokracji w długim okresie o 18 pkt proc. (wartość tę obliczono jako $\gamma/(1-\delta)$). Wyniki testów, badających poprawność warunków momentów wykorzystanych w FDGMM i SGMM (test Arellano–Bonda i testy Sargana), w żadnym z przypadków nie dają podstaw do odrzucenia H_0 o ich poprawności. Autorzy wysunęli przypuszczenie, że na wynik różnicowego testu Sargana może mieć wpływ to, że badana próba obejmuje m. in. kraje OECD, w których poziom demokracji jest wysoki i stały, a zatem zmiana wskaźnika d_{it} jest równa zero i warunek (6.14) jest trywialny. Jednak po wykluczeniu tych krajów z próby wynik testu się nie zmienił.

M. Bobba i D. Coviello [2007] zauważyli ponadto, że ocena współczynnika przy $s_{i,t-1}$, oprócz bezpośredniego wpływu tej zmiennej na d_{it}, może wyrażać wpływ pośredni, związany np. ze wzrostem dochodów (uzyskanym dzięki wyższemu poziomowi wykształcenia). Dlatego przeprowadzono badanie odporności wyników, poprzez dołączenie do modelu dodatkowych zmiennych objaśniających: logarytmu PKB *per capita*, logarytmu stopy inwestycji i logarytmu liczby ludności, opóźnionych o jeden okres. Najpierw zmienne te dołączane były pojedynczo, a następnie jednocześnie. Każdy z czterech powstałych modeli szacowany był dwustopniową SGMM. Wyniki, zawarte w tab. 2 (s. 305) w tekście oryginalnym, wskazują na brak istotnego wpływu tych zmiennych na wskaźnik demokracji, niezależnie od tego, czy występują one w modelu osobno czy jednocześnie. Wpływ poziomu wykształcenia pozostaje zaś istotny za wyjątkiem modelu, w którym dodatkową zmienną jest PKB *per capita*. Autorzy wyciągnęli stąd wniosek, że pierwszorzędny, znaczący wpływ na wskaźnik demokracji wywiera bezpośrednio jakość kapitału ludzkiego. Efekt wpływu pośredniego, wywieranego przez związany z poziomem wykształcenia wzrost zamożności ma znaczenie drugorzędne.

ZAKOŃCZENIE

Weryfikacja hipotez formułowanych we współczesnych badaniach ekonomicznych niejednokrotnie wymaga zastosowania nowoczesnych, skomplikowanych metod analitycznych. Jedną z możliwości jest zastosowanie modeli ekonometrycznych estymowanych na podstawie danych czasowo-przekrojowych, a w szczególności danych panelowych. Ich wykorzystanie stwarza możliwości przeprowadzenia wielowymiarowej analizy zjawisk ekonomicznych. Cechą wyróżniającą dynamiczne modele panelowe spośród innych narzędzi analitycznych jest niewątpliwie ich specyficzna konstrukcja, która umożliwia zbadanie wpływu na zmienną objaśnianą czynników specyficznych dla poszczególnych obiektów oraz czynnika czasu. Czynniki specyficzne dla obiektów uwzględnia się w modelu poprzez wprowadzenie, charakterystycznych dla obiektów, stałych w czasie efektów grupowych, zaś czynnik czasu może być uwzględniony poprzez wprowadzenie zmiennych z opóźnieniami czasowymi, a w szczególności opóźnionej zmiennej endogenicznej. Niewątpliwą zaletą tej metodologii jest jej uniwersalność. Możliwości zastosowania modeli panelowych, zarówno statycznych, jak i dynamicznych, obejmują bowiem analizy bardzo różnorodnych zjawisk ekonomicznych. Wykorzystanie danych panelowych lub czasowo-przekrojowych w procesie konstrukcji i estymacji modeli ekonometrycznych dostarcza korzyści w czterech głównych aspektach:

– powoduje wzrost liczby stopni swobody i redukuje problem współliniowości danych;

– ułatwia identyfikowalność modeli ekonomicznych i dokonanie wyboru między konkurencyjnymi hipotezami ekonomicznymi;

– umożliwia eliminację lub redukcję obciążenia estymatorów;

– dostarcza podstaw w skali mikro dla analiz prowadzonych na podstawie danych zagregowanych.

Z zastosowaniem danych panelowych do estymacji modeli ekonometrycznych wiąże się jednak wiele komplikacji metodologicznych, niewystępujących w modelowaniu na podstawie danych czasowych lub przekrojowych. Wynikają one z wielowymiarowości danych. W szczególności estymacja modeli dynamicznych na podstawie danych panelowych wymaga zastosowania metod estymacji, które różnią się od metod stosowanych do estymacji statycznych modeli panelowych.

Metody estymacji dynamicznych modeli panelowych rozwijane są na świecie szczególnie intensywnie od początku lat 90. XX w., przy czym liczba publikowanych opracowań anglojęzycznych z tej dziedziny jest wciąż bardzo duża. Na tym tle stan badań w Polsce prezentuje się skromnie.

Zadaniem tej monografii było więc wypełnienie istniejącej luki. Dokonano w niej przeglądu zagadnień o charakterze metodologicznym i praktycznym, związanych z wykorzystaniem danych panelowych do konstrukcji, estymacji i wnioskowania na podstawie modeli dynamicznych zarówno od strony teoretycznej i historycznej, jak i najnowszej metodologii. Szczególny nacisk położono na analizę właściwości estymatorów dynamicznych modeli panelowych pod kątem ich przydatności do analiz empirycznych. Krytyczne porównanie metod stosowanych do estymacji dynamicznych modeli panelowych było głównym celem badawczym niniejszej monografii.

W kolejnych rozdziałach zaprezentowano rozwój metodologii, od opracowań o charakterze historycznym do najbardziej współczesnych metod estymacji i wnioskowania. Omówiono też wady i zalety alternatywnych metod estymacji w zależności od własności szeregów czasowych obserwacji dla obiektów. W pracy niniejszej wskazano też oprogramowanie komputerowe, za pomocą którego możliwe jest prowadzenia analiz empirycznych na podstawie danych panelowych, a także bazy, z których czerpać można dane łączące wymiar czasowy i przekrojowy. Krytyczne omówienie i porównanie własności alternatywnych metod może być cenną wskazówką, zachęcającą do praktycznego zastosowania dynamicznych modeli panelowych w analizach ekonomicznych dotyczących różnorodnych zjawisk. Zasadniczy cel monografii wiąże się z modelowaniem dynamicznym na podstawie danych panelowych, jednak problematyka ta została osadzona w szerszym kontekście – w pracy przedstawione zostały również metody estymacji modeli statycznych jedno- i wielorównaniowych oraz zagadnienia dotyczące modelowania na podstawie danych, w których wymiar czasowy jest większy niż wymiar przekrojowy. W tym drugim wypadku do analiz stosuje się często podejście znane z analiz szeregów czasowych. W niniejszej monografii opisano metody estymacji modeli bazujących na tego typu danych, uwzględniające występowanie hereroskedastyczności grupowej i/lub korelacji przekrojowej oraz autokorelacji, jak również testowanie stacjonarności w panelach o odpowiednio długim wymiarze czasowym.

Oprócz kompleksowego i krytycznego przedstawienia zagadnień o charakterze metodologicznym, w niniejszej publikacji dużo uwagi poświęcono możliwościom praktycznego zastosowania dynamicznych modeli panelowych. Zaprezentowane zostały przykłady praktycznego zastosowania dynamicznych modeli panelowych w modelowaniu zjawisk mikro-, mezo- i makroekonomicznych. W doborze przykładów kierowano się zamiarem pokazania różnorodności obszarów badawczych, w których znalazły zastosowanie modele dynamiczne estymowane na podstawie danych panelowych. Przykłady badań z obszaru

ekonomii przemysłowej i finansów przedsiębiorstw dobrano tak, aby obrazowały korzyści badawcze płynące z zastosowania mikropaneli dla analizy ważnych aspektów funkcjonowania przedsiębiorstw: inwestycji, produktywności, zatrudnienia, innowacji, czynników finansowych, transferu kwalifikacji czy struktury zarządu firmy. Tematyka zaprezentowanych analiz sektorowych oscyluje wokół zagadnień związanych z produktywnością, a w szczególności łączną produktywnością czynników produkcji. Obok przykładów zaczerpniętych z literatury światowej, dotyczących modelowania produktywności i jej determinant przedstawiono również wyniki własnych analiz ekonometrycznych, dotyczących możliwości oszacowania i modelowania łącznej produktywności czynników produkcji według działów sekcji „przetwórstwo" w Polsce. Korzyści z zastosowania dynamicznych modeli panelowych w badaniach regionalnych ukazane zostały na przykładach wyników estymacji modeli wzrostu, konwergencji i produktywności w regionach. Są to zarówno przykłady zaczerpnięte z literatury światowej, jak i analizy własne, dotyczące konwergencji województw.

Jak wynika z powyższego opisu, własne przykłady empiryczne prezentowane w niniejszej monografii wpisują się w nurt analiz sektorowych i regionalnych, co jest odzwierciedleniem wieloletnich zainteresowań badawczych autorki. Ich tematyka koncentruje się na zagadnieniach związanych ze wzrostem gospodarczym i produktywnością. Są to obszary ściśle ze sobą związane. Wzrost łącznej produktywności czynników produkcji jest syntetycznym sposobem oceny zmian efektywności procesów produkcyjnych, zachodzących pod wpływem postępu technicznego. Postęp techniczny, wpływając na efektywność procesów produkcyjnych, wpływa również na rozwój i wzrost gospodarczy. Na gruncie neoklasycznego modelu wzrostu gospodarczego powstała koncepcja konwergencji warunkowej, czyli wyrównywania poziomu PKB między krajami lub regionami o zróżnicowanych początkowych wartościach tej zmiennej. Występowanie konwergencji regionalnej byłoby w dobie integracji europejskiej zjawiskiem pożądanym, dlatego część badań empirycznych poświęcona została zbadaniu, czy zjawisko to występuje na obszarze Polski.

Wyniki, uzyskane na podstawie własnych badań empirycznych, uznać należy za zadowalające. Zastosowanie dynamicznych modeli panelowych umożliwiło wskazanie czynników wpływających na kształtowanie się łącznej produktywności czynników produkcji zarówno według działów, jak i według województw. Analiza dynamicznych modeli wzrostu wykazała, że proces konwergencji wojewódzkiej w Polsce nie zachodzi. Uzyskane wyniki mogą, jak się wydaje, stanowić źródło informacji dla ekspertów i instytucji gospodarczych o sytuacji w analizowanych jednostkach, a prezentowane modele mogą stać się atrakcyjnym narzędziem badań.

DODATEK I

TESTY PIERWIASTKÓW JEDNOSTKOWYCH DLA DANYCH PANELOWYCH

Wraz z rosnącą liczbą zastosowań danych panelowych do analiz makroekonomicznych (patrz podrozdz. 1.2) uwaga ekonometryków zajmujących się modelowaniem na podstawie takich danych przesunęła się w kierunku asymptotycznych własności estymatorów przy $N \to \infty$ i $T \to \infty$ (analizy na podstawie mikropaneli koncentrują się na przypadku, gdy $N \to \infty$, a T jest ustalone). Ze względu na założenie, że $T \to \infty$, do analizy danych panelowych zaczęto stosować podejście znane z analiz szeregów czasowych, w szczególności rozwinęły się metody badania stacjonarności w panelach, regresji pozornych i kointegracji. Jako pierwsi badania nad własnościami statystyk testów pierwiastka jednostkowego dla danych panelowych przeprowadzili A. Levin i C. F. Lin [1992][1]. Jednym z ważniejszych wyników ich prac jest to, że w odróżnieniu do testów dla szeregów czasowych testy panelowe mają asymptotyczny rozkład normalny, przy czym zbieżność do rozkładu tego jest szybsza dla $T \to \infty$ niż dla $N \to \infty$.

Liczba praktycznych zastosowań metod analizy szeregów czasowych w odniesieniu do danych panelowych, szczególnie w zakresie testów pierwiastków jednostkowych i estymacji przeciętnych długookresowych relacji rośnie. Odnośniki do odpowiedniej literatury znaleźć można w: Baltagi [2008], s. 274. Przykładami prac przeglądowych z tej dziedziny mogą być pozycje: Phillips, Moon [2000]; Baltagi, Kao [2000]; Baltagi [2008] (rozdz. 12, s. 273–310), Breitung, Pesaran [2007]; Hsiao [2003] (podrozdz. 10.2, s. 298–302).

Poniżej zaprezentowany zostanie przegląd testów pierwiastków jednostkowych dla danych panelowych. Podzielić je można na dwie grupy: testy zakładające niezależność przekrojową i testy dopuszczające zależność miedzy jednostkami przekroju. Jak stwierdzili P. C. B. Philips i H. Moon [1999], określenie schematu, według którego każdy z wymiarów próby: N i T zbiegają do nieskończoności, jest kluczowe dla wnioskowania o asymptotycznych własnościach estymatorów i konstrukcji testów stacjonarności w panelach. Wyróżnili oni trzy przypadki:

---

[1] Więcej na temat propozycji zawartych w tej pracy, jak również omówienie i przykład praktycznego zastosowania innych testów znaleźć można w pracy K. Strzały [2005].

1) granic sekwencyjnych (ang. s*equential limits*): zakłada się, że najpierw zbiega do nieskończoności jeden z wymiarów, zazwyczaj T, a N jest ustalone. W ten sposób wyznacza się „pośrednią” wartość granicy; następnie dopuszcza się, że również $N \to \infty$. Tego typu wnioskowanie jest łatwe rachunkowo, jednak może dać mylące rezultaty asymptotyczne;

2) granic ścieżki diagonalnej (ang. *diagonal-path limits*): nakłada się restrykcje na relatywne tempa zbieżności obu wymiarów do nieskończoności. Ścieżka diagonalna może być określona przez funkcję monotoniczną typu $T = T(N)$ przy $N \to \infty$. Własności asymptotyczne zależą od postaci funkcyjnej, a założona ścieżka może nie być wystarczająco dobrą aproksymantą w konkretnym przypadku (T, N);

3) granic łącznych (ang. *joint limits*): zakłada się, że N i T dążą do nieskończoności jednocześnie. Rezultaty są na ogół bardziej odporne niż w poprzednio wymienionych przypadkach, jednak procedura ta ma następujące wady: a) łączne granice są trudniejsze do wyznaczenia, b) konieczne są silniejsze założenia (istnienie momentów wyższego rzędu) i c) granice sekwencyjne nie zawsze są równe granicy łącznej.

Na ogół zakłada się, że $N, T \to \infty$ oraz $(N/T) \to 0$, co oznacza, że wymiar czasowy próby jest relatywnie duży w stosunku do wymiaru przekrojowego. Przykładami tego typu danych mogą być odpowiednio długie szeregi czasowe dla krajów, regionów, sektorów gospodarczych (np. dane zawarte w Penn Word Table). Możliwe jest też testowanie stacjonarności przy alternatywnych założeniach, pokazane m. in. przez: Quaha [1994]; Breitunga, Mayera [1994]; Philipsa, Moona [2000]; Choi [2001]. P. C. B. Philips i H. Moon [1999] zauważyli jednak, że z praktycznego punktu widzenia rezultaty uzyskane poprzez wyznaczanie granic sekwencyjnych będą adekwatne w większości przypadków.

D.I.1. Testy zakładające niezależność przekrojową

Jedne z pierwszych testów niestacjonarności w panelach przedstawione zostały w pracach: D. Quaha [1994] oraz J. Breitunga i W. Mayera [1994]. Przyjmowano tam założenie, że zmienne są niezależne i mają jednakowe rozkłady dla wszystkich obiektów. Quah [1994] założył, że $N, T \to \infty$ oraz $(N/T) = const.$ Breitung i Mayer [1994] badali przypadek typowy dla mikropaneli, tzn. gdy $T \to \infty$ a N jest niewielkie. Pokazali oni, że statystyka testu Dickey–Fullera ma wówczas rozkład asymptotycznie normalny. Na bazie wymienionych wyżej prac A. Levin, C. F. Lin i C. Chu [2002] skonstruowali jeden z ważniejszych testów niestacjonarności. Do grupy testów zakładających niezależność przekrojową, zwanych testami pierwszej generacji (np. Baltagi

[2008], s. 274) należą też testy I m a, P e s a r a n a i S h i n a [2003] i testy typu Fishera zaproponowane przez G. S. M a d d a l ę i S. W u [1999].

Testy pierwszej generacji bazują na regresji postaci:

$$\Delta y_{it} = \rho_i y_{i,t-1} + \mathbf{d}_{it}^{\mathrm{T}} \boldsymbol{\delta} + u_{it} \qquad \text{(D.I.1)}$$

gdzie $\mathbf{d}_{it}^{\mathrm{T}}$ jest wektorem zmiennych deterministycznych, zaś u_{it} jest procesem stacjonarnym. Hipoteza zerowa tych testów wygląda następująco: $\rho_i = 0$ dla wszystkich i. Podstawową różnicą miedzy wymienionymi testami jest zakładany stopień heterogeniczności.

Test Levina, Lina i Chu (LLC)

A. L e v i n, C. F. L i n i C. C h u [2002] uogólnili model Quaha dopuszczając heterogeniczność indywidualnych efektów deterministycznych, autokorelację składnika losowego (ze współczynnikiem różniącym się względem obiektów). Założyli przy tym, że parametr autoregresyjny jest jednakowy dla wszystkich obiektów (w modelu (D.I.1) $\rho_i = \rho$), zaś $N, T \to \infty$ oraz $(N/T) \to 0$.

Test LLC weryfikuje hipotezę zerową, że szereg obserwacji dla każdego obiektu posiada pierwiastek jednostkowy (H_0: $\rho = 0$ dla wszystkich i), wobec hipotezy alternatywnej o stacjonarności każdego szeregu (H_1: $\rho < 0$ dla wszystkich i). Test bazuje na regresji postaci:

$$\Delta y_{it} = \rho y_{i,t-1} + \sum_{L=1}^{p_i} \theta_{iL} \Delta y_{i,t-L} + \boldsymbol{\delta}_{mi} \mathbf{d}_{mi} + \varepsilon_{it}, \quad m = 1, 2, 3 \qquad \text{(D.I.2)}$$

gdzie $\mathbf{d}_{mt}$ jest wektorem zmiennych deterministycznych, a $\boldsymbol{\delta}$ wektorem parametrów dla modelu m. W szczególności $d_{1i} = \emptyset$, $d_{2t} = \{1\}$, $d_{3t} = \{1, t\}$. Ponieważ nieznany jest rząd opóźnienia p_i sugeruje się trójstopniową procedurę postępowania, która jest dokładnie opisana (oprócz tekstu źródłowego) np. w: B a l t a g i [2008] (s. 275–277). W pierwszym kroku wyznacza się wartość p_i na podstawie osobnych regresji ADF (ang. *augmented Dickey–Fuller*), a następnie, na podstawie dwóch regresji pomocniczych, wyznacza się reszty ortogonalne: $\hat{e}_{it}$ z regresji Δy_{it} na $\Delta y_{i,t-L}$ i $\mathbf{d}_{mt}$ i $\hat{v}_{i,t-1}$ z regresji $y_{i,t-1}$ na ten sam zbiór. W wyniku standaryzacji tych reszt otrzymuje się wartości $\tilde{e}_{it}$ i $\tilde{v}_{i,t-1}$. W drugim kroku estymuje się stosunek odchyleń długookresowych do krótkookresowych, a w ostatnim kroku wyznacza się wartość skorygowanej statystyki t:

$$t_\rho^* = \frac{t_\rho - N\tilde{T}\hat{S}_N \hat{\sigma}_{\tilde{\varepsilon}}^{-2} \hat{\sigma}(\hat{\rho}) \mu_{m\tilde{T}}^*}{\sigma_{m\tilde{T}}^*} \tag{D.I.3}$$

gdzie: $t_\rho = \hat{\rho} / \hat{\sigma}(\hat{\rho})$ jest standardową statystyką testu t dla H_0: $\rho = 0$, $\mu_{m\tilde{T}}^*$ i $\sigma_{m\tilde{T}}^*$ są wartościami średniej i odchylenia standardowego stablicowanymi w tab. 2, w pracy: L e v i n, L i n, C h u [2002], $\tilde{T}$ jest średnią liczbą obserwacji w czasie dla obiektów panelu, $\hat{S}_N = \frac{1}{N}\sum_{i=1}^{N} \frac{\hat{\sigma}_{yi}}{\hat{\sigma}_{\varepsilon i}}$, $\hat{\sigma}_{yi}$ oznacza estymator jądrowy długookresowej wariancji obiektu i, a $\hat{\rho}$ jest estymatorem *pooled* modelu $\tilde{e}_{it} = \rho \tilde{v}_{i,t-1} + \tilde{\varepsilon}_{it}$. Statystyka (D.I.3) ma asymptotyczny rozkład normalny N(0, 1).

Asymptotyczna normalność (D.I.3) wymaga, aby ($\sqrt{N}$ /T) → 0, przy czym wymiar N jest monotonicznie rosnącą funkcją T. LLC sugerują stosowalność testu dla prób, w których $10 \leq N \leq 250$, a $25 \leq T \leq 250$. Dla większych wartości T proponują oni stosowanie standardowych testów pierwiastków jednostkowych osobno dla poszczególnych jednostek przekroju, a dla bardzo dużych wartości N i małych T – stosowanie procedur panelowych bez rozważania stacjonarności.

Zasadniczymi ograniczeniami testu LLC są: kluczowa dla jego przeprowadzenia rola założenia o niezależności przekrojowej i sformułowanie hipotez dotyczących wszystkich obiektów (dla wszystkich i). Ponadto konstrukcja testu zakłada, że ρ jest jednakowe dla wszystkich obiektów, co zostało zakwestionowane jako nierealistyczne m. in. przez G. S. M a d d a l ę [1998].

Test Ima, Pesarana i Shina (IPS)

K. S. I m, M. H. P e s a r a n i Y. S h i n [2003] rozluźniają założenia LLC, dopuszczając heterogeniczność parametru przy $y_{i,t-1}$, autokorelację reszt i zmieniając postać hipotezy alternatywnej, tzn. dopuszczając możliwość, aby tylko część obiektów miała pierwiastek jednostkowy. Zatem w teście IPS mamy: H_0: $\rho_i = 0$ dla wszystkich i, zaś

$$H_1: \begin{cases} \rho_i < 0 & \text{dla} \quad i = 1, 2, ..., N_1 \\ \rho_i = 0 & \text{dla} \quad i = N_1 + 1, ..., N \end{cases} \tag{D.I.4}$$

IPS proponują zamiast łączenia danych (ang. *pooling*) obliczenie średniej ze statystyk testu ADF dla poszczególnych jednostek przekroju osobno, przy czym

składniki losowe obserwacji w czasie dla każdego obiektu mogą wykazywać autokorelację, a jej charakterystyki mogą różnić się miedzy obiektami. Test bazuje na regresji postaci:

$$\Delta y_{it} = \rho_i y_{i,t-1} + \sum_{L=1}^{p_i} \theta_{iL} \Delta y_{i,t-L} + \boldsymbol{\delta}_{mi} \mathbf{d}_{mi} + \varepsilon_{it}, \quad m = 1, 2, 3 \qquad \text{(D.I.5)}$$

gdzie oznaczenia są takie, jak we wzorze (D.I.1). Statystyka *t*-bar jest średnią statystyk ADF dla poszczególnych obiektów:

$$\bar{t} = \frac{1}{N} \sum_{i=1}^{N} t_{\rho_i} \qquad \text{(D.I.6)}$$

gdzie t_{ρ_i} jest wartością statystyki *t* dla H_0: $\rho_i = 0$. Jeśli rząd opóźnienia p_i jest niezerowy przynajmniej dla części obiektów, statystyka (D.I.6) ma asymptotyczny rozkład normalny N(0, 1). W celu poprawnej standaryzacji statystyki $\bar{t}$ należy korzystać z tab. 3 (I m, P e s a r a n, S h i n [2003], s. 66), w której zawarte są wartości oczekiwane i odchylenia standardowe wyznaczone metodami Monte Carlo dla różnych wartości *T* i p_i. Jeśli $p_i = 0$ dla wszystkich *i*, to należy korzystać z wartości krytycznych określonych dla $\bar{t}$ w: I m, P e s a r a n, S h i n [2003] (tab. 2, s. 61). IPS pokazali ponadto, że jeśli rząd opóźnienia jest wystarczająco duży, to wyniki testu IPS w małych próbach są lepsze niż testu LLC.

J. B r e i t u n g [2000] wykazał, że testy LLC i IPS tracą moc, jeśli w modelu obecne są trendy specyficzne dla obiektów. Zaproponował on też test, którego moc jest wyższa niż poprzednio wymienionych testów. Jego trójstopniowa procedura jest zbliżona do procedury stosowanej w teście LLC, a statystyka empiryczna *t* dla H_0: $\rho = 0$ ma asymptotyczny rozkład normalny N(0, 1).

Test typu Fishera kombinacji *p-value*

G. S. M a d d a l a i S. W u [1999] i I. C h o i [2001] zaproponowali test pierwiastka jednostkowego typu Fishera oparty na kombinacji *p-value* dla statystyk empirycznych[2] obliczonych dla poszczególnych obiektów (oznaczonych symbolem p_i):

$$P = -2 \sum_{i=1}^{N} \ln p_i \qquad \text{(D.I.7)}$$

[2] Można wykorzystać test ADF lub inny test niestacjonarności.

Jeśli liczba obserwacji w czasie dla i-tego obiektu $T_i \to \infty$, a N jest skończone, to P ma rozkład χ^2 z $2N$ stopniami swobody.

Hipoteza alternatywna testu kombinacji *p-value* ma postać (D.I.4), co stanowi przewagę tego testu nad testem LLC. C h o i [2001] wskazał następujące zalety testu typu Fischera w stosunku do testu IPS: 1) test Fishera można stosować, jeśli panel nie jest zbilansowany, 2) dopuszczalne są różne długości opóźnień w indywidualnych testach niestacjonarności – ADF lub innych, 3) N może być skończone lub nie, 4) dopuszcza się, by różne obiekty miały składowe niestochastyczne różnych typów. C h o i [2001] zaproponował też zastosowanie modyfikacji statystyki (D.I.7) postaci:

$$P_m = \frac{1}{2\sqrt{N}} \sum_{i=1}^{N} (-2 \ln p_i - 2) \tag{D.I.8}$$

P_m zbiega asymptotycznie do rozkładu N(0, 1), jeśli $T_i \to \infty$, a następnie $N \to \infty$, tzn. zachodzi zbieżność sekwencyjna. Zastosowanie statystyki (D.I.8) zamiast (D.I.7) jest koniczne, jeśli $N \to \infty$, ponieważ wówczas graniczny rozkład statystyki P byłby zdegenerowany.

Test stacjonarności Hadri'ego

K. H a d r i [2000] zaproponował test, w którym weryfikowana jest hipoteza zerowa o tym, że wszystkie szeregi czasowe w panelu są niestacjonarne, wobec alternatywy, że w panelu występuje pierwiastek jednostkowy. Test bazuje na resztach KMNK regresji y_{it} na stałą r_{it} lub na stałą i trend:

$$y_{it} = r_{it} + \beta_i t + \varepsilon_{it} \tag{D.I.9}$$

gdzie $r_{it} = r_{i,t-1} + u_{it}$ jest procesem błądzenia losowego, ε_{it} i u_{it} mają jednakowe, niezależne rozkłady normalne ze stałymi wariancjami: σ_ε^2 i σ_u^2. Model (D.I.9) można zapisać w postaci:

$$y_{it} = r_{i0} + \beta_i t + \sum_{s=1}^{t} u_{is} + \varepsilon_{it} = r_{i0} + \beta_i t + v_{it}, \quad \text{gdzie} \quad v_{it} = \sum_{s=1}^{t} u_{is} + \varepsilon_{it} \tag{D.I.10}$$

Hipoteza zerowa o niestacjonarności ma więc postać H_0: $\sigma_u^2 = 0$. Statystyką empiryczną jest:

$$LM = \frac{1}{N}\left(\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}S_{it}^2\right)/\hat{\sigma}_{\varepsilon}^2 \qquad \text{(D.I.11)}$$

gdzie: $\hat{\varepsilon}_{is}$ są resztami równania (D.I.10), $S_{it} = \sum_{s=1}^{t}\hat{\varepsilon}_{is}$, $\hat{\sigma}_{\varepsilon}^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\varepsilon}_{it}^2$ jest zgodnym estymatorem σ_{ε}^2 przy założeniu prawdziwości H_0.

K. Hadri [2000] przedstawił też alternatywną wersję statystyki (D.I.11) dopuszczającą heteroskedastyczność grupową. Jej postać podają np.: Baltagi [2008], s. 283; Yin, Wu [2000].

D.I.2. Testy zakładające występowanie zależności przekrojowej

Podstawową wadą testów pierwszej generacji jest to, że ich konstrukcja zakłada, iż szeregi czasowe obserwacji dla poszczególnych obiektów panelu są przekrojowo niezależne. Jest to założenie mało realistyczne, gdyż literatura ekonomiczna dostarcza dowodów ich wzajemnych powiązań. Dlatego powstały tzw. testy drugiej generacji, dopuszczające zależność przekrojową[3]. Kluczowym problemem konstrukcji tych testów jest specyfikacja zależności przekrojowej, co jest utrudnione ze względu na brak naturalnego uporządkowania przekroju. Można wyróżnić dwa główne nurty w podejściu do tego zagadnienia:

1) minimalizacja liczby restrykcji nakładanych na macierz wariancji-kowariancji składnika losowego – jest ich niewiele lub wcale. Przykładami tego typu podejścia są m. in. prace: Maddali, Wu [1999]; Chang [2002, 2004],

2) modelowanie zależności przekrojowej na podstawie modelu wspólnych czynników, którego estymacji dokonuje się przed konstrukcją testu pierwiastka jednostkowego. Przykładami tego podejścia są m. in. prace: Moon, Perron [2004]; Philips, Sul [2003]; Pesaran [2007].

Minimalizacja liczby restrykcji

G. S. Maddala i S. Wu [1999] wyznaczyli wartości krytyczne testów LLC, IPS i kombinacji *p-value* za pomocą procedur bootstrapowych. W ten sposób określono rozkłady z próby, co umożliwia wnioskowanie. Takie podejście powoduje zmniejszenie (choć nie całkowitą eliminację) błędów spowodowanych korelacją przekrojową. Techniczne problemy z praktycznym zastoso-

[3] Test zależności przekrojowej, który ma zastosowanie w różnorodnych modelach panelowych zaproponował M. H. Pesaran [2004].

waniem tej metody przesądzają o jej małym znaczeniu praktycznym. Innym przykładem zastosowania metod bootstrapowych w celu rozwiązania problemu powstającego w panelach z zależnością przekrojową z powodu parametrów zakłócających jest praca Y. C h a n g [2004].

Y. C h a n g [2002] zaproponowała rozwiązanie problemu parametrów zakłócających poprzez konstrukcję statystyki panelowej asymptotycznie niewrażliwej na zależność przekrojową. Dla każdego obiektu oszacowano współczynnik autokorelacji na podstawie regresji ADF używając instrumentów generowanych przez przekształcenie całkowe opóźnionych wartości zmiennej objaśnianej. Na podstawie uzyskanych w ten sposób *N* estymatorów nieliniowej metody zmiennych instrumentalnych skonstruowano następnie indywidualne statystyki *t*, których rozkład graniczny jest standaryzowanym rozkładem normalnym. Statystyka testu Chang jest ich średnią obliczoną po obiektach. Y. C h a n g [2002] podkreśla dobre własności testu w małych próbach i znacznie wyższą moc dyskryminacyjną w stosunku do testu IPS.

Wykorzystanie modelu wspólnych czynników

W podejściu tym zależność przekrojową uwzględnia się przez zastosowanie dynamicznych modeli czynnikowych. Za pomocą takich modeli osiąga się redukcję liczby źródeł zależności przekrojowej i dopuszcza heterogeniczną względem obiektów reakcję na wspólne szoki (poprzez heterogeniczne ładunki czynnikowe).

H. R. M o o n i B. P e r r o n [2004] rozważali model postaci:

$$\begin{cases} y_{it} = \alpha_i + y_{it}^0 \\ y_{it}^0 = \rho_i y_{i,t-1}^0 + \varepsilon_{it} \end{cases} \tag{D.I.12}$$

gdzie: ε_{it} są nieobserwowalnymi składnikami losowym z określoną strukturą czynnikową, α_i – efektami grupowymi. ε_{it} są generowane przez *M* nieobserwowalnych losowych czynników f_i i specyficzne zakłócenia e_{it}: $\varepsilon_{it} = \Lambda_i^{\mathrm{T}} \mathbf{f}_t + e_{it}$, gdzie $\mathbf{\Lambda}_i$ jest wektorem parametrów, a ogólna liczba czynników *M* jest nieznana. Siła zależności przekrojowej jest określona przez wektor $\mathbf{\Lambda}_i$: $\mathrm{E}(y_{it} y_{jt}) = \Lambda_i^{\mathrm{T}} \mathrm{E}(\mathbf{f}_t \mathbf{f}_t^{\mathrm{T}}) \Lambda_j$. Czynniki $\mathbf{f}_t$ są traktowane jako parametry zakłócające, a statystyki testowe[4] konstruowane są na podstawie połączonych danych, z których usunięto wspólne składowe (ang. *pooled defactored data*). Służą one

[4] H. R. M o o n i B. P e r r o n [2004] zaproponowali dwie alternatywne statystyki testowe. Odpowiednie wzory znaleźć można np. w: B a l t a g i [2008], s. 285.

do weryfikacji hipotezy zerowej: H_0: $\rho_i = 1$ dla każdego i, wobec H_1: istnieje takie i, że $\rho_i < 1$. Do obliczenia statystyk M o o n a i P e r r o n a [2004] konieczne jest uprzednie oszacowanie współczynnika korelacji. Również ten estymator bazuje na estymatorze typu *pooled* wyznaczonym na podstawie danych, z których usunięto wspólne składowe. Rozkład graniczny obu statystyk Moona i Perrona jest standaryzowanym rozkładem normalnym, o ile N i $T \to \infty$, przy czym $N/T \to 0$. Asymptotyczna moc testu jest duża, o ile w modelu nie występują trendy deterministyczne. Ponadto, ze względu na zastosowanie procedury usuwania wspólnych składowych, zaleca się stosowanie tego testu, gdy przypuszcza się, że zależność przekrojowa jest spowodowana występowaniem kilku wspólnych czynników.

P. C. B. P h i l l i p s i D. S u l [2003] rozważali model, który zawiera tylko jeden wspólny czynnik f_t o niezależnym w czasie rozkładzie N(0, 1). Podobnie jak H. R. M o o n i B. P e r r o n [2004] proponują oni eliminację wspólnego czynnika, jednak nie za pomocą analizy głównych składowych, ale na podstawie iteracyjnej metody momentów. Po wyeliminowaniu wspólnego czynnika dane wykorzystano do konstrukcji kilku testów pierwiastka jednostkowego. W symulacjach najlepsze wyniki osiągnięto dla statystyki będącej kombinacją wartości p z indywidualnych (obliczonych dla poszczególnych obiektów) testów pierwiastków jednostkowych, postaci:

$$Z = \frac{1}{\sqrt{N}} \sum_{i=1}^{N-1} \Phi^{-1}(p_i) \tag{D.I.13}$$

gdzie: p_i jest wartością p dla statystyki empirycznej testu ADF dla i-tego obiektu, a $\Phi^{-1}(\cdot)$ – funkcją odwrotną do dystrybuanty standaryzowanego rozkładu normalnego. Wyrażenie (D.I.13) zbiega do rozkładu N(0, 1).

Podejście J. B a i i S. N g [2004] jest odmienne od poprzednio prezentowanych, ponieważ nie traktuje ono zależności przekrojowej jako zakłócenia. Model ma postać analogiczną do (D.I.12), ale z ograniczeniem do jednego wspólnego czynnika. Każdy szereg czasowy y_{it} (dla wszystkich i) dekomponowany jest na trzy składowe: deterministyczną d_{it}, K-wymiarowy wektor wspólnych czynników $\mathbf{f}_t$ oraz składową indywidualną (składnik losowy) e_{it}:

$$y_{it} = d_{it} + \mathbf{\Lambda}_i^{\mathrm{T}} \mathbf{f}_t + e_{it} \tag{D.I.14}$$

Szereg ze strukturą czynnikową postaci (D.I.14) jest niestacjonarny, jeśli przynajmniej jedna składowa wektora $\mathbf{f}_t$ lub e_{it} są niestacjonarne. Stacjonarność czynników i składnika losowego jest testowana osobno. Aby było to możliwe, zgodność estymatorów wspólnych czynników musi być niezależna od tego, czy e_{it} są stacjonarne czy nie. Osiągnięto to poprzez konstrukcję estymatorów na

podstawie pierwszych różnic zmiennych i kumulację estymowanych wspólnych czynników. Statystyka testu Bai i Ng jest kombinacją wartości *p* indywidualnych testów ADF przeprowadzonych na podstawie danych, z których wyeliminowano wspólne składowe i ma postać:

$$P = \frac{-2\sum_{i=1}^{N} \ln p_{\hat{e}}^{c}(i) - 2N}{2\sqrt{N}} \tag{D.I.15}$$

gdzie $p_{\hat{e}}^{c}(i)$ jest wartością *p* testu ADF dla estymowanej składowej indywidualnej $\hat{e}_{it}$. Jeśli *N* i $T \to \infty$ to statystyka (D.I.15) zbieżna jest do rozkładu N(0, 1).

M. H. Pesaran [2007] sugeruje metodę eliminacji zależności przekrojowej, która nie wymaga estymacji ładunków czynników i nazywa ją przekrojowym uogólnieniem testu Dickey–Fullera (ang. *cross-sectionally augmented Dickey–Fuller test*, CADF). Test Pesarana bazuje na regresji postaci:

$$\Delta y_{it} = \alpha_i + \rho_i y_{i,t-1} + d_0 \bar{y}_{t-1} + d_1 \bar{y}_t + \varepsilon_{it} \tag{D.I.16}$$

gdzie $\bar{y}_t = \sum_{i=1}^{N} y_{it}$, $\bar{y}_{t-1}$ określone jest analogicznie. Obecność w regresji (D.I.16) $\bar{y}_{t-1}$ i $y_{i,t-1}$ odzwierciedla występowanie zależności przekrojowej. W przypadku występowania autokorelacji składnika losowego lub czynnika regresja (D.I.16) musi być uogólniona do postaci:

$$\Delta y_{it} = \alpha_i + \rho_i y_{i,t-1} + d_0 \bar{y}_{t-1} + \sum_{j=0}^{p} d_{j+1} \Delta \bar{y}_{t-j} + \sum_{k=1}^{p} c_k \Delta y_{i,t-k} + \varepsilon_{it} \tag{D.I.17}$$

Niech $CADF_i$ będzie wartością statystyki ADF dla *i*-tego obiektu określoną jako wartość statystyki *t* dla oceny KMNK parametru ρ_i w regresji (D.I.16) lub (D.I.17). Wartości $CADF_i$ są wykorzystywane do konstrukcji statystyki typu IPS, nazwanej przekrojowym uogólnieniem testu IPS (ang. *cross-sectionally augmented IPS*, CIPS), która jednocześnie uwzględnia zależność przekrojową i autokorelację reszt:

$$CIPS = \frac{1}{N} \sum_{i=1}^{N} CADF_i \tag{D.I.18}$$

Rozkład statystyki *CIPS* jest niestandardowy, a wartości krytyczne uzyskane drogą symulacji dla różnych *T* i *N* oraz trzech alternatywnych wariantów składowych deterministycznych przedstawione są w: P e s a r a n [2007].

Powyżej omówiono możliwości zastosowania alternatywnych testów pierwiastka jednostkowego w podziale na testy zakładające niezależność przekrojową i testy dopuszczające taką zależność. Testy drugiej generacji są chronologicznie późniejsze, przy czym wciąż są one przedmiotem badań. Wszystkie one zakładają taką samą dynamiczną strukturę danych. Podstawowe różnice ująć można w poniższe punkty:

– podejście J. B a i i S. N g [2004] jest ogólniejsze niż H. R. M o o n a i B. P e r r o n a [2004] i M. H. P e s a r a n a [2007], ponieważ dopuszcza, że niestacjonarność pochodzi ze wspólnych czynników lub składnika losowego, przy czym dopuszcza różne stopnie ich integracji oraz kointegrację między y_{it}, jak również między wartościami obserwowalnymi i wspólnymi czynnikami;

– testy P e s a r a n a [2003] oraz B a i i N g [2004] dopuszczają występowanie stałej i trendu liniowego, zaś test M o o n a i P e r r o n a [2004] – tylko stałej.

Ze względu na powyższe różnice, jedynie test Bai i Ng umożliwia wykrycie niestacjonarności szeregu w przypadku, gdy wynika ona tylko z niestacjonarności wspólnych czynników.

C. G e n g e n b a c h, F. P a l m i J. P. U r b a n [2010] badali zachowanie alternatywnych testów w małych próbach. Wnioski doprowadziły do przedstawienia następującej procedury testowania pierwiastka jednostkowego:

1) zastosowanie testu *CADF* Pesarana w przypadku podejrzenia, że zależność przekrojowa jest generowana przez jeden wspólny czynnik lub testów Moona i Perrona, gdy podejrzewa się działanie wielu wspólnych czynników;

2) zastosowanie testów ADF Bai i Ng do testowania występowania pierwiastków jednostkowych w wektorze wspólnych czynników i składowej indywidualnej.

Ponadto C. G e n g e n b a c h, F. P a l m i J. P. U r b a n [2010] stwierdzili, że moc wszystkich testów jest słaba, jeśli proces zawiera trend deterministyczny.

D O D A T E K I I

UOGÓLNIONA METODA MOMENTÓW (GMM) DLA MODELI DANYCH CZASOWYCH

Uogólniona metoda momentów (ang. *Generalized Method of Moments*, GMM) została zaproponowana przez J. P. H a n s e n a [1982], a więc jest metodą stosunkowo nową. Być może z tego powodu jest ona wykorzystywana relatywnie rzadko. Tymczasem jest to metoda wysoce uniwersalna, m. in. dlatego, że warunki momentów często wynikają bezpośrednio z modelu lub z teorii, a własności jej estymatorów nie zależą od założeń o rozkładzie składnika losowego. Umożliwia ona oszacowanie parametrów nawet w sytuacji, gdy modelu (nieliniowego) nie można rozwiązać analitycznie.

W kontekście rozważań przedstawionych w niniejszej publikacji znaczenie GMM jest niezwykle istotne, ponieważ jest ona bazą dla całej grupy metod stosowanych do estymacji dynamicznych modeli panelowych. Poniżej przedstawiony zostanie zarys idei GMM dla modeli danych czasowych i jej związki z uogólnioną metodą zmiennych instrumentalnych.

D.II.1. Warunki momentów i uogólniona metoda zmiennych instrumentalnych

W wielu modelach ekonometrycznych niespełnione jest założenie o braku korelacji między wszystkimi zmiennymi objaśniającymi modelu a składnikiem losowym, co oznacza, że istnieje przynajmniej jedna zmienna x_k, dla której nie jest spełniony warunek:

$$\mathrm{E}(\varepsilon_t\, x_{kt}) = 0 \text{ dla } t = 1, ..., T,\ \ k = 1, ..., K \qquad \text{(D.II.1)}$$

Zastosowanie do estymacji takich modeli klasycznej metody najmniejszych kwadratów (KMNK) skutkowałoby utratą pożądanych własności estymatorów. Przykładami sytuacji, kiedy założenie (D.II.1) nie jest spełnione są:

- Modele z opóźnioną zmienną objaśnianą i z autokorelacją składnika losowego.

Załóżmy dla uproszczenia, że rozważany jest model z jedną zmienną egzogeniczną: $y_t = \alpha_0 + \alpha_1 x_t + \alpha_2 y_{t-1} + \varepsilon_t$ i że składnik losowy podlega procesowi AR(1): $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$.

Zachodzi wtedy również: $y_{t-1} = \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 y_{t-2} + \varepsilon_{t-1}$, skąd wynika, że składnik losowy ε_t jest skorelowany ze zmienną objaśniającą y_{t-1}, gdyż obie wielkości zależą od ε_{t-1}.

- Nieprecyzyjny pomiar zmiennych objaśniających.

Rozważmy model postaci: $y_t = \alpha_0 + \alpha_1 x_t^* + \xi_t$, w którym obserwacje na zmiennej x_t obarczone są błędem pomiaru: $x_t = x_t^* + u_t$. Skutkiem zastosowania KMNK do estymacji modelu $y_t = \alpha_0 + \alpha_1 x_t + \varepsilon_t$ zamiast modelu wyjściowego byłoby uzyskanie estymatorów obciążonych, gdyż x_t i ε_t są skorelowane[1].

- Endogeniczność zmiennych objaśniających modelu.

Wzajemne powiązania wielu zmiennych ekonomicznych, mające charakter sprzężeń zwrotnych, skutkują tym, że zmienne objaśniające modeli ekonometrycznych często są zmiennymi endogenicznymi (są skorelowane ze składnikiem losowym). Najprostszym przykładem może być model konsumpcji postaci: $c_t = \alpha_0 + \alpha_1 y_t + \varepsilon_t$, gdzie c_t oznacza konsumpcję *per capita*, a y_t – dochód *per capita*. Ze względu na zależność: $y_t = c_t + i_t$, gdzie i_t oznacza inwestycje *per capita*, zmienna y_t jest skorelowana ze składnikiem losowym ε_t. Estymator KMNK parametru α_1 przeszacowywałby rzeczywistą krańcową skłonność do konsumpcji (por. V e r b e e k [2004], s. 129–130).

- Dynamiczne modele panelowe.

Estymatory stosowane dla statycznych modeli panelowych są w przypadku modeli dynamicznych niezgodne i obciążone, co zostało uzasadnione szczegółowo w podrozdz. 3.1 niniejszej publikacji. Obecnie przypomnimy jedynie, że utrata zgodności i nieobciążoności estymatora nie wynika z założeń przyjmowanych odnośnie do efektów grupowych α_i (tzn. jest niezależna od tego, czy efekty uznajemy za losowe, czy nielosowe), ale z faktu, że transformowany za pomocą operatora wewnątrzgrupowego wektor obserwacji na opóźnionej zmiennej objaśnianej[2] jest skorelowany z transformowanym w ten sam sposób wektorem składników losowych.

Metodami, które mogą być stosowane do estymacji parametrów modelu, który nie spełnia założenia (D.II.1) są m. in.: metoda zmiennych instrumentalnych (ang. *Instrumental Variables Method*, IV) lub uogólniona metoda momen-

---

[1] Wynika to z faktu, że: $\varepsilon_t = y_t - (\alpha_0 + \alpha_1 x_t) = \alpha_0 + \alpha_1 x_t^* + \xi_t - (\alpha_0 + \alpha_1 x_t^* + u_t) = \xi_t - u_t$, zaś z konstrukcji $x_t = x_t^* + u_t$.

[2] Wynikiem takiej transformacji jest wektor, którego elementami są odchylenia od średnich grupowych; i-ta współrzędna tego wektora ma postać: $y_{iT} - \bar{y}_{i.}$, gdzie $\bar{y}_{i.}$ oznacza średnią arytmetyczną zmiennej y po czasie.

tów (ang. *Generalized Method of Moments*, GMM). Ta druga metoda, opracowana przez J. P. Hansena [1982], jest rozszerzeniem znanej ze statystyki klasycznej metody momentów. Uogólnienie GMM w stosunku do klasycznej metody momentów oznacza dopuszczenie istnienia w modelu większej liczby momentów z próby niż parametrów (możliwa jest także sytuacja odwrotna, ale ma ona w ekonometrii mniejsze znaczenie). Możliwość taka jest szczególnie istotna ze względu na często spotykane zagadnienia z dużą liczbą warunków początkowych. Zastosowanie GMM umożliwia znalezienie rozwiązania, które możliwie najlepiej – dzięki odpowiedniemu doborowi metryki – spełnia podane założenia. Znalezienie zaś rozwiązania za pomocą klasycznej metody momentów jest w przypadku różnej liczby równań i niewiadomych niemożliwe.

O ile stosowalność metody momentów wydaje się mało rozpowszechniona, to metoda zmiennych instrumentalnych jest powszechnie znana i stosowana. Dlatego zasadnicza idea GMM zostanie tu opisana w kontekście estymatora IV dla modelu liniowego.

Rozważmy model:

$$y_t = \mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta} + \varepsilon_t, \quad t = 1, 2, ..., T \tag{D.II.2}$$

gdzie: $\mathbf{x}_t$ jest wektorem zmiennych objaśniających o wymiarach ($k \times 1$); $\boldsymbol{\beta}$ – wektorem nieznanych parametrów; ε_t – składnikiem losowym.

Najprostszą ilustracją idei metody momentów może być przykład poszukiwania estymatorów wartości oczekiwanej i wariancji składnika losowego. Niezależnie od założeń dotyczących rozkładu składnika losowego, model (D.II.2) generuje tzw. warunki momentów:

$$\mathrm{E}(\varepsilon_t) - \mu_0 = 0,$$
$$\mathrm{E}(\varepsilon_t^2) - (\sigma_0^2 + \mu_0^2) = 0,$$

gdzie μ_0, σ_0^2 oznaczają odpowiednio wartość oczekiwaną i wariancję składnika losowego.

Ich odpowiednikami z próby są warunki:

$$T^{-1}\sum_{t=1}^{T}\varepsilon_t - \hat{\mu}_t = 0,$$
$$T^{-1}\sum_{t=1}^{T}\varepsilon_t^2 - (\hat{\sigma}_t^2 + \hat{\mu}_t^{\,2}) = 0.$$

Stąd estymatorami parametrów rozkładu są:

$$\hat{\mu}_t = T^{-1}\sum_{t=1}^{T}\varepsilon_t,$$

$$\hat{\sigma}_t^2 = T^{-1}\sum_{t=1}^{T}(\varepsilon_t^2 - \hat{\mu}_t)^2.$$

Wyznaczone one zostały jedynie na podstawie warunków momentów.

Uogólniając, układ K warunków postaci:

$$\mathrm{E}(\varepsilon_t \boldsymbol{x}_t) = \mathbf{0} \quad \text{dla } t = 1, 2, ..., T \tag{D.II.3}$$

jest przykładem warunków momentów z populacji (ang. *population moment conditions*). Są to warunki związane z wartością oczekiwaną (pierwszym momentem), wynikające z konstrukcji modelu.

Jeśli dowolna zmienna x_{kt} nie spełnia warunku (D.II.3), to estymator KMNK parametrów modelu (D.II.2) nie jest zgodny. Ideą metody zmiennych instrumentalnych jest, jak wiadomo, zastąpienie jej inną zmienną z_t, zwaną zmienną instrumentalną, która jest nieskorelowana ze składnikiem losowym, ale skorelowana, i to jak najsilniej, ze zmienną x_{kt} (a więc spełnia (D.II.3)).

W ogólnym przypadku, do dyspozycji jest R zmiennych instrumentalnych[3], tworzących wektor $\mathbf{z}_t = [z_{1t}, ..., z_{Rt}]^{\mathrm{T}}$, którego pewne składowe mogą być równe składowym wektora $\mathbf{x}_t$. Warunek (D.II.3) przyjmuje wówczas postać:

$$\mathrm{E}(\varepsilon_t \mathbf{z}_t) = \mathbf{0} \quad \text{dla } t = 1, 2, ..., T \tag{D.II.4}$$

Jeśli $R = K$, a więc liczba warunków (D.II.4) jest równa liczbie nieznanych parametrów, to estymator metody zmiennych instrumentalnych $\hat{\boldsymbol{\beta}}_{\mathrm{IV}}$ wyznacza się tak, by spełniał układ R równań, będących warunkami momentów z próby (ang. *sample moments conditions*):

$$\frac{1}{T}\sum_{t=1}^{T}(y_t - \mathbf{x}_t^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{\mathrm{IV}})z_t = 0 \tag{D.II.5}$$

Jeśli $R > K$, to układ (D.II.5) jest nieoznaczony. Pominięcie dowolnych $R - K$ równań zapewniłoby uzyskanie rozwiązania jednoznacznego, ale skutkowałoby też stratami efektywności. Dlatego estymatorów wektora $\boldsymbol{\beta}$ poszukuje się

[3] Zakładamy, że $R \geq K$, gdyż w przeciwnym wypadku układ byłby nieidentyfikowalny.

w taki sposób, aby każde z równań (D.II.5) przyjmowało wartości jak najbliższe 0. Osiąga się to minimalizując formę kwadratową:

$$Q_T(\boldsymbol{\beta}) = \left[\frac{1}{T}\sum_{t=1}^{T}(y_t - \mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta})\mathbf{z}_t\right]^{\mathrm{T}} \mathbf{W}_T \left[\frac{1}{T}\sum_{t=1}^{T}(y_t - \mathbf{x}_t^{\mathrm{T}}\boldsymbol{\beta})\mathbf{z}_t\right]$$

lub w notacji macierzowej:

$$Q_T(\boldsymbol{\beta}) = \left[\frac{1}{T}\sum_{t=1}^{T}\mathbf{Z}^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\right]^{\mathrm{T}} \mathbf{W}_T \left[\frac{1}{T}\sum_{t=1}^{T}\mathbf{Z}^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\right] \qquad \text{(D.II.6)}$$

gdzie $\mathbf{W}_T$ jest macierzą wag stopnia R, dodatnio określoną, symetryczną. Macierz $\mathbf{W}_T$ określa wagi, nadawane poszczególnym kombinacjom liniowym warunków momentów z próby. W ogólnym przypadku postać $\mathbf{W}_T$ zależy od liczebności próby T, ponieważ macierz ta może być estymowana. Różniczkowanie formy kwadratowej (D.II.6) względem $\boldsymbol{\beta}$ daje warunki pierwszego rzędu postaci:

$$-2\mathbf{X}^{\mathrm{T}}\mathbf{Z}\mathbf{W}_T\mathbf{Z}^{\mathrm{T}}\mathbf{y} + 2\mathbf{X}^{\mathrm{T}}\mathbf{Z}\mathbf{W}_T\mathbf{Z}^{\mathrm{T}}\mathbf{X}\hat{\boldsymbol{\beta}}_{\mathrm{IV}} = \mathbf{0} \qquad \text{(D.II.7)}$$

Jest to układ K warunków z K nieznanymi elementami wektora $\hat{\boldsymbol{\beta}}_{\mathrm{IV}}$. Jeśli macierz $\mathbf{X}^{\mathrm{T}}\mathbf{Z}$ jest rzędu K, to rozwiązanie układu (D.II.7) istnieje i ma postać:

$$\hat{\boldsymbol{\beta}}_{\mathrm{IV}} = (\mathbf{X}^{\mathrm{T}}\mathbf{Z}\mathbf{W}_T\mathbf{Z}^{\mathrm{T}}\mathbf{X})^{-1}(\mathbf{X}^{\mathrm{T}}\mathbf{Z}\mathbf{W}_T\mathbf{Z}^{\mathrm{T}}\mathbf{y}) \qquad \text{(D.II.8)}$$

Ważną własnością estymatorów postaci (D.II.8) jest to, że o ile macierz $\mathbf{W} = p\lim \mathbf{W}_T$ jest dodatnio określona[4], to są one zgodne, niezależnie od doboru macierzy wag $\mathbf{W}_T$. Intuicyjne uzasadnienie tej własności jest następujące: estymator (D.II.8) uzyskuje się minimalizując formę kwadratową (D.II.6), zależną od momentów z próby. Momenty te są asymptotycznie zbieżne do momentów z populacji[5], które dla rzeczywistych wartości parametrów przyjmują wartość 0 (warunki (D.II.4)). Idea ta leży u podstaw metody momentów.

---

[4] Dowolny ciąg zmiennych losowych $\{h_t\}$ jest zbieżny według prawdopodobieństwa do zmiennej losowej h, $p\lim h_t = h$, jeśli dla wszystkich $\varepsilon > 0$ zachodzi: $\lim_{t\to\infty} P(|h_t - h| < \varepsilon) = 1$. Macierz symetryczna $\mathbf{A}$ jest dodatnio określona, jeśli dla dowolnego wektora $\mathbf{x}$ zachodzi: $\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} > 0$.

[5] Wynika to z praw wielkich liczb – twierdzeń dotyczących sytuacji, w których momenty z próby dążą do momentów z populacji.

Wzór (D.II.8) określa zbiór estymatorów, zależnych od macierzy wag $\mathbf{W}_T$. Można wykazać, że najefektywniejszy estymator uzyskuje się dla macierzy $\mathbf{W}_T^{\text{ef}} = \left(\frac{1}{T}\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\right)^{-1}$ (por. V e r b e e k [2004], s. 144). Estymator (D.II.8) przyjmuje wówczas postać:

$$\hat{\boldsymbol{\beta}}_{\text{GIV}} = (\mathbf{X}^{\mathrm{T}}\mathbf{Z}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{X})^{-1}(\mathbf{X}^{\mathrm{T}}\mathbf{Z}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{y}) \tag{D.II.9}$$

Estymator (D.II.9) nazywany jest estymatorem uogólnionej metody zmiennych instrumentalnych (*Generalized Instrumental Variables Method*, GIV). Jest on równoważny estymatorowi dwustopniowej metody najmniejszych kwadratów (2MNK). Macierz wariancji-kowariancji tego estymatora ma postać:

$$\mathrm{D}^2(\hat{\boldsymbol{\beta}}_{\text{GIV}}) = \hat{\sigma}_2(\mathbf{X}^{\mathrm{T}}\mathbf{Z}(\mathbf{Z}^{\mathrm{T}}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{X})^{-1} \tag{D.II.10}$$

gdzie $\hat{\sigma}_2$ uzyskuje się na podstawie reszt estymatora IV.

D.II.2. Uogólniona metoda momentów

Wektor momentów z populacji (momentów teoretycznych) jest w najbardziej ogólnym przypadku funkcją nieznanych parametrów:

$$\mathrm{E}[\mathbf{f}(\mathbf{z}_t, \boldsymbol{\theta})] = \mathbf{0} \tag{D.II.11}$$

gdzie $\mathbf{f}$ jest wektorem funkcji o R składowych, $\mathbf{z}_t$ – wektorem zmiennych objaśniających lub instrumentalnych, $\boldsymbol{\theta}$ – wektorem nieznanych parametrów.

Parametry są zatem estymowane bezpośrednio z warunków momentów, które mogą być liniowe lub nieliniowe względem parametrów. Załóżmy, że dysponujemy R warunkami momentów postaci (D.II.11). Wektor momentów z próby odpowiadający tym warunkom ma postać:

$$\mathrm{g}(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{f}(\mathbf{z}_t, \boldsymbol{\theta}) \tag{D.II.12}$$

Podobnie jak w metodzie zmiennych instrumentalnych, aby zapewnić identyfikowalność układu, należy założyć, że $R \geq K$. Jeśli $R = K$, to po przyrównaniu (D.II.12) do zera możliwe jest wyznaczenie zgodnych estymatorów $\boldsymbol{\theta}$ z uzyskanego układu R równań. W takim przypadku GMM sprowadza się do klasycznej metody momentów.

Jeśli $R > K$, to rozwiązanie jednoznaczne układu (D.II.12) nie istnieje. Usunięcie $R - K$ równań byłoby nieefektywne, gdyż każde z nich wnosi pewne informacje dotyczące parametrów, a ponadto wybór tych równań dokonywałby się arbitralnie. Zamiast tego korzysta się ze wszystkich R warunków i poszukuje średniej ważonej estymatorów dla niejednoznacznie identyfikowalnego wektora $\boldsymbol{\theta}$ tak, aby wariancja była jak najmniejsza. W uogólnionej metodzie momentów estymator wektora $\boldsymbol{\theta}$ wyznacza się tak, aby wektor momentów z próby przyjmował wartości najbliższe zeru, co zapewnia się minimalizując ze względu na $\boldsymbol{\theta}$ formę kwadratową:

$$Q_T(\boldsymbol{\theta}) = \left[\frac{1}{T}\sum_{t=1}^{T}\mathbf{f}(\boldsymbol{\theta},\mathbf{z}_t)\right]^{\mathrm{T}}\mathbf{W}_T\left[\frac{1}{T}\sum_{t=1}^{T}\mathbf{f}(\boldsymbol{\theta},\mathbf{z}_t)\right] \tag{D.II.13}$$

gdzie $\mathbf{W}_T$ jest pewną dodatnio określoną, symetryczną macierzą wag stopnia R, spełniającą warunek $p\lim \mathbf{W}_T = \mathbf{W}$, gdzie macierz $\mathbf{W}$ jest pewną macierzą dodatnio określoną i symetryczną. Jest więc to postępowanie analogiczne, jak w przypadku wyznaczania estymatora GIV. Estymator GMM wektora $\boldsymbol{\theta}$ ma formę:

$$\hat{\boldsymbol{\theta}}_{\mathrm{GMM}} = \min_{\theta} Q_T(\boldsymbol{\theta}) = \min_{\theta} \boldsymbol{g}(\boldsymbol{\theta})^{\mathrm{T}}\mathbf{W}_T\mathbf{g}(\boldsymbol{\theta}) \tag{D.II.14}$$

Ponieważ estymator (D.II.14) zależy od macierzy $\mathbf{W}_T$, to poprawne jej wyznaczenie ma kluczowe znaczenie. Macierz ta determinuje, które K, spośród wszystkich R warunków momentów, zostanie wykorzystanych w procesie estymacji. Warunki te, zwane warunkami identyfikującymi (ang. *identyfying restrictions*, *identyfying conditions*) określone są poprzez zależność (por. S o w e l l [1996]; H a l l [2005]):

$$\mathbf{F}(\mathbf{F}^{\mathrm{T}}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{W}^{1/2}\mathrm{E}(\mathbf{z}_t\boldsymbol{\varepsilon}_t) = \mathbf{0} \tag{D.II.15}$$

gdzie: $\mathbf{F}^{\mathrm{T}} = \mathrm{E}(\mathbf{x}_t\mathbf{z}_t^{\mathrm{T}})\mathbf{W}^{1/2^{\mathrm{T}}}$, $\mathbf{W}^{1/2}$ jest macierzą stopnia R, spełniającą warunek $\mathbf{W}_T = \mathbf{W}^{1/2^{\mathrm{T}}}\mathbf{W}^{1/2}$; jej istnienie jest zagwarantowane przez dodatnią określoność macierzy $\mathbf{W}_T$. Układ warunków (D.II.15) składa się z R równań, spośród których tylko K jest (z założenia) liniowo niezależnych. Pozostałe warunki momentów są niewykorzystane w procesie estymacji. Określone one są przez zależność:

$$\left(\mathbf{I}_R - \mathbf{F}(\mathbf{F}^{\mathrm{T}}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\right)\mathbf{W}^{1/2}\mathrm{E}(\mathbf{z}_t\boldsymbol{\varepsilon}_t) = \mathbf{0} \tag{D.II.16}$$

F. Sowell [1996] nazywa je warunkami ponadidentyfikującymi (ang. *overidentyfying restrictions*). Warunki identyfikujące i ponadidentyfikujące są wzajemnie ortogonalne, gdyż

$$\mathbf{F}(\mathbf{F}^{\mathrm{T}}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\left(\mathbf{I}_R - \mathbf{F}(\mathbf{F}^{\mathrm{T}}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\right) = \mathbf{0}.$$

Hansen [1982] wykazał, że najmniejszą asymptotyczną wariancję ma estymator $\hat{\boldsymbol{\theta}}_{\mathrm{GMM}}$ uzyskany przez podstawienie w (D.II.14) macierzy wag:

$$\mathbf{W}^{\mathrm{opt}} = \left(\mathrm{E}\left[\mathbf{f}(\boldsymbol{\theta}, \mathbf{z}_t)\right]\left[\mathbf{f}(\boldsymbol{\theta}, \mathbf{z}_t)\right]^T\right)^{-1} \tag{D.II.17}$$

Macierz ta jest jednak nieznana, gdyż zależy od nieznanego wektora $\boldsymbol{\theta}$. Najczęściej spotykanym sposobem praktycznego wyznaczenia macierzy wag $\mathbf{W}_T$ jest metoda iteracyjna.

W pierwszym kroku wyznacza się estymator $\hat{\boldsymbol{\theta}}_1$ zastępując w (D.II.14) macierz $\mathbf{W}_T$ dowolną macierzą symetryczną, dodatnio określoną. Zazwyczaj przyjmuje się $\mathbf{W}_T = \mathbf{I}$.

W drugim kroku wartość $\hat{\boldsymbol{\theta}}_1$ wstawia się do (D.II.17), uzyskując kolejne oszacowanie macierzy wag. Procedurę tę można powtarzać ustaloną ilość razy (zazwyczaj 2 lub 3) albo do uzyskania odpowiednio małej wartości funkcji (D.II.14). Ostatecznie wyznaczoną wartość $\hat{\boldsymbol{\theta}}_*$ wstawia się do (D.II.17), otrzymując zgodny estymator optymalnej macierzy wag:

$$\mathbf{W}_T^{\mathrm{opt}} = \left(\frac{1}{T}\sum_{t=1}^{T}\left[\mathbf{f}(\boldsymbol{\theta}^*, \mathbf{z}_t)\right]\left[\mathbf{f}(\boldsymbol{\theta}^*, \mathbf{z}_t)\right]^{\mathrm{T}}\right)^{-1} \tag{D.II.18}$$

Następnie wyznacza się asymptotycznie efektywny estymator $\hat{\boldsymbol{\theta}}_{\mathrm{GMM}}$, wstawiając do wzoru (D.II.14) zależność $\mathbf{W}_T = \mathbf{W}_T^{\mathrm{opt}}$.

Można wykazać (por. Hall [2005]), że estymator $\hat{\boldsymbol{\theta}}_{\mathrm{GMM}}$ jest zgodny i ma rozkład asymptotycznie normalny:

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{\mathrm{GMM}} - \boldsymbol{\theta}\right) \to \mathrm{N}(0, \mathbf{MSM}^{\mathrm{T}}) \tag{D.II.19}$$

gdzie $\mathbf{M} = \left\{\mathrm{E}(\mathbf{x}_t\mathbf{z}_t^{\mathrm{T}})\mathbf{W}\mathrm{E}(\mathbf{x}_t\mathbf{z}_t^{\mathrm{T}})\right\}^{-1}\mathrm{E}(\mathbf{x}_t\mathbf{z}_t^{\mathrm{T}})\mathbf{W}$, zaś

$$\mathbf{S} = \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{E}(u_t u_s \mathbf{z}_t\mathbf{z}_s) = \mathrm{E}(u^2\mathbf{z}\mathbf{z}^{\mathrm{T}}).$$

H a n s e n [1982] udowodnił, że $\mathbf{S}^{-1} = \mathbf{W}_T^{\text{opt}}$. Oznacza to, że momenty z próby, których wariancja jest niewielka, a zatem informacje uzyskane na ich podstawie o parametrach θ_i są precyzyjne, mają większą wagę w procesie estymacji niż momenty charakteryzujące się dużą wariancją[6].

Zgodny estymator $\mathbf{MSM}^{\mathrm{T}}$ można, na podstawie twierdzenia Slutsky'ego, otrzymać jako iloczyn estymatorów poszczególnych czynników. Zgodnym estymatorem macierzy $\mathbf{M}$ jest: $\hat{\mathbf{M}} = \left\{(T^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Z})\mathbf{W}_T(T^{-1}\mathbf{Z}^{\mathrm{T}}\mathbf{X})\right\}^{-1}(T^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Z})\mathbf{W}_T$, a zgodnym estymatorem $\mathbf{S}$: $\hat{\mathbf{S}}_T = T^{-1}\sum_{t=1}^{T}\hat{u}_t^2\mathbf{z}_t\mathbf{z}_t^{\mathrm{T}}$, gdzie $\hat{u}_t = y_t - \mathbf{x}_t^T\hat{\boldsymbol{\theta}}_{\mathrm{GMM}}$.

Zalety GMM, w porównaniu z KMNK, są następujące: 1) metoda ta nie wymaga założeń o rozkładzie składnika losowego, 2) dopuszcza heteroskedastyczność dowolnej postaci, 3) umożliwia oszacowanie parametrów, nawet jeśli modelu (nieliniowego) nie można rozwiązać analitycznie, 4) warunki momentów często wynikają bezpośrednio z modelu lub z teorii.

GMM posiada też co najmniej dwie wady. Jedna z nich dotyczy rozkładu estymatora $\hat{\boldsymbol{\theta}}_{\mathrm{GMM}}$. Rozkład asymptotycznie normalny (D.II.19) stanowi na ogół bardzo słabe przybliżenie dla niezbyt licznych prób, którymi najczęściej posługuje się w praktycznych zastosowaniach (por. H a n s e n, H e a t o n, Y a r o n [1996]). Innym poważnym problemem jest problem tzw. „słabych instrumentów" (ang. *weak instruments*). Jeśli zmienne instrumentalne są zbyt słabo skorelowane ze zmienną objaśniającą, to estymator $\hat{\boldsymbol{\theta}}_{\mathrm{GMM}}$ nie jest zgodny, a jego rozkład nie jest rozkładem asymptotycznie normalnym, nawet dla bardzo dużych prób (por. S t o c k, W r i g h t [2003]).

D.II.3. Test Sargana

Ponieważ w uogólnionej metodzie momentów parametry estymowane są bezpośrednio z warunków momentów, szczególnej wagi nabiera problem sprawdzenia, czy warunki te sformułowane są poprawnie. Innymi słowy sprawdzeniu podlegać może właściwość użytych w estymacji instrumentów.

Podstawowym testem jest tzw. test na warunki ponadidentyfikujące (ang. *overidentifying restrictions test*) zwany też testem Sargana (S a r g a n [1958]; H a n s e n [1982]). Stosuje się go, jeśli $\boldsymbol{\theta}$ jest identyfikowalne z nadmiarem ($R > K$). W przypadku, gdy $R = K$, wektor momentów z próby jest zerowy z konstrukcji, niezależnie od tego, czy momenty z populacji rzeczywiście

[6] Interpretacja powyższa jest analogiczna do idei ważonej metody najmniejszych kwadratów. Podstawowa różnica jest taka, że w omawianym tu przypadku wagi przypisywane są momentom z próby, a nie obserwacjom.

przyjmują wartość zero. Oznacza to, że w takim przypadku poprawność warunków momentów jest niesprawdzalna. Testowane są warunki „nadliczbowe”, ponad te, konieczne do identyfikacji.

Hipoteza zerowa testu Sargana zakłada, że wektor momentów z populacji jest wektorem zerowym (warunki (D.II.11)). Oznacza to, że instrumenty wykorzystane w procesie estymacji są właściwe, tzn. nieskorelowane ze składnikiem losowym modelu. Jeśli $\boldsymbol{\theta}$ jest identyfikowalne z nadmiarem ($R > K$) i hipoteza zerowa jest prawdziwa, to elementy wektora momentów z próby $\mathbf{g}(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{f}(\mathbf{z}_t, \boldsymbol{\theta})$ mają wartości bliskie zeru. Na tej podstawie opiera się konstrukcja testu. Statystyka empiryczna ma postać:

$$ST = T \cdot Q_T(\boldsymbol{\theta}) = T \cdot \left[\frac{1}{T}\sum_{t=1}^{T}\mathbf{f}(\boldsymbol{\theta}, \mathbf{z}_t)\right]^{\mathrm{T}} \mathbf{W}_T \left[\frac{1}{T}\sum_{t=1}^{T}\mathbf{f}(\boldsymbol{\theta}, \mathbf{z}_t)\right] \qquad \text{(D.II.20)}$$

Ma ona rozkład χ^2 z $R - K$ stopniami swobody. Liczba stopni swobody jest więc równa liczbie zmiennych instrumentalnych pomniejszonej o liczbę parametrów modelu. Inaczej, liczba stopni swobody jest równa liczbie warunków ponad te, które są konieczne do identyfikacji K nieznanych parametrów.

WYKAZ OZNACZEŃ

$t = 1, ..., T$	– numery okresów
$i = 1, ..., N$	– numery obiektów
α_i	– efekt grupowy właściwy dla i-tego obiektu
λ_t	– efekt czasowy właściwy dla okresu t
ε_{it}	– właściwy składnik losowy
u_{it}	– składnik losowy ($u_{it} = \alpha_i + \varepsilon_{it}$ lub $u_{it} = \alpha_i + \lambda_t + \varepsilon_{it}$), w którym wyróżnia się efekty grupowe (i czasowe) oraz „właściwy” składnik losowy
$\otimes$	– iloczyn Kroneckera macierzy
$\mathbf{I}_N$	– macierz jednostkowa stopnia N
$\mathbf{J}_T$	– macierz jedynkowa stopnia T
$\mathrm{N}(0, \sigma^2)$	– rozkład normalny o wartości oczekiwanej 0 i wariancji σ^2
$\mathrm{E}(y)$	– wartość oczekiwana zmiennej y
$\mathrm{cov}(xy)$	– kowariancja zmiennych x i y
$\mathrm{D}^2(y) = \mathrm{E}(y^2) = \sigma^2$	– wariancja zmiennej y
$\mathrm{E}(y\|\alpha)$	– warunkowa wartość oczekiwana
Δy_t	– przyrost wartości zmiennej y w okresie t
$p\lim h_t$	– granica według prawdopodobieństwa ciągu zmiennych losowych $\{h_t\}$
$\mathrm{diag}(a_{11}, a_{22}, ..., a_{nn})$	– macierz diagonalna, której elementy głównej przekątnej są równe $a_{11}, a_{22}, ..., a_{nn}$
$O(f_2(x)) = f_1(x)$	– f_1 jest co najwyżej rzędu f_2
$\mathbf{Q}$	– operator wewnątrzgrupowy
$\mathbf{P}$	– operator międzygrupowy
$\mathbf{\Omega}$	– macierz wariancji-kowariancji składnika losowego
FE	– ze zmiennymi sztucznymi (efektami nielosowymi, *fixed effects*)
RE	– dekompozycją składnika losowego (efektami losowymi, *random effects*)
$\hat{\boldsymbol{\beta}}_{\mathrm{W}}$	– estymator wewnątrzgrupowy wektora parametrów strukturalnych (dla modeli FE)

$\hat{\boldsymbol{\beta}}_{B}$	– estymator międzygrupowy wektora parametrów strukturalnych
$\hat{\boldsymbol{\beta}}_{UMNK}$	– estymator panelowy UMNK wektora parametrów strukturalnych (dla modeli RE)
$\hat{\boldsymbol{\beta}}_{AH}$	– estymator Andersona–Hsiao wektora parametrów strukturalnych
$\hat{\boldsymbol{\beta}}_{IV}$	– estymator metody zmiennych instrumentalnych
$\hat{\boldsymbol{\beta}}_{FDGMM}$	– estymator FDGMM wektora parametrów strukturalnych
$\hat{\boldsymbol{\beta}}_{SGMM}$	– estymator SGMM wektora parametrów strukturalnych
$\hat{\boldsymbol{\delta}}_{1,PM2MNK}$	– estymator dwustopniowej metody najmniejszych kwadratów dla modeli panelowych (2MNK-MP)
$\hat{\boldsymbol{\delta}}_{PM3MNK}$	– estymator trójstopniowej metody najmniejszych kwadratów dla modeli panelowych (PM3MNK)
WG	– (estymator) wewnątrzgrupowy
SGMM	– systemowa GMM Blundella i Bonda
FDGMM	– GMM pierwszych różnic Arellano i Bonda
IV	– metoda zmiennych instrumentalnych
GMM	– uogólniona metoda momentów
UNMK	– uogólniona metoda najmniejszych kwadratów
MNW	– metoda największej wiarygodności
2MNK	– dwustopniowa metoda najmniejszych kwadratów
3MNK	– trójstopniowa metoda najmniejszych kwadratów
KMNK	– klasyczna metoda najmniejszych kwadratów
AR(1)	– proces autoregresyjny pierwszego rzędu

BIBLIOGRAFIA

Abramovitz M. [1986], *Catching up, forging ahead, and falling behind*, „The Journal of Economic History", Vol. 46, No. 2, s. 385–406

Acemoglu D., Johnson S., Robinson J. A., Yared P. [2005], *From education to democracy*, „American Economic Review", Vol. 95, s. 44–49

Acharya R. C., Keller W. [2007], *Technology Transfer through Imports*, „NBER Working Paper", No. 13086

Aghion P., Blundell R., Griffith R., Howitt P., Prantl S. [2004], *Entry and productivity growth: evidence from microlevel panel data*, „Journal of the European Economic Association", Vol. 2, No. 2–3, s. 265–276

Ahn S. C., Schmidt P. [1995], *Efficient estimation of models for dynamic panel data*, „Journal of Econometrics", Vol. 68, s. 5–28

Allison P. [2000], *Inferring casual order from panel data*, paper presented at the Ninth International Conference on Panel Data, June 22–23, 2000, Geneva, Switzerland

Alonzo-Borrego C., Arellano M. [1999], *Symmetrically normalized instrumental variable estimation using panel data*, „Journal of Business and Economic Statistics", Vol. 17, s. 36–49

Alvarez J., Arellano M. [2003], *The time series and cross-section asymptotics of dynamic panel data estimators*, „Econometrica", Vol. 71, s. 1121–1159

Amemiya T. [1971], *The estimation of the variances in a variance-components model*, „International Economic Review", Vol. 12, s. 1–13

Anderson T. W., Hsiao C. [1981], *Estimation of Dynamic models with Error Components*, „Journal of the American Statistical Association", Vol. 76, s. 598–606

Anderson T. W., Hsiao C. [1982], *Formulation and Estimation of Dynamic Models Using Panel Data*, „Journal of Econometrics", Vol. 18, s. 47–82

Andini C. [2009], *How fast do wages adjust to human-capital productivity? Dynamic panel – data evidence from Belgium, Denmark and Finland*, „IZA Discussion Paper", No. 4583

Andrews D. W. [1987], *Asymptotic results for generalized Wald tests*, „Econometric Theory", Vol. 3, No. 3, s. 348–358

Anselin L., Le Gallo J., Jayet H. [2008], *Spatial panel econometrics*, [w:] L. Matyas, P. Sevestre (eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, 3rd Ed., Springer, New York, Chapter 19

Arellano M. [1989], *A note on the Anderson–Hsiao estimator for panel data*, „Economics Letters", Vol. 31, s. 337–341

Arellano M. [1990], *Some testing for autocorrelation in dynamic random effects model*, „Review of Economic Studies", Vol. 57, s. 127–134

Arellano M. [2003], *Panel Data Econometrics*, Oxford University Press, Oxford

Arellano M., Bond S. [1991], *Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations*, „Review of Economic Studies", Vol. 58, s. 277–297

Arellano M., Bover O. [1995], *Another look at the instrumental variable estimation of error – components models*, „Journal of Econometrics", Vol. 68, s. 29–52

Arellano M., Honoré B. [2001], *Panel data models: some recent developments*, [w:] J. Heckman, E. Leamer (eds.), *Handbook of Econometrics*, Vol. 5, Chapter 53, North-Holland, Amsterdam

Arizala F., Cavallo E., Galindo A. [2009], *Financial development and TFP growth: cross-country and industry-level evidence*, „Inter-American Development Bank Research Department Working Paper", #682 RG-K1097

Ascari G., Di Cosmo V. [2004], *Determinants of Total Factor Productivity in the Italian Regions*, „University of Pavia, Department of Economics, Working Paper", #170 (12-04), http://129.3.20.41/eps/mac/papers/0511/0511009.pdf, 8.03.2009

Avery R. [1977], *Error components and seemingly unrelated regressions*, „Econometrica", Vol. 45, s. 199–209

Bai J., Ng S. [2004], *A PANIC attack on unit roots and cointegration*, „Econometrica", Vol. 72, s. 1127–1177

Bal-Domańska B. [2008], *Rozwój regionalny a kapitał ludzki – analiza ekonometryczna z wykorzystaniem modeli dla danych panelowych*, [w:] *Metody ilościowe w ekonomii*, S. Forlicz (red.), Wydawnictwo Wyższej Szkoły Bankowej w Poznaniu, Poznań, s. 192–208

Bal-Domańska B. [2009], *Wpływ gospodarki opartej na wiedzy na rozwój regionalny*, „Wiadomosci Statystyczne", vol. 54, nr 6, s. 84–98

Balestra P., Nerlove M. [1966], *Pooling Cross Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for natural Gas*, „Econometrica", Vol. 34, s. 585–612

Baltagi B. H. [1980], *On seemingly unrelated regressions with error components*, „Econometrica", Vol. 48, s. 1547–1551

Baltagi B. H. [1981], *Simultaneous equations with error components*, „Journal of Econometrics", Vol. 17, s. 189–200

Baltagi B. H. [2008], *Econometric Analysis of Panel Data*, 4th Ed., Wiley&Sons, Chichester

Baltagi B. H., Griffin J. M. [1983], *Gasoline demand in the OECD: An application of pooling and testing procedures*, „European Economic Review", Vol. 22, s. 117–137

Baltagi B. H., Griffin J. M. [1984], *Short and long run effects in pooled models*, „International Economic Review", Vol. 25, s. 631–645

Baltagi B. H., Griffin J. M. [1997], *Pooled estimators vs. their heterogenous counterparts in the context of dynamic demand for gasoline*, „Journal of Econometrics", Vol. 77, s. 303–327

Baltagi B. H., Kao C. [2000], *Nonstationary panels, cointegration in panels and dynamic panels: a survey*, „Advances in Econometrics", Vol. 15, s. 7–51

Baltagi B. H., Bresson G., Griffin J. M., Pirotte A. [2003], *Homogenous, heterogenous or shrinkage estimators? Some empirical evidence from French regional gasoline consumption*, „Empirical Economics", Vol. 28, s. 795–811

Baltagi B. H., Griffin J. M., Rich D. P. [1995], *The measurement of firm-specific indexes of technical change*, „The Review of Economics and Statistics", Vol. 77, No. 4, s. 654–663

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

Barro R. J. [1990], *Government spending in a simple model of endogenous growth*, „Journal of Political Economy", Vol. 98, s. 103–125

Barro R. J. [1991], *Economic growth in cross-section of countries*, „Quarterly Journal of Economics", Vol. 106, s. 407–443

Barro R. J., Sala-i-Martin X. [1992], *Convergence*, „Journal of Political Economy", Vol. 100, s. 223–251

Barro R. J., Sala-i-Martin X. [2003], *Economic Growth*, MIT Press, Cambridge (MA)

Bartkowiak R. [2003], *Historia myśli ekonomicznej*, PWE, Warszawa

Bassanini A., Scarpetta S., Hemmings P. [2001], *Economic growth: The role of policies and institutions. Panel data evidence from OECD countries*, „OECD Economic Department Working Papers", No. 283

Baumol W. J. [1986], *Productivity growth, convergence and welfare: what the long-run data show*, „American Economic Review", Vol. 76, s. 1072–1085

Beck N., Katz L. N. [1995], *What to Do (and Not to Do) with Time-Series-Cross-Section Data*, „American Political Science Review", Vol. 89, s. 634–647

Beck N., Katz L. N. [1996], *Nuisance vs. Substance: Specifying and Estimating Time-Series--Cross-Section Models*, „Political Analysis", Vol. 6, s. 1–34

Bekker P. A. [1994], *Alternative approximations to the distributions of instrumental variable estimators*, „Econometrica", Vol. 62, s. 657–681

Benito A., Hernando I. [2007], *Firm behaviour and financial pressure: evidence from Spanish panel data*, „Bulletin of Economic Research", Vol. 59, No. 4, s. 283–311

Bernanke B. S., Gurkaynak R. S. [2001], *Is growth exogenous? Taking Mankiw, Romer, and weil seriously*, „Working Papers", No. 8365, www.nber.org/papers/w8365

Bhargava A., Sargan J. D. [1983], *Estimating dynamic random effects models from panel data covering short time periods*, „Econometrica", Vol. 51, s. 1635–1659

Binder M., Hsiao C., Pesaran M. H. [2005], *Estimation and inference in short panel vector autoregressions with unit roots and cointegration*, „Econometric Theory", Vol. 21, s. 795–837

Bloom N., Schankerman M., Van Reenen J. [2005], *Identifying technology spillovers and product market rivalry*, „LSE STICERD Research Paper", No. EI40

Blundell R., Bond S. [1998], *Initial conditions and moment restrictions in dynamic panel data models*, „Journal of Econometrics", Vol. 87, No. 1, s. 115–143

Blundell R., Bond S. [2000], *GMM estimation with persistent panel data: An application to production functions*, „Econometric Reviews", Vol. 19, s. 321–340

Blundell R., Smith R. [1991], *Initial conditions and efficient estimation in dynamic panel data models: An application to company investment behaviour*, „Annales d'Economie et de Statistique", Vol. 21, No. 20, s. 109–123

Blundell R., Bond S., Windmeijer F. [2000], *Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator*, „Advances in Econometrics", Vol. 15, s. 53–91

Bo H. [2007], *Nonlinear effects of debt on investment: evidence from Dutch listed firms*, „European Journal of Finance", Vol. 13, s. 669–687

Bobba M., Coviello D. [2007], *Weak instruments and weak identification in estimating the effects of education on democracy*, „Economic Letters", Vol. 96, s. 301–306

Bond S. [2002], *Dynamic panel data models: A guide to micro data methods and practice*, Centre for Microdata and Practice (CEMMAP), „Working Paper" CWP09/02

Bond S., Windmeijer F. [2005], *Reliable inference for GMM estimators? Finite sample properties of alternative test procedures in linear panel data models*, „Econometric Reviews", Vol. 24, No. 1, s. 1–37

Bond S., Hoeffler A., Temple J. [2001], *GMM estimation of empirical growth models*, „CEPR Discussion Paper" 3048, Centre for Economic Policy Research, London

Bond S., Elton J., Mairesse J., Mulkay B. [1997], *Financial factor and inwestment in Belgium, France, Germany and the UK: A comparison using company panel data*, „NBER Working Paper", No. 5900

Bover O., Bentolila S., Arellano M. [2001], *The distribution of earnings in Spain during the 1980s: The effects of skill, unemployment, and union power*, „CEPR Discussion Paper", No. 2770

Bowsher C. [2002], *On testing overidentifying restrictions in dynamic panel data models*, „Economic Letters", Vol. 77, s. 211–220

Breitung J. [2000], *The local power of some unit root tests for panel data*, [w:] B. H. Baltagi (ed.), *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, „Advances in Econometrics", Vol. 15, (Elsevier Science), s. 161–178

Breitung J., Mayer W. [1994], *Testing for unit roots using panel data: are wages on different bargaining levels cointegrated?*, „Applied Economics", Vol. 26, s. 353–361

Breitung J., Pesaran M. H. [2008], *Unit roots and cointegration in panels*, [w:] L. Matyas, P. Sevestre (eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, 3rd Ed., Springer, New York, Chapter 8

Breusch T. S., Pagan A. R. [1980], *The Lagrange Multiplier Test and its Applicarions to Model Specification in Econometrics*, „Review of Economic Studies", Vol. 47, s. 239–253

Breusch T. S., Mizon G. E., Schmidt P. [1989], *Efficient estimation using panel data*, „Econometrica", Vol. 57, s. 695–700

Brown J. R., Fazzari S. M., Petersen B. C. [2009], *Financing Innovation and Growth: Cash Flow, External Equity and the 1990s R&D Boom*, „Journal of Finance", Vol. 64, s. 151–185

Bun M., Carree M. [2006], *Bias-corrected estimation in dynamic panel data models with heteroscedasticity*, „Economics Letters", Vol. 92, No. 2, s. 220–227

Byrne J., Fazio G., Piacentino D. [2005], *Convergence in TFP among Italian Regions: Panel Unit Roots with Heterogeneity and Cross Sectional Dependence*, „ERSA Conference Papers Eersa" 05p591, (European Regional Science Association)

Cameron G. [2006], *Openness, R&D, and growth at the industry level*, [w:] *Perspectives on Economic Growth*, L. A. Finley (ed.), Nova Publishers, Hauppauge (NY), Chapter 5, s. 137–158

Caselli F., Esquivel G., Lefort F. [1996], *Reopening the convergence debate: A new look at cross-country growth empirics*, „Journal of Economic Growth", Vol. 1, s. 363–389

Casu B., Girardone C. [2007], *Does competition lead to efficiency? The case of EU commercial banks*, „University of Essex Discussion Paper", No. 07-01

Chamberlain G. [1982], *Multivaraiate regression models for panel data*, „Journal of Econometrics", Vol. 18, s. 5–46

Chamberlain G. [1984], *Panel data*, [rozdz. 22], [w:] Z. Grilliches, M. Intrilligator (eds.), *Handbook of Econometrics*, North-Holland, Amsterdam

Chamberlain G., Grilliches Z. [1975], *Unobservables with a variance-components structure: ability, schooling and the economic success of brothers*, „International Economic Review", Vol. 6, s. 422–450

Chang Y. [2002], *Nonlinear IV unit root tests in panels with cross-sectional dependency*, „Journal of Econometrics", Vol. 110, s. 261–292

Chang Y. [2004], *Bootstrap unit root tests in panels with cross-sectional dependency*, „Journal of Econometrics", Vol. 120, s. 263–293

Choi I. [2001], *Unit root tests for panel data*, „Journal of International Money and Finance", Vol. 20, s. 249–272

Chojnicki Z., Czyż T. [2004], *Główne aspekty regionalnego rozwoju społeczno--gospodarczego*, [w:] J. J. Parysek (red.), *Rozwój regionalny i lokalny w Polsce w latach 1989–2002*, Bogucki Wydawnictwo Naukowe, Poznań, s. 13–24

Ciecieląg J., Tomaszewski A. [2003], *Ekonometryczna analiza danych panelowych*, WNE UW, Warszawa

Ciołek D. [2004], *Szacowanie regresji wzrostu i konwergencji dla danych panelowych. Metody ilościowe w naukach ekonomicznych*, t. 4, (red. A. Welfe), SGH, Oficyna Wydawnicza, Warszawa

Ciołek D., Golejewska A. [2005], *Wpływ bezpośrednich inwestycji zagranicznych na polski przemysł przetwórczy w oparciu o dane panelowe z lat 1993–2002. Dynamiczne Modele Ekonometryczne*, Wydawnictwo Uniwersytetu Mikołaja Kopernika, Toruń

Coe D., Helpman E. [1995], *International R&D Spillovers*, „European Economic Review", Vol. 39, s. 859–887.

Coviello D., Islam R. [2006], *Does aid help improve economic institutions?*, „World Bank Policy Research Working Paper", No. 3990

Criscuolo C. [2003], *Does ownership structure affect productivity? Evidence from the UK*, University College London, *mimeo*

Dahl C. M., Kongsted H. C., Sorensen A. [2010], *ICT and Productivity Growth in the 1990's: Panel Data Evidence on Europe*, „CREATES Research Papers", No. 2010–47, s. 1–36

Dańska B. [2000], *Przestrzenno-czasowe modelowanie zmian w działalności produkcyjnej w Polsce. Zastosowanie modeli panelowych*, t. I, [w:] *Dane panelowe i modelowanie wielowymiarowe w badaniach ekonomicznych*, B. Suchecki (red.), Absolwent, Łódź

Dańska-Borsiak B. [2008], *Wzrost gospodarczy województw. Analiza na podstawie panelowego modelu dynamicznego*, referat przedstawiony na konferencji: „Modele i prognozy w ekonomii i finansach", Olsztyn–Łańsk

Dańska-Borsiak B. [2009a], *Analiza wzrostu gospodarczego województw*, „Wiadomości Statystyczne", nr 1, s. 50–59

Dańska-Borsiak B. [2009b], *Determinanty TFP według działów przemysłu w Polsce. Dynamiczna analiza panelowa*, „Acta Universitatis Nicolai Copernici", Ekonomia XXXIX, Wydawnictwo Naukowe UMK, Toruń, s. 115–124

Dańska-Borsiak B. [2009c], *Zastosowania panelowych modeli dynamicznych w badaniach mikroekonomicznych i makroekonomicznych*, „Przegląd Statystyczny", nr 56, s. 25–41

David Simmonds Consultancy [2001], Appendix C: *Longitudinal data in transport research*, [w:] *A new look at multi-modal modeling: modeling possibilities report*, s. 79–108 http://www.davidsimmonds.com/index.php?action=download_resource&id=13&module=resourcesmodule&src..., 10.02.2009

De la Fuente A. [2000], *Convergence cross countries and regions: theory and empirics*, „CEPR Discussion Paper", No. 2465, Centre for Economic Policy Research, London

De la Fuente A., Doménech R. [2000], *Human capital in growth regressions: how much difference does data quality make?*, „OECD Economics Department Working Papers", No. 262

Deaton A. [1985], *Panel data from time series of cross-sections*, „Journal of Econometrics", Vol. 30, s. 109–26

Di Liberto A., Mura R., Pigliaru F. [2004], *How to measure the unobservable: A panel technique for the analysis of TFP convergence*, „CRENOS Working Papers", May

Drobetz W., Wanzenried G. [2006], *What determines the speed of adjustment to the target capital structure?*, „Applied Financial Economics", Vol. 16, s. 941–958

Egger P., Pfaffermayr M. [2004], *Estimating long run and short run effects in static panel models*, „Econometric Reviews", Vol. 23, s. 199–214

Engle R. F., Granger C. [1987], *Co-integration and error correction: representation, estimation and testing*, „Econometrica", Vol. 55, s. 251–276.

Everaert G., Pozzi L. [2007], *Bootstrap-based bias correction for dynamic panels*, „Journal of Economic Dynamics and Control", Vol. 31, No. 4, s. 1160–1184

Florczak W. [2007], *Kapitał ludzki a rozwój gospodarczy*, [w:] W. Welfe (red.), *Gospodarka oparta na wiedzy*, PWE, Warszawa, rozdz. 5, s. 112–172

Freeman R. B., Medoff J. L. [1981], *The impact of collective bargaining: illusion or reality*, Harvard University, *mimeo*

Frees E. W. [2005], *Longitudinal and Panel Data Analysis and Applications in the Social Science*, Cambridge University Press, Cambridge

Fujiki H., Hsiao C., Shen Y. [2002], *Is there a stable money demand function under the low interest rate policy? – A Panel Data Analysis*, „Monetary and Economic Studies", Vol. 20, No. 2, s. 1–23

Gajewski P. [2006], *Nowe koncepcje konwergencji*, [w:] S. Krajewski, P. Kaczorowski (red.), *Wzrost gospodarczy, restrukturyzacja i bezrobocie w Polsce. Ujęcie teoretyczne i empiryczne*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź

Gengenbach C., Palm F. C., Urban J. P. [2010], *Panel Unit Root Tests in the Presence of Cross-Sectional Dependencies: Comparison and Implications for Modelling*, „Econometric Reviews", Vol. 29, No. 2, s. 111–145

Glaeser E. L. [2004], *Do institutions cause, growth?*, „Journal of Economic Growth", Vol. 9, s. 271–303

Gorg H., Strobl E. [2005], *Spillovers from foreign firms through worker mobility: An empirical investigation*, „Scandinavian Journal of Economics", Vol. 107, No. 4, (Blackwell Publishing), s. 693–709

Griffith R., Redding S., Van Reen J. [2003], *R&D and absorptive capacity: theory and empirical evidence*, „Scandinavian Journal of Economics", Vol. 105, s. 99–118

Guellec D., Pottelsberghe de la Potterie B. van [2001], *R&D and productivity growth: panel data analysis of 16 OECD countries*, „OECD Economic Studies", Vol. 33, March, s. 103–126

Guest P. M. [2009], *The impact of mergers and acquisitions on executive pay in the United Kingdom*, „Economica", Vol. 76, s. 149–175

Hadri K. [2000], *Testing for unit roots in heterogeneous panel data*, „Econometrics Journal", Vol. 3, s. 148–161

Hahn J. [1999], *How informative is the initial condition in the dynamic panel model with fixed effects?* „Journal of Econometrics", Vol. 93, s. 309–326

Halaby C. [2004], *Panel models in sociological research: theory into practice*, „Annual Review of Sociology", Vol. 30, s. 507–544

Hall A. R. [2005], *Generalized Method of Moments*, Oxford University Press, Oxford

Halpern L., Murakozy B. [2005], *Does Distance Matter in Spillover?*, „CEPR Discussion Papers", No. 4857

Hansen L. P. [1982], *Large sample properties of Generalized Method of Moments estimators*, „Econometrica", Vol. 50, s. 1029–1054

Hansen L. P., Heaton J., Yaron A. [1996], *Finite sample properties of some alternative GMM estimators*, „Journal of Business and Economic Statistics", Vol. 14, s. 262–280

Harris M., Matyas L. [2004], *A comparative analysis of different IV and GMM estimators of dynamic panel data models*, „International Statistical Review", Vol. 72, No. 3, s. 397–408

Hausman J. A. [1978], *Specification tests in Econometrics*, „Econometrica", Vol. 46, s. 1251–1272

Hausman J. A., Taylor W. E. [1981], *Panel data and unobservable individual effects*, „Econometrica", Vol. 49, s. 1377–1398.

Heckman J. [1978], *Simple statistical models for discrete panel data developed and applied to test the hypothesis of true state dependence against the hypothesis of spurious state dependence*, „Colloque International du CNRS", s. 227–269

Hildreth C. [1950], *Combining cross-section data and time-series*, „Cowles Comission Discussion Paper", May 15, No. 347, [za:] Nerlove M. [2000], *An essay on the history of panel data econometrics*, paper presented at the Ninth International Conference on Panel Data, June 22–23, 2000, Geneve, Switzerland, mimeo

H i g o n D. A. [2007], *The impact of R&D spillovers on UK manufacturing TFP: A Dynamic Panel Approach*, „Research Policy", Vol. 36, No. 7, s. 964–979

H o c h I. [1962], *Estimation of production function parameters combining time-series and cross--section data*, „Econometrica", Vol. 30, s. 35–53

H o l z-E a k i n D., N e w e y W., R o s e n H. S. [1988], *Estimating vector autoregressions with panel data*, „Econometrica", Vol. 56, s. 1371–1395

H s i a o C. [2003], *Analysis of Panel Data*, 2nd Ed., Cambridge University Press, Cambridge

H s i a o C., P e s a r a n M. H. [2008], *Random coefficient panel data models*, [w:] L. M a t y a s, P. S e v e s t r e (eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, 3rd Ed., Springer, New York, Chapter 6

H s i a o C., P e s a r a n M. H., T a h m i s c i o g l u A. [1999], *Bayes estimation of short-run coefficients in dynamic panel data models*, [w:] *Analysis of Panels and Limited Dependent Variables Models*, C. Hsiao, K. Lahiri, L. F. Lee, M. H. Pesaran (eds.), Cambridge University Press, Cambridge, s. 268–296

H s i a o C., P e s a r a n M. H., T a h m i s c i o g l u A. [2002], *Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods*, „Journal of Econometrics", Vol. 109, s. 107–150

I m K. S., P e s a r a n M. H., S h i n Y. [2003], *Testing for unit roots in heterogeneous panels*, „Journal of Econometrics", Vol. 115, s. 53–74

I s l a m N. [1995], *Growth empirics: A Panel Data Approach*, „Quarterly Journal of Economics", Vol. 110, No. 4, s. 1127–1170

J a n g S. J., T h o r n b e r r y T. P. [1998], *Self-esteem, delinquent peers, and delinquency: A Test of Self-enhancement Thesis*, „American Sociological Review", Vol. 63, s. 586–598

J u d s o n R., O w e n A. [1996], *Estimating dynamic panel data models: A guide for macroeconomists*, „Economic Letters", Vol. 65, s. 9–15

J u s z c z a k-S z u m a c h e r G. [1996], *Makroekonomiczna analiza procesu produkcyjnego*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź

J u s z c z a k-S z u m a c h e r G. [2000], *Pomiar wielkości nieobserwowalnych przez analizę regresji. Postęp techniczny i efektywność postępu produkcyjnego*, [w:] A. W e l f e (red.), *Gospodarka polska w okresie transformacji*, PWE, Warszawa, rozdz. 6

K a o C. [1999], *Spurious regression and residual-based tests for cointegration in panel data*, „Journal of Econometrics", Vol. 90, s. 1–44

K e a n e M., R u n k l e D. [1992], *On the estimation of panel data models with serial correlation when instruments are not strictly exogenous*, „Journal of Business and Economic Statistics", Vol. 10, s. 1–9

K i v i e t J. F. [1995], *On bias, inconsistency, and efficiency of various estimators in dynamic panel data models*, „Journal of Econometrics", Vol. 68, s. 53–78

K n i e s n e r T., L i Q. [2002], *Semiparametric panel data models with heterogenous dynamic adjustment: theoretical consideration and an application to labor supply*, „Empirical Economics", Vol. 27, s. 131–148

K o o p G., O s i e w a l s k i J., S t e e l M. F. J. [2000], *Modeling the sources of output growth in a panel of countries*, „Journal of Business and Economic Studies", Vol. 18, No. 3, s. 284–299

K o r o l J. [2009], *Ekonometryczne modelowanie wzrostu gospodarczego regionów opartego na wiedzy*, [w:] *Taksonomia nr. 16, Klasyfikacja i analiza danych – teoria i zastosowania*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu", nr 47, s. 173–180

K u c h E. [1959], *The validity of cross-sectionally estimated behavior equations in time – series applications*, „Econometrica", Vol. 27, s. 197–214

L a c h e n m a i e r S., R o t t m a n n H. [2007], *Effects of innovation on employment: A dynamic panel analysis*, „CESIFO Working Paper", No. 2015

L a h i r i K., L i u F. [2006], *Modelling multi – period inflation ucertainty using a panel of density forecasts*, „Journal of Applied Econometrics", Vol. 21, s. 1111–1326

L e e J. [2009], *Trade, FDI, and productivity convergence: A Dynamic Panel Data Approach in 25 Countries*, „Japan and the World Economy", Vol. 21, s. 226–238

L e e J. [2002], *Panel Data Econometrics. Methods-of-Moments and Limited Dependent Variables*, An Elsevier Sciene Imprint, San Diego–San Francisco–New York

L e v i n A., L i n C. F. [1992], *Unit root test in panel data: asymptotic and finite-sample properties*, „Discussion Paper", No. 92–93, Department of Economics, University of California at San Diego, San Diego

L e v i n A., L i n C. F., C h u C. [2002], *Unit root test in panel data: asymptotic and finite sample properties*, „Journal of Econometrics", Vol. 108, s. 1–24

L o s h k i n B., B e l d e r b o s R., C a r r e e M. [2008], *The productivity effects of internal and external R&D: evidence from a dynamic panel data model*, „Oxford Bulletin of Economics and Statistics", Vol. 70 , No. 3, s. 399–413

L u c a s R. [1988], *On the mechanism of Economic development*, „Journal of Monetary Economics", Vol. 22, s. 3–42

L u c a s R. [1990], *Why doesn't capital flow from rich to poor countries*, „Americal Economic Review", Vol. 80, s. 92–96

M a C u r d y T. E. [1981], *An Empirical Model of Labour Supply in a Life Cycle Setting*, „Journal of Political Economy", Vol. 89, s. 1059–1085

M a d d a l a G. S. [1971], *The use of the variance components in pooling cross section and time series data*, „Econometrica", Vol. 39, s. 341–358

M a d d a l a G. S. [1975], *Some problems arising in pooling cross section and time series data*, „Discussion Paper University of Rochester", No. 61

M a d d a l a G. S. [1987], *Recent developments in the econometrics of panel data analysis*, „Transportation Research", Vol. 21A, s. 303–326

M a d d a l a G. S. [1994], *Some problems arising in pooling cross section and time series data*, [w:] G. S. M a d d a l a, *Econometric Methods and Applications*, Vol. 1, Edward Elgar, Aldershot, s. 223–245

M a d d a l a G. S. [1998], *Recent developments in dynamic econometric modeling: A Personal Viewpoint*, „Political Analysis", Vol. 7, s. 59–87

M a d d a l a G. S., W u S. [1999a], *A comparative study of unit root tests and a new simple test*, „Oxford Bulletin of Economics and Statistics", Vol. 61, s. 631–652

M a d d a l a G. S., W u S. [1999b], *On the use of panel data methods with cross country data*, „Annales d'Economie et de Statistique", Vol. 55–56, s. 429–448

M a d d a l a G. S., T r o s t R., L i H., J o u t z F. [1997], *Estimation of a short-run and long-run elasticities of energy demand from panel data using shrinkage estimators*, „Journal of Business and Economic Statistics", Vol. 15, s. 90–100

M a l a g a K. [2004], *Konwergencja gospodarcza w krajach OECD w świetle zagregowanych modeli wzrostu*, „Prace Habilitacyjne", nr 10, Wydawnictwo AE w Poznaniu, Poznań

M a l a g a K., K l i b e r P. [2007], *Konwergencja i nierówności regionalne w Polsce w świetle neoklasycznych modeli wzrostu*, Wydawnictwo Akademii Ekonomicznej, Poznań

M a n k i w N., R o m e r D., W e i l D. [1992], *A contribution to the empirics of economic growth*, „Quarterly Journal of Economics", Vol. 107, No. 2, s. 407–437

M a t y a s L., S e v e s t r e P. (eds.) [2008], *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, 3rd Ed., Springer, New York

M e g h i r C., W i n d m e i j e r F. [1999], *Moment conditions for dynamic panel data models with multiplicative individual effects in the conditional variance*, „Annales d'Economie et de Statistique", Vol. 55–56, s. 317–330

M i c h a ł e k J., S i w i ń s k i W., S o c h a M. (red.) [2003], *Od liberalizacji do integracji Polski z Unią Europejską. Mechanizmy i skutki gospodarcze*, Wydawnictwo Naukowe PWN, Warszawa

M i c h a ł e k J., S i w i ń s k i W., S o c h a M. (red.) [2007], *Polska w Unii Europejskiej. Dynamika konwergencji ekonomicznej*, Wydawnictwo Naukowe PNW, Warszawa

M i n c e r J. [1974], *Schooling, Experience and Earnings*, Columbia University Press, New York

M o o n H. R., P e r r o n B. [2004], *Testing for a unit root in panel with dynamic factors*, „Journal of Econometrics", Vol. 122, s. 81–126

M o r i m u n e K. [1983], *Approximate distribution of k-class estimators when the degree of overidentifiability is large compared with the sample size*, „Econometrica", Vol. 51, s. 821–841

M u n d l a k Y. [1961], *Empirical production function free of management bias*, „Journal of Farm Economics", Vol. 43, s. 44–56

N e r l o v e M. [1965], *Estimation and identification of Cobb-Douglas production function*, Rand McNally, Chicago

N e r l o v e M. [1967], *Experimental evidence on the estimation of dynamic economic relations from a time series of cross-sections*, „Economic Studies Quarterly", Vol. 18, s. 42–74

N e r l o v e M. [1971], *Further Evidence on the Estimation of Dynamic Economic Relations from a Time Series of Cross-Sections*, „Econometrica", Vol. 39, s. 359–382

N e r l o v e M. [2000], *An essay on the history of panel data econometrics*, paper presented at the Ninth International Conference on Panel Data, June 22–23, 2000, Geneva, Switzerland, *mimeo*

N e r l o v e M. [2002], *Essays in Panel Data Econometrics*, Cambridge University Press, Cambridge

N i c k e l l S. [1981], *Biases in dynamic models with fixed effects*, „Econometrica", Vol. 49, s. 1417–1426

O'M a h o n y M., V e c c h i M. [2003], *Is there an ICT impact on TFP? A heterogeneous dynamic panel approach*, „NIESR Discussion Papers", No. 219

O r d i n e P., R o s e G. [2008], *Local Banks Efficiency and Employment*, „Review of Labour Economics and Industrial Relations", Vol. 22, s. 469–493

P a r e n t e S., P r e s c o t t E. [2004], *A unified theory of the evolution of international income levels*, „Federal Reserve Bank of Minneapolis Staff Reports", No. 333

P a r k B., S i c k l e s R., S i m a r L. [2007], *Semiparametric efficient estimation of AR(1) panel data models*, „Journal of Econometrics", Vol. 117, s. 279–309

P e d r o n i P. [2004], *Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis*, „Econometric Theory", Vol. 20, s. 597–625

P e d r o n i P. [2007], *Social capital, barriers to production and capital shares: implication for the importance of parameter heterogeneity from a nonstationary panel approach*, „Journal of Applied Econometrics", Vol. 27, s. 429–451

P e s a r a n M. H. [2004], *General diagnostic tests for of cross-section dependence in panels*, „CESifo Working Paper", No. 1229

P e s a r a n M. H. [2007], *A simple panel unit root test in the presence of cross-section dependence*, „Journal of Applied Econometrics", Vol. 27, s. 265–312

P e s a r a n M. H., S m i t h R. [1995], *Estimating long-run relationships from dynamic heterogeneous panels*, „Journal of Econometrics", Vol. 68, s. 79–113

P e s a r a n M. H., S h i n Y., S m i t h R. [1999], *Pooled mean group estimation of dynamic heterogeneous panels*, „Journal of the American Statistical Association", Vol. 94, s. 621–634

P h i l i p s P. C. B., M o o n H. R. [1999], *Linear regression limit theory for nonstationary panel data*, „Econometrica", Vol. 67, s. 1057–1111

Philips P. C. B., Moon H. R. [2000], *Nonstationary panel data analysis: An Overview of Some Recent Developments*, „Econometric Reviews", Vol. 19, s. 263–286

Phillips P. C. B., Sul D. [2003], *Dynamic panel estimation and homogeneity testing under cross section dependence*, „Econometrics Journal", Vol. 6, s. 217–259

Pindyck R. S., Rubinfeld D. L. [1991], *Econometric Models and Economic Forecasts*, McGraw–Hills, New York

Pirotte A. [1999], *Convergence of the static estimation toward the long run effects of dynamic panel data models*, „Economics Letters", Vol. 63, s. 151–215

Próchniak M., Witkowski B. [2006], *Modelowanie ścieżek realnej konwergencji w skali międzynarodowej*, „Gospodarka Narodowa", nr 10, s. 1–32

Prucha I. R. [1984], *On the asymptotic efficiency of feasible Aitken estimators for seemingly unrelated regression models with error components*, „Econometrica", No. 52, s. 203–207

Quah D. [1994], *Exploiting cross-section variation for unit root inference in dynamic data*, „Economics Letters", Vol. 44, s. 9–19

Quella N. [2006], *Intra- and inter-sectoral knowledge spillovers and TFP growth rates*, „Society for Economic Dynamics", Ser. „2006 Meeting Papers", No. 431

Rapacki R. (red.) [2009], *Wzrost gospodarczy w krajach transformacji: konwergencja czy dywergencja?*, Polskie Wydawnictwo Ekonomiczne, Warszawa

Ridder G., Wansbeek T. J. [1990], *Dynamic Models for Panel Data*, [w:] R. van der Ploeg (ed.), *Advanced Lectures in Quantitative Economics*, Academic Press, New York, s. 557–582

Robertson D., Symons J. [1992], *Some strange properties of panel data estimators*, „Journal of Applied Econometrics", Vol. 7, s. 175–189

Romer P. M. [1986], *Increasing returns and long-run growth*, „Journal of Political Economy", Vol. 94, s. 1002–1037

Romer P. M. [1990], *Endogenous technological change*, „Journal of Political Economy", Vol. 98, s. 71–102

Sargan J. D. [1958], *The estimation of economic relationships using instrumental variables*, „Econometrica", Vol. 26, s. 329–338

Sarno L., Taylor M. P. [1998], *Real exchange rates under the current float: unequivocal evidence of m e a n reversion*, „Economics Letters", Vol. 60, s. 131–137

Sevestre P., Trognon A. [1985], *A note on autoregressive error-components models*, „Journal of Econometrics", Vol. 29, s. 231–245

Sevestre P., Trognon A. [1996], *Linear dynamic models*, [w:] L. Matyas, P. Sevestre (eds.), *The Econometrics of Panel Data: Handbook of Theory and Applications*, Kluwer, Boston, Chapter 7

Shehzad C. T., De Haan J., Scholtens B. [2009], *Growth and earnings persistence in banking firms: A Dynamic Panel Investigation*, „CESIFO Working Paper", No. 2772

Shell K. [1966], *Toward a theory of inventive activity and capital accumulation*, „American Economic Review", Vol. 56, No. 2, s. 62–68

Simmonds D. [2001], Appendix C: *Longitudinal data in transport research*, [w:] *David Simmonds Consultancy. A new look at multi-modal modeling: modeling possibilities report*, s. 79–108, http://www.davidsimmonds.com/index.php?action=download_resource&id=13&module=resourcesmodule&src..., 10.02.2009

Solow R. [1957], *Technical change and the aggregate production function*, „The Review of Economics and Statistics", Vol. 39, No. 3, s. 312–320

Soukiazis E., Castro V. [2004], *How the Maastricht criteria and the stability and growth pact affected the convergence process in the European Union. A panel data analysis*, „CEUNEUROP Working Paper", No. 24

Sowell F. [1996], *Optimal tests in parameter variation in the Generalized Method oh Moments framework*, „Econometrica", Vol. 64, s. 1085–1108

S t a i g e r D., S t o c k J. H. [1997], *Instrumental variables regression with weak instruments*, „Econometrica”, Vol. 65, s. 557–586

S t o c k J. H., W r i g h t J. H. [2003], *GMM with weak identification*, „Econometrica”, Vol. 68, s. 1055–1096

S t r a h l D. (red.) [2006], *Metody oceny rozwoju regionalnego*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław

S t r a h l D., B a l-D o m a ń s k a B. [2009], *Modele panelowe Produktu Krajowego Brutto w modelowaniu regionalnej innowacyjności*, „Prace Naukowe Wałbrzyskiej Wyższej Szkoły Zarządzania i Przedsiębiorczości”, nr 4 (Wałbrzych)

S t r z a ł a K. [2005], *Korelacja inwestycji i oszczędności w krajach Unii Europejskiej – weryfikacja empiryczna z zastosowaniem podejścia panelowego*, „Prace i Materiały Wydziału Zarządzania UG”: *Ekonometryczne modelowanie i prognozowanie wzrostu gospodarczego*, nr 1, s. 141–157

S u i t s D. [1984], *Dummy Variables: Mechanics vs. Interpretation*, „Review of Economics and Statistics”, Vol. 66, s. 177–180

S w a m y P. A. V. B. [1970], *Efficient Inference in a Random Coefficient Regression Model*, „Econometrica”, Vol. 38, s. 311–323.

S z t a u d y n g e r J. J. [2005], *Wzrost gospodarczy a kapitał społeczny, prywatyzacja i inflacja*, Wydawnictwo Naukowe PWN, Warszawa

Ś w i e c z e w s k a D. [2007], *Łączna produktywność czynników produkcji. Ucieleśniony kapitał wiedzy*, [w:] W. W e l f e (red.), *Gospodarka oparta na wiedzy*, PWE, Warszawa, rozdz. 4, s. 58–111

T a r l i n g R. [2008], *Statistical modelling for social researchers: principles and practice*, Taylor & Francis, London

T e s l e r L. G. [1964], *Iterative simultaneous estimation of set of linear regressions*, „Journal of American Statistical Association”, Vol. 59, s. 845–862

T h a n g a v e l u S. M., O w y o n g D. T. [2003], *The impact of export growth and scale economies on productivity in Singapore's manufacturing industries*, „Journal of Economic Studies”, Vol. 30, No. 6, s. 623–635

T o k a r s k i T. [1995], *Nakłady na badania i rozwój a wzrost dochodu narodowego i wydajność pracy*, „Wiadomości Statystyczne”, nr 1, s. 33–41

T o k a r s k i T. [1996], *Postęp techniczny a wzrost gospodarczy w modelach endogenicznych*, „Ekonomista”, nr 5, s. 581–604

T o k a r s k i T. [2002], *Determinanty wzrostu gospodarczego w warunkach stałych efektów skali*, Katedra Ekonomii Uniwersytetu Łódzkiego, Łódź

T o k a r s k i T. [2008a], *Efekty skali a wzrost gospodarczy*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków

T o k a r s k i T. [2008b], *Oszacowanie regionalnych funkcji produkcji*, „Wiadomości Statystyczne”, Vol. 10, s. 38–53

T r o g n o n A. [1978], *Miscellaneous asymptotic properties of Ordinary Least Squares and Maximum Likelihood methods in dynamic error components models*, „Annales de l’INSEE”, Vol. 30–31, s. 631–657

V e r b e e k M. [2004], *A Guide to Modern Econometrics*, 2nd Ed., Wiley & Sons, Chichester

V e r b e e k M. [2008], *Pseudo panels and repeated cross-sections*, [w:] L. M a t y a s, P. S e v e s t r e (eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, 3rd Ed., Springer, New York, Chapter 11

W a n s b e e k T. J., K a p t e y n A. [1982], *A simple way to obtain the spectral decomposition of variance components models for balanced data*, „Communications in Statistics”, Vol. A11, s. 2105–2112

W e l f e W. [2000], *Empiryczne modele wzrostu gospodarczego*, „Ekonomista”, nr 4, s. 483–499

W e l f e W. [2003], *Łączna produktywność czynników produkcji a postęp techniczny*, „Studia Ekonomiczne”, vol. 36–37, nr 1–2, s. 94–115

W e l f e W. (red.) [2001], *Ekonometryczny model wzrostu gospodarczego*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź

W e l f e W. (red.) [2007], *Gospodarka oparta na wiedzy*, PWE, Warszawa

W i n d m e i j e r F. [2005], *A finite sample correction for the variance of linear efficient two-step GMM estimators*, „Journal of Econometrics”, Vol. 126, s. 25–51

W o o l d r i d g e J. M., [2002], *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge (MA)

W o r r a l l J. L., P r a t t T. C. [2004], *Estimation Issues with Time-Series – Cross-Section Analysis in Criminology*, „Western Criminology Review”, Vol. 5, No. 1, s. 35–49

W ó j c i k P. [2004], *Konwergencja regionów Polski w latach 1990–2001*, „Gospodarka Narodowa”, nr 11–12, s. 69–85

W ó j c i k P. [2008], *Dywergencja czy konwergencja: dynamika rozwoju polskich regionów*, „Studia Regionalne i Lokalne”, Vol. 2, No. 32, s. 41–60

Y i n Y., W u S. [2000], *Stationarity tests in heterogeneous panels*, „Advances in Econometrics”, Vol. 15, s. 275–296

Z e l l n e r A. [1962], *An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias*, „Journal of the American Statistical Association”, Vol. 57, s. 348–368

Z i e n k o w s k i L. (red.) [2003], *Wiedza a wzrost gospodarczy*, Scholar, Warszawa

Z i l i a k J. P. [1997], *Efficient estimation with panel data when instruments are predetermined: An Empirical Comparison of Moment-condition Estimators*, „Journal of Business and Economic Statistics”, Vol. 15, s. 419–431

INDEKS NAZWISK Z POWOŁAŃ NA BIBLIOGRAFIĘ

Abramovitz M. 183
Acemoglu D. 208–209
Acharya R. C. 175
Aghion P. 152
Ahn S. C. 81, 87, 102–107, 120
Allison P. 21
Alonzo-Borrego C. 120
Alvarez J. 118, 120
Amemiya T. 47, 68
Anderson T. W. 30, 36, 38, 79, 81, 84, 87–88, 97, 106, 111–112, 114, 118, 148
Andini C. 124, 136–137
Andrews D. W. 55
Anselin L. 31
Arellano M. 18, 30, 32, 34–35, 38, 79, 80, 84, 87–89, 91, 93, 95, 97–98, 100, 102–103, 105–106, 109, 111, 114–120, 123–124, 126, 131–134, 137, 186
Arizala F. 152, 160
Ascari G. 168, 173, 174, 182, 190, 192, 196
Avery R. 66, 68

Bai J. 223, 225
Bal-Domańska B. 193
Balestra P. 25–27, 29
Baltagi B. H. 16–20, 31–32, 40, 68, 70, 72, 77, 103, 215–217, 221–222
Barro R. J. 183, 184–185
Bartkowiak R. 181
Bassanini A. 182
Baumol W. J. 183
Beck N. 60–61, 63–64
Bekker P. A. 120
Belderbos R. 123, 128
Benito A. 123
Bentolila S. 124, 137
Bernanke B. S. 160
Bhargava A. 38, 148
Binder M. 32
Bloom N. 124
Blundell R. 18, 34, 80, 87, 94, 97, 106–107, 109, 111, 117–118, 120, 123, 129–131, 133–134, 140, 148, 152, 186–187
Bo H. 123
Bobba M. 80, 208–210
Bond S. 17–18, 34, 78, 80, 87–89, 91, 93–95, 97, 102–103, 106–107, 109, 111, 114–118, 120–121, 123–127, 129–134, 140, 148–149, 182, 184, 186–189, 194, 197, 207–208
Bover O. 34, 79, 80, 97–98, 100, 102, 105–106, 109, 119, 124, 137
Bowsher C. 117
Breitung J. 215–216, 219
Bresson G. 31–32
Breusch T. S. 48, 53, 101
Brown J. R. 123
Bun M. 114
Byrne J. 182, 192

Cameron G. 152, 157, 175
Carree M. 114, 123, 128
Caselli F. 182, 186–188, 195, 197
Castro V. 182
Casu B. 123, 179
Cavallo E. 152, 160
Chamberlain G. 21, 35, 104, 185–186
Chang Y. 221–222
Choi I. 216, 219, 220
Chojnicki Z. 181
Chu C. 193, 216–218
Ciotek D. 193, 197
Coe D. 175
Coviello D. 80, 179, 208–210
Criscuolo C. 124, 139, 141
Czyż T. 181

Dahl C. M. 152
Dańska B. 8, 40, 66, 67, 70, 72

Dańska-Borsiak B. 152, 172, 202
De Haan J. 123
De la Fuente A. 182, 184
Deaton A. 30
Di Cosmo V. 168, 173–174, 182, 190, 192, 196
Di Liberto A. 179
Doménech R. 182
Drobetz W. 124, 142

Egger P. 78
Elton J. 123, 124
Engle R. F. 77
Esquivel G. 182, 186–188, 195, 197
Everaert G. 114

Fazio G. 182, 192
Fazzari S. M. 123
Florczak W. 156
Freeman R. B. 20
Frees E. W. 33
Fujiki H. 21–22

Gajewski P. 183
Galindo A. 152, 160
Gengenbach C. 225
Girardone C. 123, 179
Glaeser E. L. 209
Gorg H. 124, 144
Granger C. 77
Griffin J. M. 31–32, 77
Griffith R. 152, 175
Grilliches Z. 21
Guellec D. 182
Guest P. M. 124, 143
Gurkaynak R. S. 160

Hadri K. 220–221
Hahn J. 111
Halaby C. 208
Hali A. R. 233, 234
Hansen L P. 79, 88, 100, 116, 227, 229, 234–236
Harris M. 118–119
Hausman J. A. 39, 50–55, 93, 98–100, 105
Heaton J. 235
Heckman J. 85
Helpman E. 175
Hemmings P. 182
Hernando I. 123
Higon D. A. 124
Hildreth C. 23
Hoch I. 23–24
Hoeffler A. 18, 91, 95, 121, 148–149, 182, 184, 186–189, 194, 197, 207–208
Holz-Eakin D. 88
Honoré B. 32, 34, 38
Howitt P. 152
Hsiao C. 16, 18, 20–22, 30, 32, 34, 36, 38, 75–76, 79, 81, 84, 88, 97, 106, 111–112, 114, 118, 128, 148, 180, 215

Im K. S. 193, 217– 219
Islam N. 182, 185–186
Islam R. 179

Jang S. J. 21
Jayet H. 31
Joutz F. 31
Judson R. 118–119
Juszczak-Szumacher G. 154, 156, 195

Kao C. 215
Kapteyn A. 46
Katz L. N. 60–61, 63–64
Keane M. 35
Keller W. 175
Kiviet J. F. 81, 111–114, 118, 148
Kliber P. 11, 184, 193
Kniesner T. 35
Kongsted H. C. 152
Koop G. 182
Kuch E. 23

Lachenmaier S. 123, 135
Lahiri K. 32
Le Gallo J. 31
Lee J. 22, 55, 182
Lefort F. 182, 186–188, 195, 197
Levin A. 193, 215–218
Li Q. 31, 35, 158–159, 173
Lin C. F. 193, 215–218
Liu F. 32
Loshkin B. 123, 128
Lucas R. 183

Maddala G. S. 31, 36, 62, 217–219, 221
Mairesse J. 123–124

Malaga K. 11, 184, 193
Mankiw N. 185, 188
Matyas L. 118–119
Matyas L. (red.) 32
Mayer W. 216
Medoff J. L. 20
Meghir C. 35
Michałek J. (red.) 193
Mincer J. 136
Mizon G. E. 53, 101
Moon H. R. 32, 215–216, 221–223, 225
Morimune K. 120
Mulkay B. 123–124
Mundlak Y. 23
Mura R. 179

Nerlove M. 22–23, 25– 27, 29, 37, 77, 83
Newey W. 88
Ng S. 223, 225
Nickell S. 74–75, 112, 158, 180

O'Mahony M. 152, 162
Ordine P. 123
Osiewalski J. 182
Owen A. 118–119
Owyong D. T. 152

Pagan A. R. 48
Palm F. C. 225
Park B. 35
Pedroni P. 32
Perron B. 32, 221–223, 225
Pesaran M. H. 30–32, 75, 77, 128, 148, 163, 193, 215, 217–219, 221, 224–225
Petersen B. C. 123
Pfaffermayr M. 78
Philips P. C. B. 215–216, 221, 223
Piacentino D. 182, 192
Pigliaru F. 179
Pindyck R. S. 169
Pirotte A. 31–32, 77–78
Pottelsberghe de la Potterie B. van 182
Pozzi L. 114
Prantl S. 152
Pratt T. C. 62
Próchniak M. 193
Prucha I. R. 68

Quah D. 216
Quella N. 152

Rapacki R. (red.) 193
Redding S. 175
Ridder G. 37–38, 84
Robertson D. 31
Romer D. 185
Romer P. M. 183, 185, 188
Rose G. 123
Rosen H. S. 88
Rottmann H. 123, 135
Rubinfeld D. L. 169
Runkle D. 35

Sala-i-Martin X. 183–185
Sargan J. D. 38, 79, 116, 148, 236
Sarno L. 193
Scarpetta S. 182
Schankerman M. 124
Schmidt P. 53, 81, 87, 101–107, 120
Scholtens B. 123
Sevestre P. 36–37, 75, 95, 112
Sevestre P. (red.) 32
Shehzad C. T. 123
Shen Y. 21–22
Shin Y. 163, 193, 217–219
Sickles R. 35
Simar L. 35
Siwiński W. (red.) 193
Smith R. 31, 75, 77, 148, 163
Socha M. (red.) 193
Solow R. 152, 182–183
Sorensen A. 152
Soukiazis E. 182
Sowell F. 233–234
Staiger D. 80
Steel M. F. J. 182
Stock J. H. 80, 235
Strahl D. 193
Strobl E. 124, 144
Strzała K. 215
Suits D. 43
Sul D. 221, 223
Swamy P. A. V. B. 30, 40
Symons J. 31
Sztaudynger J. J. 193, 195

Świeczewska D. 152–154, 156

Tahmiscioglu A. 32, 128, 148
Tarling R. 208

Taylor W. E. 39, 51–54, 93, 98–99, 100, 105, 193
Temple J. 18, 91, 95, 121, 148–149, 182, 184, 186–189, 194, 197, 207–208
Tesler L. G. 27
Thangavelu S. M. 152
Thornberry T. P. 21
Tokarski T. 152, 168, 172–173, 180, 193, 195, 201
Trognon A. 36–37, 75, 77, 95, 112
Trost R. 31

Urban J. P. 225

Van Reen J. 175
Van Reenen J. 124
Vecchi M. 152, 162
Verbeek M. 30, 228, 232

Wansbeek T. J. 37–38, 46, 84
Wanzenried G. 124, 142
Weil D. 185, 188
Welfe W. 152, 156, 168
Welfe W. (red.) 152, 156, 168, 183, 193, 195
Windmeijer F. 35, 81, 96–97, 111, 117–118, 127, 133
Witkowski B. 193
Wooldridge J. M. 21, 32, 52, 108
Worrall J. L. 62
Wójcik P. 193, 201–202
Wright J. H. 235
Wu S. 217, 219, 221

Yaron A. 235
YinY. 221

Zellner A. 27, 66, 68
Zienkowski L. (red.) 152, 156, 168
Ziliak J. P. 120

INDEKS RZECZOWY

błędy pomiaru zmiennych 18, 21, 129, 146–149
dane czasowo-przekrojowe 14–15, 55–66
dane wzdłużne (ang. *longitudinal data*) 14, 29
dwustopniowy estymator FDGMM 88, 91, 94, 114, 121, 127, 132–133, 176, 200
dwustopniowy estymator SGMM 110, 121, 176, 200, 209
efekty czasowe 15, 23, 33
efekty grupowe 15, 22–24, 33, 35, 38, 40, 43–45, 48–53, 73–74, 76, 80, 83, 88, 92, 94, 98, 105, 112, 119
estymator Ahna i Schmidta 102–106, 120
estymator Andersona–Hsiao 81–87, 119–120, 132
estymator Arellano i Bover 97–102
estymator dwustopniowej metody najmniejszych kwadratów dla modeli panelowych (2MNK-MP) 70
estymator GMM pierwszych różnic (FDGMM) 80, 87, 90–91, 94–95, 97, 107, 110–111, 118, 120
estymator Hausmana-Taylora 39, 51–55, 98, 100
estymator międzygrupowy 69
estymator minimalnej odległości Chamberlaina 16, 104
estymator SGMM 80, 109–111, 118–121
estymator *shrinkage* 31, 32
estymator trójstopniowej metody najmniejszych kwadratów dla modeli panelowych (PM3MNK) 71
estymator UMNK dla modelu RE 48
estymator wewnątrzgrupowy 43, 48, 54–55, 74–78, 85, 95, 111–114, 119
funkcja Cobba–Douglasa 24, 128–129, 133, 140, 154–155, 160, 168, 173, 182, 190–193
jednostopniowy estymator FDGMM 88, 114, 119–121, 126, 133, 142, 176, 204, 209
jednostopniowy estymator SGMM 107, 110, 119–121, 134, 176, 204, 209
klasyczna metoda najmniejszych kwadratów 26, 29, 42–43, 44, 47, 57, 61–62, 76–77, 95, 119–120
konwergencja 181, 182–185, 188–189, 192–193, 194–195, 197, 200–202, 207
korekta Kivieta 111–114, 120, 181, 197
korekta Windmeijera (skorygowany estymator wariancji estymatora FDGMM2) 96–97, 127, 133, 135
łączna produktywność czynników produkcji (TFP) 141, 145–146, 151–180, 182, 189–193
makropanel 17–18, 30, 123
metoda GMM pierwszych różnic (FDGMM) 34, 80, 87–94, 118–121, 125–126, 128, 130–134, 147–149, 176, 179, 200, 207
metoda największej wiarygodności (MNW) 16, 27, 34, 81–83
metoda *panel corrected standard errors* (PCSE) 61–62
metoda systemowa GMM (SGMM) 34, 106–110, 119–121, 128, 130, 133–134, 137, 140, 147–149, 176, 179, 200, 207
metoda zmiennych instrumentalnych (IV) 51, 53, 81–86, 228, 230
mikropanel 17–18, 30
model autokorelacyjny 85–86
model autoregresyjny 35
model korekty błędem 77, 163
model *one-way* 40
model *pooled* 26, 44, 48, 57, 179
model *two-way* 33, 40
model z dekompozycją składnika losowego (z efektami losowymi, RE) 27, 30, 33, 44–18

model ze zmiennymi sztucznymi (z efektami nielosowymi, FE) 27, 30, 40–44
model zależności stanowej 85–86, 112
modele cenzurowane 15
modele danych niekompletnych 16
modele proste o równaniach pozornie niezależnych (SUR) 66–68
modele statyczne 39, 73–78
modele ucięte 15
modele wielorównaniowe o równaniach łącznie współzależnych 68–72
modele z heteroskedastycznością grupową i korelacją przekrojową składnika losowego 58–62
modele z heteroskedastycznością grupową 56–58, 170
modele z heteroskedastycznością grupową, korelacją przekrojową i autokorelacją składnika losowego 62–66
modele z losowymi parametrami 30
modele zmiennych jakościowych 15
operator międzygrupowy 46, 68–69
operator wewnątrzgrupowy 42, 46, 68–69
panel niezbilansowany 14
panel zbilansowany 14
panelowe modele przestrzenne 30–31
przekształcenie Praisa–Winstena 65
pseudopanel 30
różnicowy test Sargana 117–118, 141, 179, 188, 209, 210
stacjonarność w panelach 162, 215–225
szeregi czasowe o wysokim stopniu trwałości (ang. *persistent time series*) 18, 80, 95, 106, 130, 134, 140, 185, 191, 200, 209
test autokorelacji Arellano–Bonda 115–116, 126–127, 132, 134, 179, 200, 207, 210
test Chowa 44
test Durbina–Watsona 63
test Hausmana 50–51
test Hausmana–Taylora warunków identyfikujących 54–55
test ilorazu wiarygodności 57, 60
test mnożnika Lagrange'a 57, 63
test mnożnika Lagrange'a (Breuscha–Pagana) 48, 59
testy pierwiastków jednostkowych w panelach 215–225
– test Phillipsa i Sula 223
– test Moona i Perrona 221–223, 225
– test Chang 222
– test Hadri'ego 162, 220–221
– test Ima, Pesarana i Shina (IPS) 162, 193, 217–219, 221, 222
– test Levina, Lina i Chu (LLC) 193, 217–218
– test Bai i Ng kombiancji *p-value* 223–224, 225
– test Maddali i Wu 217, 219–220
– tesr Pesarana CADF 224–225
– test typu Fishera 217, 219–220
test Sargana 116–117, 126–127, 130, 132–134, 141, 149, 179, 188, 200, 207, 210, 235–236
test Walda heteroskedastyczności grupowej 58
uogólniona metoda momentów (GMM) 34, 36, 78, 87, 146, 227–236
uogólniona metoda najmniejszych kwadratów (UMNK) 28–29, 34–37, 47, 68, 70–71
uogólniona metoda najmniejszych kwadratów z estymacją (EUMNK) 57–58, 60
wartości początkowe 35–38, 76, 80–85,106–108, 112
warunki momentów 79–80, 87–90, 92–93, 96, 99, 101–104, 106–107, 109–110, 115, 119–120, 179, 227, 229–233, 235–236
warunki początkowe 35–38, 79–81, 106–108
ważona metoda najmniejszych kwadratów (WMNK) 58
zmienna instrumentalna 51, 53, 79, 84, 87–89, 92–94, 99–102, 109–110, 186–189, 230, 235
zmienna objaśniająca endogeniczna 17, 52, 78, 79, 132, 135, 137, 141, 148–149, 176, 179, 184–186, 188, 200, 204, 228
zmienna objaśniająca ściśle egzogeniczna 34, 51, 79, 92, 105, 110, 176, 179, 200, 204
zmienna objaśniająca z góry ustalona 17, 79, 92, 176, 179, 200, 204