



UNIWERSYTET
EKONOMICZNY
W POZNANIU

mgr Tomasz Józefowski

Rynek pracy na obszarach funkcjonalnych
miast wojewódzkich w świetle
estymacji typu SPREE

The labor market in functional areas of provincial
capital cities in the light of SPREE estimation

Praca doktorska

Promotor: dr hab. Grażyna Dehnel, prof. UEP

Promotor pomocniczy: dr hab. Marcin Szymkowiak, prof. UEP

Pracę przyjęto dnia:

Podpis promotora:

Poznań 2022

Spis treści

Skróty	1
Wstęp	3
1. Wprowadzenie do problematyki miejskich obszarów funkcjonalnych	11
1.1. Istota miejskich obszarów funkcjonalnych	12
1.2. Identyfikacja obszarów funkcjonalnych w świetle ustaw i strategii krajowych	16
1.3. Wymiar terytorialny na poziomie Unii Europejskiej	24
1.4. Delimitacja miejskich obszarów funkcjonalnych	31
1.5. Informacja statystyczna w kontekście miejskich obszarów funkcjonalnych .	36
1.6. Podsumowanie	42
2. Wybrane źródła danych w obszarze rynku pracy	45
2.1. Badania prowadzone poprzez gospodarstwa domowe	46
2.2. Administracyjne źródła danych	53
2.3. Badanie bezrobocia rejestrowanego	59
2.4. Badania przedsiębiorstw i jednostek sfery budżetowej	60
2.5. Spisy powszechne	62
2.6. Możliwości i ograniczenia poszczególnych źródeł danych - analiza porów- nawcza	64
2.7. Podsumowanie	73

3. Estymacja typu SPREE	75
3.1. Wprowadzenie do estymacji pośredniej	76
3.2. Idea estymacji typu SPREE	84
3.3. Estymacja typu SPREE – ujęcie klasyczne	87
3.4. GSPREE dla tabeli dwuwymiarowej	89
3.5. GSPREE dla trójwymiarowej tabeli kontyngencji	92
3.6. Estymacja parametrów modeli GSPREE	95
3.7. Inne rozszerzenia estymatora typu SPREE	99
3.8. Przegląd praktycznych zastosowań estymatorów typu SPREE	102
3.9. Podsumowanie	107
4. Rynek pracy na obszarach funkcjonalnych miast wojewódzkich - procedura badawcza	109
4.1. Założenia procedury badawczej	110
4.2. Opis zakresu wykorzystanych źródeł danych	114
4.3. Zakres szacowanych kategorii i podkategorii zmiennych	117
4.4. Rozważane estymatory i ocena ich jakości	119
4.5. Rozważane układy tabel	123
4.6. Analiza statystyczna uzyskanych szacunków dla kategorii rynku pracy	128
4.7. Podsumowanie	136
5. Wskaźniki rynku pracy dla obszarów funkcjonalnych miast wojewódzkich	139
5.1. Ocena jakości oszacowań analizowanych wskaźników rynku pracy	140
5.2. Ocena sytuacji na rynku pracy – ujęcie przestrzenne	149
5.3. Taksonomiczna analiza zróżnicowania rynku pracy w miejskich obszarach funkcjonalnych ośrodków wojewódzkich	159
5.4. Podsumowanie	164
Zakończenie	172
Bibliografia	173
Aneks	191

Spis treści	iii
Spis tablic	209
Spis rysunków	210

Skróty

AZD	administracyjne źródła danych
BAEL	Badanie Aktywności Ekonomicznej Ludności
BJS	Baza Jednostek Statystycznych
BR	Bezrobocie rejestrowane
CRP KEP	Centralny Rejestr Podmiotów Krajowej Ewidencji Podatników
CV	współczynniki zmienności (ang. <i>Coefficient of Variation</i>)
GLSM	uogólniony liniowy model strukturalnym (ang. <i>Generalized Linear Structural Model</i>)
GLSMM	uogólniony mieszany liniowy model strukturalny (ang. <i>Generalized Linear Structural Mixed Model</i>)
GUS	Główny Urząd Statystyczny
IPF	metoda iteracyjnego proporcjonalnego dopasowania (ang. <i>Iterative Proportional Fitting</i>)
KPM 2023	Krajowa Polityka Miejska 2023
KPZK 2030	Koncepcja Przestrzennego Zagospodarowania Kraju 2030
KSRR 2030	Krajowa Strategia Rozwoju Regionalnego 2030
MOF OR	miejskie obszary funkcjonalne ośrodków regionalnych
MOF OS	miejskie obszary funkcjonalne ośrodków subregionalnych
MOF OW	miejskie obszary funkcjonalne ośrodków wojewódzkich
MOP	Międzynarodowa Organizacja Pracy
NSP 2011	Narodowy Spis Powszechny Ludności i Mieszkań 2011
NSP 2021	Narodowy Spis Powszechny Ludności i Mieszkań 2021
POJT	Pozostały obszar jednostki terytorialnej
PwGN	Pracujący w gospodarce narodowej
REGON	Krajowy Rejestr Urzędowy Podmiotów Gospodarki Narodowej
RSE	względny błąd standardowy (ang. <i>Relative Standard Error</i>)

SMO	statystyka małych obszarów
SOR	Strategia na rzecz Odpowiedzialnego Rozwoju
SPREE	estymator zachowujący strukturę (ang. <i>Structure Preserving Estimation</i>)
ZIT	Zintegrowane Inwestycje Terytorialne
ZROPiB	Zintegrowany Rejestr Osób Pracujących i Bezrobotnych

Wstęp

Monitorowanie i diagnoza sytuacji na rynku pracy to jeden z kluczowych elementów sprawnego zarządzania państwem. Rzetelne informacje w tym zakresie umożliwiają podejmowanie działań pozwalających na ograniczanie negatywnych skutków zachodzących zmian. Z kolei bycie członkiem Unii Europejskiej obliguje Polskę do przekazywania takich danych o rynku pracy, które umożliwiają dokonywanie wiarygodnych porównań pomiędzy krajami członkowskimi. Stąd też zachodzi konieczność wypracowania metodyki badań pozwalających na charakterystykę polskiego rynku pracy, która będzie respektowała wspólne wytyczne.

Główny Urząd Statystyczny (GUS), realizując swoje zadania statutowe, prowadzi cykliczne Badanie Aktywności Ekonomicznej Ludności (BAEL). Podstawowym celem tego badania jest pozyskanie informacji o rozmiarach i strukturze populacji osób pracujących, bezrobotnych oraz biernych zawodowo z uwzględnieniem ich cech demograficzno-społecznych. Jednak przy analizie wyników z BAEL należy brać pod uwagę ograniczenia, które wynikają z reprezentacyjnej metody badania. Z powodu zastosowanego podejścia, badanie to nie jest w stanie zapewnić wiarygodnych oszacowań dla bardziej szczegółowych przekrojów, gdyż mała liczba obserwacji w odpowiednich domenach może prowadzić do dużych błędów szacunku. Dlatego też najniższym poziomem agregacji przestrzennej, na którym udostępnia się wyniki z badania BAEL, jest poziom województwa, gdzie jak do tej pory nie występują pozycyjne braki informacji dla podstawowych charakterystyk rynku pracy (GUS, 2018d).

Obecny poziom dostępności danych wynika też z ograniczonych środków finansowych na realizację badania oraz dążenia do zmniejszania obciążeń respondentów, którzy są coraz mniej skłonni do brania udziału w badaniach ankietowych. Z drugiej jednak strony warto dodać, że odbiorcy danych statystycznych oczekują większego zakresu dostarczanych informacji niż tego, jaki można zapewnić na podstawie realizowanego badania reprezentacyjnego.

Wobec rosnących potrzeb informacyjnych, zarówno na arenie krajowej, jak i międzynarodowej, oczekuje się zatem, że statystyka publiczna dostarczy rzetelnych informacji dla wcześniej nieplanowanych domen w badaniach reprezentacyjnych. Jednym z pożądanych przekrojów, dla którego istnieje obecnie duże zapotrzebowanie na wszelkiego rodzaju informacje (w tym na temat sytuacji na rynku pracy), są miasta wraz ze swoimi obszarami funkcjonalnymi. W literaturze przedmiotu istnieje wiele sposobów definiowania i wyznaczania tych obszarów (Śleszyński, 2015). Ze względu na rosnące ich znaczenie pewne regulacje zostały wprowadzone w krajowych dokumentach strategicznych. W Koncepcji Przestrzennego Zagospodarowania Kraju 2030 (KPZK 2030) (Rada Ministrów, 2012) miejski obszar funkcjonalny (MOF) rozumiany jest jako ciągły przestrzennie układ osadniczy złożony z odrębnych jednostek administracyjnych. Obejmuje on zwarty obszar miejski oraz powiązaną z nim funkcjonalnie strefę zurbanizowaną. Wyróżniono w nim cztery podtypy MOF: obszary funkcjonalne ośrodków wojewódzkich, regionalnych, subregionalnych oraz lokalnych. Popyt na informacje w tych przekrojach terytorialnych wynika przede wszystkim z faktu, iż polityka miejska stanowi integralną część wymiaru terytorialnego w okresie realizacji programu polityki spójności na lata 2014–2020 i pozostaje jednym z ważniejszych priorytetów wsparcia finansowego realizowanego w ramach Unii Europejskiej. W tym kontekście miejskie obszary funkcjonalne nabrały szczególnego znaczenia, a zmiana podejścia do planowania rozwoju stworzyła nową możliwość wykorzystywania środków i instrumentów wspólnotowych służących wspieraniu miejskich obszarów funkcjonalnych. Szczególnym przykładem takiego instrumentu zaplanowanego przez Komisję Europejską są Zintegrowane Inwestycje Terytorialne (ZIT). W Polsce w latach 2014–2020 ZIT-y realizowano na obszarach funkcjonalnych miast wojewódzkich. O aktualności tematyki związanej z obszarami funkcjonalnymi świadczą m.in. artykuły naukowe (Batóg i in., 2015; Kociuba, 2015, 2017b,a, 2018b,a; Szafranek, 2015, 2017b,a; Szafranek i Kociuba,

2018; Szlachta, 2017; Śleszyński, 2015) podejmujące temat ich delimitacji i funkcjonowania.

Dla tak wyznaczonych jednostek przestrzennych dostępne są obecnie jedynie informacje pochodzące z badań reprezentacyjnych, które nie zapewniają szczegółowych danych, m.in. na temat rynku pracy. Przyczyną tego jest zbyt mała liczba reprezentantów dla bardziej szczegółowych przekrojów. W takiej sytuacji klasyczna estymacja bezpośrednia stosowana w statystyce publicznej prowadzi do szacunków charakteryzujących się niską precyzją.

Remedium na tego typu ograniczenia mogą stanowić metody statystyki małych obszarów (SMO)¹, które umożliwiają uzyskiwanie wiarygodnych szacunków na niższych poziomach agregacji przestrzennej lub dla bardziej szczegółowo zdefiniowanych domen. Metody te daje się stosować nawet w przypadkach, gdy w danym przekroju nie ma żadnego reprezentanta w próbie (Rao i Molina, 2015). Metody SMO znajdują zastosowanie w wielu obszarach badawczych, takich jak ubóstwo, niepełnosprawność, statystyka gospodarcza czy rynek pracy (Dehnel i Wawrowski, 2020; Gołata, 2010; Szymkowiak i in., 2017b; Klimanek, 2017; Wilak, 2020; Józefowski i Szymkowiak, 2013).

W tym miejscu warto podkreślić, że chociaż szacunki otrzymane za pomocą różnych metod SMO charakteryzują się na ogół lepszą precyzją w porównaniu do estymacji klasycznej, wartości uzyskane dla mniejszych jednostek terytorialnych w wyniku ich agregacji często nie są zgodne z wynikami otrzymanymi na podstawie estymacji bezpośredniej w ramach metody reprezentacyjnej na wyższym poziomie². Jest to jeden z elementów, który ogranicza powszechne wykorzystanie tych technik w badaniach realizowanych w statystyce publicznej. Jednym ze sposobów radzenia sobie z powyżej opisanym brakiem spójności oszacowań jest zastosowanie estymatorów typu SPREE (ang. *Structure Preserving Estimation*). Estymatory te przenoszą znaną z rejestrów administracyjnych, bądź spisów powszechnych, strukturę powiązań dla kategorii zmiennych pomocniczych i badanych domen, przy jednoczesnym zachowaniu spójności z szacunkami na wyższym poziomie agregacji. Jest to bardzo pożądana cecha estymatora, zwłaszcza z punktu widzenia wymogów co do jakości informacji publikowanych przez krajowe urzędy statystyczne (Paradysz, 2009).

¹W literaturze przedmiotu określane również mianem estymacji pośredniej czy estymacji dla małych domen.

²Przykładowo oszacowana z wykorzystaniem metod SMO liczba osób bezrobotnych w powiatach pewnego województwa może nie sumować się do oszacowanej liczby bezrobotnych z wykorzystaniem estymacji bezpośredniej na tym poziomie agregacji przestrzennej.

Technika SPREE wydaje się być szczególnie wartościowa w przypadku estymacji dla nietypowych przekrojów, które nie pokrywają się z przyjętym podziałem administracyjnym. Do tej pory, na przykład, nie została wypracowana metoda, która pozwoliłaby na publikowanie informacji o osobach pracujących, bezrobotnych oraz biernych zawodowo (według definicji Międzynarodowej Organizacji Pracy) na poziomie miejskich obszarów funkcjonalnych miast wojewódzkich (MOF OW).

Niniejsza praca ma przyczynić się do częściowego wypełnienia tej luki. Stanowi bowiem próbę wypracowania metody estymacji charakterystyk rynku pracy na obszarach funkcjonalnych miast wojewódzkich z wykorzystaniem estymatorów typu SPREE, które zapewniają zgodność i spójność informacyjną z szacunkami bezpośrednimi (w tym kalibracyjnymi) na wyższym poziomie agregacji przestrzennej. Jest to zarazem jedna z pierwszych tak kompleksowych prac w Polsce poświęconych tej klasie estymatorów pośrednich, a o ile mi wiadomo pierwsza dotycząca miejskich obszarów funkcjonalnych w kontekście rynku pracy.

Cele i hipotezy badawcze

Głównym celem rozprawy jest pomiar i ocena sytuacji na rynku pracy na obszarach funkcjonalnych miast wojewódzkich w 2017 roku w świetle estymacji typu SPREE. Została ona przeprowadzona na podstawie trzech wskaźników: współczynnika aktywności ekonomicznej, wskaźnika zatrudnienia i stopy bezrobocia w przekrojach zdefiniowanych przez płeć i grupy wieku (15–29, 30–39, 40–49, 50 lat i więcej). Ważnym aspektem jest także uzyskanie oszacowań i ich precyzji dla liczby osób pracujących, bezrobotnych oraz biernych zawodowo w zakładanych przekrojach. Informacje te są niezbędnym elementem do wyznaczenia wymienionych wyżej wskaźników dotyczących rynku pracy. Istotnym celem pracy była również wszechstronna charakterystyka estymacji typu SPREE.

Realizacji tak sformułowanych celów podporządkowana została struktura niniejszej rozprawy, którą podzielono na pięć zasadniczych rozdziałów. Pracę rozpoczyna wprowadzenie do problematyki obszarów funkcjonalnych. W pierwszym rozdziale dokonany został także przegląd definicji i metodyki wyznaczania miejskich obszarów funkcjonalnych. W szczególności analizie poddane zostały miejskie obszary funkcjonalne objęte Zintegrowanymi Inwestycjami Terytorialnymi, z uwzględnieniem ich znaczenia w kształtowaniu polityki spójności Unii Europejskiej. Naświetlone zostało również znaczenie tych obszarów w świetle

krajowych dokumentów strategicznych. Przeanalizowano obecne pokrycie informacyjne, szczególnie w odniesieniu do rynku pracy. Rozpoznano także zgłaszane potrzeby odbiorców danych w tym zakresie.

Kolejny etap realizacji prac stanowiła analiza źródeł danych dotyczących rynku pracy w odniesieniu do miejskich obszarów funkcjonalnych. Drugi rozdział obejmuje zatem takie zagadnienia jak: przestrzenny poziom dostępności danych, częstotliwość ich uzyskiwania oraz ich zgodność definicyjną, a także jakość uzyskiwanych informacji o liczbie osób pracujących, bezrobotnych oraz biernych zawodowo na podstawie badań pełnych i reprezentacyjnych, spisów powszechnych i administracyjnych źródeł danych. Przeprowadzono analizę porównawczą źródeł zmiennych pomocniczych (spisy powszechne, rejestry administracyjne), szczególnie pod względem ich aktualności i spójności definicyjnej oraz zależności względem szacowanych parametrów zmiennych. Działania te posłużyły do wytypowania potencjalnych źródeł zmiennych pomocniczych przydatnych do wykorzystania w toku dalszych prac związanych z procesem estymacji pośredniej.

Trzeci rozdział zawiera kompleksowy przegląd zagadnień związanych ze statystyką małych obszarów oraz ideą estymacji typu SPREE. Przedstawiono szczegółowe opisy estymatorów typu SPREE, zarówno w ujęciu klasycznym, jak i modelowym, z uwzględnieniem strony matematycznej oraz własności, którymi się charakteryzują. Syntetycznie zaprezentowano też dalszą ewolucję tych metod. Omówiono kwestie oceny jakości estymacji oraz wskazania i rekomendacje co do stosowania metod pośrednich. Przedstawiono również dotychczasowe, wybrane przypadki użycia wskazanej wyżej grupy estymatorów, w różnych sferach społeczno-ekonomicznych.

Rozdział czwarty opisuje z kolei cały proces badawczy, wykorzystane źródła danych oraz zakres zmiennych wziętych pod uwagę w pracach empirycznych (w tym ustalony zbiór miejskich obszarów funkcjonalnych), zastosowane warianty estymatorów oraz przyjęte metody ich oceny, a także rozważane układy tabel wynikowych. W rozdziale tym szczegółowo przeanalizowano wyniki oszacowań (liczbę osób pracujących, bezrobotnych oraz biernych zawodowo) dla wszystkich rozważanych wariantów estymatorów pod względem jakości oszacowań. W ocenie uwzględniono przyjęte przekroje przestrzenne i kategorie rynku pracy z uwzględnieniem płci i czterech grup wieku. Bazując na przedstawionych wynikach analizy jakości uzyskanych szacunków wybrano „optymalny” estymator, który posłużył do

wyznaczenia najważniejszych wskaźników: współczynnika aktywności zawodowej, wskaźnika zatrudnienia oraz stopy bezrobocia.

Ostatni, piąty rozdział rozprawy, zawiera analizę jakości oszacowań współczynnika aktywności ekonomicznej, wskaźnika zatrudnienia oraz stopy bezrobocia dla przyjętego w niniejszej pracy docelowego przekroju. Ponadto dokonano przestrzennej oceny zróżnicowania sytuacji na rynku pracy w miejskich obszarach funkcjonalnych ośrodków wojewódzkich z wykorzystaniem wybranych technik taksonomicznych. Analizie poddana została również sytuacja na rynku pracy na obszarach MOF OW w porównaniu do pozostałej części województwa. Uzyskane wyniki oszacowań zostały ocenione w odniesieniu do sformułowanych hipotez badawczych.

W związku z realizacją celów postawionych w rozprawie sformułowane zostały następujące cztery hipotezy badawcze:

- 1) Estymacja typu SPREE pozwala na oszacowanie podstawowych charakterystyk rynku pracy na poziomie agregacji przestrzennej, dla którego obecnie nie ma pokrycia informacyjnego,
- 2) Estymacja typu SPREE cechuje się lepszą precyzją szacunków charakterystyk rynku pracy w porównaniu z estymacją kalibracyjną stosowaną w badaniu BAEL,
- 3) Sytuacja na rynku pracy na obszarach funkcjonalnych miast wojewódzkich, z uwzględnieniem płci i grup wieku, jest lepsza niż w pozostałej części województwa,
- 4) Rynek pracy w ujęciu obszarów funkcjonalnych miast wojewódzkich cechuje się wyraźnym zróżnicowaniem przestrzennym.

Novum niniejszej rozprawy stanowi propozycja autorskiego przebiegu procesu badawczego w zakresie estymacji pośredniej, który bazuje na wskazaniach międzynarodowych. Empiryczne przeprowadzenie zaproponowanego procesu badawczego jest, o ile nam wiadomo, pierwszą w Polsce próbą uzyskania i analizy informacji na tym poziomie agregacji. Oryginalnym wkładem niniejszej rozprawy jest także pierwsze kompleksowe zastosowanie estymatorów typu SPREE do danych o polskim rynku pracy w układach, dla których informacje nie są publikowane w ramach statystyki publicznej. Warto również podkreślić, że przetestowane i przyjęte w rozprawie metody mogą stanowić uniwersalne narzędzie do estymacji różnych charakterystyk rynku pracy na poziomie obszarów funkcjonalnych miast wojewódzkich.

Zaproponowane w pracy wykorzystanie połączonych danych pochodzących z badania reprezentacyjnego, z rejestrów administracyjnych oraz szacunków bieżących struktur ludności jest zgodne ze sformułowanymi przez Eurostat wytycznymi na temat procesów integracji danych w celu zwiększenia zakresu informacyjnego. Stanowi ono również, w naszej ocenie wkład w rozwój wykorzystania nieklasycznych metod szacunku w praktyce badań statystycznych.

Metody badawcze i źródła danych

Zastosowane w pracy estymatory typu SPREE stanowią uogólnioną klasę estymatorów syntetycznych, wykorzystujących pełną informację o ocenach estymatora bezpośredniego. W ich budowie uwzględnia się znaną technikę przeważania danych w tabelach kontyngencji: tzw. metodę iteracyjnego proporcjonalnego dopasowania (ang. *Iterative Proportional Fitting* – IPF). Zastosowanie tej metody wymaga korekty liczebności znajdujących się w komórkach wielowymiarowej tabeli kontyngencji w taki sposób, aby skorygowane wartości sumowały się do znanych liczebności brzegowych. Wyjściowe liczebności w poszczególnych komórkach tabeli kontyngencji mogą pochodzić z różnych źródeł danych, na przykład ze spisu powszechnego, czy z rejestru administracyjnego. Natomiast liczebności brzegowe to szacunki uzyskane za pomocą estymatora bezpośredniego na podstawie danych pochodzących najczęściej z badania reprezentacyjnego. Purcell i Kish (1980) byli jednymi z pierwszych badaczy, którzy zaproponowali wykorzystanie estymatora zachowującego strukturę (SPREE) do uaktualniania danych w tabeli kontyngencji w oparciu o dane ze spisu powszechnego i aktualnego rejestru przy pomocy metody IPF. Zaproponowane przez nich podejście nawiązywało do pracy Deminga i Stephana (1940). Kluczową pracą w rozwoju metodologii SPREE jest także publikacja Zhanga i Chambersa (2004), w której autorzy zdefiniowali klasę uogólnionych liniowych modeli strukturalnych (GLSM) i rozszerzyli rozpatrywane podejście do modeli mieszanych z efektami losowymi o rozkładach normalnych (GLSMM).

W niniejszej rozprawie zastosowane zostały różne warianty estymatorów typu SPREE bazujące na modelach GLSM oraz GLSMM w celu oszacowania najważniejszych charakterystyk rynku pracy na poziomie obszarów funkcjonalnych miast wojewódzkich. Dane wykorzystane w pracy pochodzą z Badania Aktywności Ekonomicznej Ludności i rejestru ZROPiB wyprowadzonego na podstawie źródeł administracyjnych (m.in. ZUS, KRUS, KEP oraz rejestrów bezrobocia), a także obejmują szacunki bieżących struktur ludno-

ści. W celu oceny przestrzennego zróżnicowania rynku pracy w ujęciu miejskich obszarów funkcjonalnych wykorzystano wybrane metody analizy taksonomicznej: Topsis oraz hierarchiczną metodę grupowania Warda.

Rozdział 1

Wprowadzenie do problematyki miejskich obszarów funkcjonalnych

Celem niniejszego rozdziału jest wprowadzenie do zagadnień związanych z miejskimi obszarami funkcjonalnymi. W pierwszej części przedstawiona zostanie geneza oraz obecne postrzeganie i definiowanie miejskich obszarów funkcjonalnych w świetle istniejącej literatury przedmiotu. Przedmiotem rozważań będzie również kategoryzacja miejskich obszarów funkcjonalnych z uwzględnieniem zagadnienia ich zarządzania. Następnie wskazane będzie ich rosnące znaczenie w kontekście ustaw i strategii krajowych, a także w odniesieniu do polityki prowadzonej przez Unię Europejską, w której wymiar terytorialny i wzmacnianie endogenicznych potencjałów wykraczających poza przyjęte granice administracyjne nabiera szczególnego znaczenia. Opisane zostaną także metody wyznaczania tych obszarów. Ostatnia część rozdziału przedstawia charakterystykę obecnego pokrycia informacyjnego, zwracając szczególną uwagę na zgłaszane przez odbiorców danych zapotrzebowanie na informację statystyczną w przekroju miejskich obszarów funkcjonalnych, w tym szczególnie w obszarze rynku pracy.

1.1. Istota miejskich obszarów funkcjonalnych

1.1.1. Geneza miejskich obszarów funkcjonalnych

Miasta oraz ich obszary funkcjonalne są nieodzownym elementem rozwoju gospodarczego, stając się w coraz większej mierze źródłem rozwoju współczesnego świata (Morawiecki, 2017). Obecnie miasto, określone jedynie granicami administracyjnymi, przestaje być jednostką samodzielną i samowystarczalną pod względem zasobów ludzkich i dziennego rytmu życia. Terminologia pojawiająca się w literaturze przedmiotu, dotycząca terytorialnej identyfikacji stref lub obszarów oddziaływania miast jest bardzo szeroka i składają się na nią takie pojęcia jak: *aglomeracja*, *obszar metropolitalny*, *region metropolitalny*, *funkcjonalny obszar miejski*, *funkcjonalny region miejski*, *pole miejskie*, *strefa podmiejska*, *strefa żywicielska*, *residuum*. Jak wskazuje Śleszyński (2013a), przytoczone określenia nie wyczerpują wszystkich pojęć związanych z tym zagadnieniem. W światowej literaturze od lat toczy się bowiem dyskusja na temat relacji zachodzących pomiędzy miastem a jego otoczeniem. W tym kontekście warto tu przytoczyć chociażby prace takich autorów jak McKenzie (1933), Friedmann i Miller (1965), Berry (1967), Huff (1973), Hay i Hall (1977). Również w polskiej literaturze można znaleźć prace, które odnoszą się do tego nurtu badań. Wśród nich należy wskazać na opracowania Korcelliego (1981), Dziewońskiego (1987), Potrykowskiej (1989), Zuzańskiej-Zysko (2012) czy Korcelliego-Olejniczak (2012). Obszerne zestawienie głównych koncepcji z tego obszaru badań prowadzonych przez polskich naukowców znajduje się także w opracowaniu Śleszyńskiego (2015).

Jak można zauważyć stosowana w obszarze przestrzennej identyfikacji jednostek terminologia jest dość zróżnicowana. Obecnie miejskie obszary funkcjonalne najczęściej utożsamiane są z obszarami metropolitalnymi lub aglomeracjami¹ (Batóg i in., 2015) i w tym ujęciu są one przedmiotem analizy w niniejszej rozprawie.

Warto podkreślić, że już samo rozróżnienie wyżej wskazanych terminów jest dość istotną kwestią i sprawia trudności nie tylko naukowcom, ale też urzędnikom, planistom, urbani-

¹Należy jednak również zaznaczyć, iż szczególnie w ostatnim czasie pojawiły się opracowania podkreślające znaczenie miast średnich i ich obszarów funkcjonalnych (Śleszyński, 2018; Wałachowski i Król, 2019).

stom, czy też politykom (Kaczmarek i Bul, 2014). Próbę określenia definicji pojęć można odnaleźć chociażby w opracowaniu Markowskiego i Marszała (2006).

Szczególnie istotną kwestią w tym zakresie wydaje się być określenie różnic w pojmowaniu tych obszarów. Churski (2018b), bazując na dokonanym bogatym przeglądzie literatury, podkreślił, że „metropolia i jej obszar metropolitalny nie jest synonimem aglomeracji”, a cechami, które wyróżniają metropolię w stosunku do aglomeracji są przede wszystkim:

- 1) silniejsze relacje wewnętrzne, które nie ograniczają się do bezpośredniego otoczenia miasta głównego, ale dotyczą całego obszaru metropolitalnego, budując jego całościowy potencjał w relacjach z otoczeniem;
- 2) bardzo dobrze wykształcony podsystem dziennych powiązań mieszkaniowych, pracy, wypoczynku i współżycia zbiorowego;
- 3) dynamiczny rozwój wielofunkcyjnych relacji zewnętrznych, w tym o zasięgu ponadkrajowym.

Biorąc pod uwagę powyższe wskazania, szczególnie w kontekście ostatniego kryterium, trudno nie zgodzić się ze stwierdzeniem Churskiego (2018b), że „w dyskusji o roli dużych miast w procesach rozwoju w Polsce mówimy bardziej o aglomeracjach i obszarach aglomeracyjnych niż o metropoliach i obszarach metropolitalnych”, a na miano metropolii w polskich realiach zasługuje obecnie jedynie Warszawa. Do tego grona jednak chcą być zaliczane pozostałe miasta wojewódzkie, niezależnie od tego, czy wydaje się to uzasadnione, czy też nie. Dotyczy to szczególnie miast wojewódzkich położonych we wschodniej części Polski, które wspólnie podejmowały różne działania w kierunku postrzegania ich jako metropolie (Kuć-Czajkowska, 2019).

1.1.2. Miejski obszar funkcjonalny w literaturze przedmiotu

W odniesieniu do pojęć i uwag poczynionych w poprzednim podrozdziale, wydaje się zasadne, że ujmując najbardziej ogólnie, miejski obszar funkcjonalny możemy zdefiniować za Kaczmarkiem (2015) jako „obszar zajmowany przez duże miasto i otaczające go obszary, dla których jest głównym rynkiem pracy i źródłem specjalistycznych usług”.

Nieco odmiennie definiuje to pojęcie Śleszyński (2013a), który dokonując na zlecenie Ministerstwa Rozwoju Regionalnego delimitacji miast wojewódzkich, przyjął definicję miej-

skiego obszaru funkcjonalnego jako spójnej pod względem przestrzennym strefy oddziaływania miasta. Definicja ta obejmuje nie tylko istniejące powiązania funkcjonalne, bazujące niemal wyłącznie na dojazdach do pracy, ale bierze również pod uwagę proces zaawansowania procesów urbanizacyjnych. Miejski obszar funkcjonalny pojmowany jest zatem nie tylko jako strefa oddziaływania, w sensie dziennego systemu miejskiego, ale także jako ukształtowany w procesie historycznym zespół jednostek terytorialnych. Zespół ten wyróżnia się z otoczenia i upodabnia się pod pewnymi względami do głównego miasta określanego jako rdzeń. Z kolei Kuc-Czajkowska (2019) w swojej monografii definiuje (miejski) obszar funkcjonalny jako wyraźnie powiązany z miastem wojewódzkim (kilkoma miastami w przypadku ośrodków policentrycznych) obszar zurbanizowany, który z perspektywy struktury terytorialno-organizacyjnej państwa obejmuje gminy miejskie, miejsko-wiejskie i wiejskie położone w pierwszym, drugim, ewentualnie trzecim paśmie (pierścieniu).

Przywołując różne definicje obszaru funkcjonalnego warto też zwrócić uwagę na oddziaływanie miasta rdzeniowego na jego otoczenie. Według Noworóla (2014), wiąże się to ze zjawiskiem, które autor określa jako „funkcjonalne zajmowanie nowych terenów”. Abstrahuje ono od powiększania granic administracyjnych miasta. Związane jest zaś z kreowaniem formalnych i nieformalnych związków gospodarczych oraz społecznych, które wymagają coraz więcej przestrzeni. Ekspansja na nowe tereny w celu rozwoju działalności gospodarczej, czy też w związku z powstawaniem terenów mieszkaniowych powiększa obszar oddziaływania miasta rdzeniowego, kształtując miejski obszar funkcjonalny.

Warto też podkreślić, że miejskie obszary funkcjonalne, a szczególnie metropolitalne, tworzą złożone struktury w wymiarze administracyjnym, osadniczym, a także społeczno-gospodarczym. Istotę tych obszarów wyrażają takie cechy jak (Kaczmarek, 2017):

- 1) wielkomiejski układ osadniczy składający się z dużego miasta centralnego (lub kilku miast centralnych), strefy podmiejskiej oraz powiązanej z nimi strefy funkcjonalnej (najczęściej dojazdów do pracy);
- 2) wielość jednostek administracyjnych szczebla lokalnego, na które składają się gminy wiejskie i miejskie oraz często także jednostki szczebla subregionalnego;
- 3) powiązania przestrzenno-funkcjonalne, wyrażające się przepływem dóbr, osób, kapitału i informacji, skierowane nie tylko dośrodkowo, ale i odśrodkowo.

Postrzegając obszary funkcjonalne przez pryzmat polityki miejskiej należy podkreślić, że powinny one obejmować obszary, dla których należy prowadzić interwencję na rzecz rozwoju i zagospodarowania przestrzennego. W tym kontekście można wyznaczyć następujące typy miejskich obszarów funkcjonalnych (Kuźnik, 2015):

- obszar funkcjonalny miasta rozumiany jako miasto i jego strefa podmiejska, w której przebiegają procesy suburbanizacji oraz integracji i\lub dezintegracji w zbiorze zespołu jednostek osadniczych, w tym o profilu miejsko-wiejskim i wiejskim;
- aglomeracja miejska jako duży zespół miejski składający się z wielu różnej rangi miast i zorganizowany przestrzennie wokół miasta rangi regionalnej lub ponadregionalnej;
- aglomeracja policentryczna\konurbacja jako zespół miejski, w którym wielość miast o zbliżonej wielkości i randze wytwarza dodatkowe, w porównaniu z klasyczną aglomeracją, wyzwania rozwojowe i warunki współdziałania;
- obszar metropolitalny jako zespół miejski wyższej rangi funkcjonalnej, dający się uplasować w przestrzeni narodowej i międzynarodowej jako ważny element krajowej sieci osadniczej;
- obszar funkcjonalny miejsko-wiejski, gdzie duża gęstość zaludnienia i bogata sieć osadnicza małych i średnich miast oraz wiejskich jednostek osadniczych pozwala wyróżnić zespoły miejsko-wiejskie ważne dla identyfikacji i rozumienia procesów urbanizacji, w tym urbanizacji wsi.

Jak wynika z powyższego zestawienia, miejskie obszary funkcjonalne ewidentnie abstrahują od przyjętego podziału administracyjnego kraju, tworząc układy, które jako całość nie mają swojego odpowiednika w strukturze podziału administracyjnego Polski. Utrudnia to z pewnością zarządzanie nimi, a także diagnozowanie oraz monitorowanie zmian zachodzących w tych strukturach.

1.1.3. Zarządzanie miejskimi obszarami funkcjonalnymi

Obserwowany, szczególnie w ostatnich latach, dynamiczny rozwój relacji funkcjonalno-przestrzennych między miastem, a jego otoczeniem daje nie tylko podstawy do intensywnej współpracy między jednostkami komunalnymi, ale umożliwia również stosowanie zinstytucjonalizowanych form współzarządzania. Integracja ta jest obecnie coraz bardziej

kluczowym warunkiem rozwoju społeczno-gospodarczego, przede wszystkim w kontekście dynamicznych procesów suburbanizacji (Kaczmarek i Bul, 2014). W literaturze przedmiotu rozróżnia się dwa podejścia do procesu formalizowania współpracy w ramach miejskich obszarów funkcjonalnych. Pierwszym z nich jest podejście oddolne (ang. *bottom-up*) polegające na dobrowolnej współpracy poszczególnych samorządów w ramach tworzonego przez nie obszaru funkcjonalnego. Drugie to tzw. podejście odgórne (ang. *top-down*) polegające na narzuceniu z poziomu centralnego sposobu wyznaczenia i ukonstytuowania się miejskich obszarów funkcjonalnych zgodnie z obowiązującymi przepisami prawa (Kaczmarek, 2017; Kuć-Czajkowska, 2019).

Mając powyższe na uwadze, zmiany w tym obszarze można wprowadzać za pomocą odpowiednich ustaw na poziomie krajowym, bądź też w formie dobrowolnej współpracy jednostek samorządu terytorialnego. Wśród form dobrowolnej współpracy rozróżnić możemy następujące postaci (Kuć-Czajkowska, 2019):

- zinstytucjonalizowane (związki, stowarzyszenia),
- słabo zinstytucjonalizowane (porozumienia),
- niezinstytucjonalizowane (rady, fora, czyli wszelkiego rodzaju umowy typu *gentlemen agreements*).

1.2. Identyfikacja obszarów funkcjonalnych w świetle ustaw i strategii krajowych

1.2.1. Charakterystyka obecnie funkcjonującego podziału administracyjnego kraju

Właściwe zrozumienie i odpowiednie umiejscowienie miejskich obszarów funkcjonalnych należałoby rozpocząć od krótkiej genezy oraz analizy obecnie funkcjonującego podziału administracyjnego Polski. Warto również przyjrzeć się działaniom podejmowanym w celu nawiązywania współpracy pomiędzy jednostkami samorządowymi, które mogłyby wchodzić w skład miejskiego obszaru funkcjonalnego.

Podział terytorialny, zaliczany do kluczowych reform ustrojowych państwa, wdrożono w dwóch etapach. W pierwszym zmiana odbyła się w 1990 roku², po transformacji ustrojowej z 1989 roku. Umożliwiła wprowadzenie samorządowych gmin. Następnie w roku 1999 przeprowadzono reformę układu i liczby województw oraz wprowadzono dodatkowy pośredni szczebel podziału administracyjnego, jakim są powiaty. Tym samym Polska znalazła się w grupie krajów o trójszczeblowej strukturze podziału administracyjnego. W porównaniu do pozostałych państw europejskich, polskie gminy mają relatywnie dużą liczbę mieszkańców. Podczas gdy ludność województw jest zbliżona do średniej europejskiej dla porównywalnych jednostek administracyjnych i wyróżnia się na tym tle najmniejszymi pod względem liczby mieszkańców jednostkami szczebla pośredniego (Kachniarz i Babczuk, 2015). Od czasu ostatniej reformy przeprowadzane były jedynie niewielkie korekty granic jednostek samorządu terytorialnego. Powoływano również kilka nowych powiatów i gmin oraz likwidowano niektóre dotychczas funkcjonujące. Dokonywano również zmian w charakterze gmin przekształcając je głównie z gmin wiejskich na miejsko-wiejskie, z wydzielonym miastem i częścią wiejską gminy. Według stanu na 1 stycznia 2017 roku (rok referencyjny dla analiz przeprowadzonych w niniejszej pracy) Polska podzielona była na 16 województw, 380 powiatów i 2478 gmin.

Jednak, jak wskazano w raporcie *Samorząd terytorialny w Polsce* (Wojciechowski i in., 2014), decyzje polityczne miały duży wpływ na wyznaczanie granic województw oraz powiatów, które nie często układają się w spójny funkcjonalnie obszar w wymiarze społecznym, czy gospodarczym. Podkreślono również, między innymi, że wraz z przekazywaniem zadań jednostkom samorządu terytorialnego, nie zawsze są zabezpieczone adekwatne środki finansowe. Konsekwencją tego są znaczne ograniczenia finansów samorządów przy jednoczesnym zwiększeniu zakresu zadań lub delegowaniu nowych zadań bez odpowiednich środków. Jednocześnie zaznaczono, że w przeciwieństwie do innych krajów europejskich (np. Finlandii, Portugalii, Włoch), w Polsce brakuje polityki krajowej dotyczącej stymulowania współpracy międzysamorządowej. Jak wskazuje Kuć-Czajkowska (2019), godnym rozważenia rozwiązaniem w tym zakresie mogłoby być przygotowanie odpowiednich zachęt finansowych dla tak zdefiniowanego obszaru, jak na przykład ma to miejsce w Niemczech.

²Dnia 8 marca 1990 roku przyjęto ustawę o samorządzie terytorialnym (Dz.U.1990, nr 16, poz. 95).

Należy przypuszczać, że nawiązana forma współpracy pomiędzy jednostkami samorządu terytorialnego mogłaby ułatwić w szczególności realizację zadań, które swoim zakresem wykraczają poza granice administracyjne poszczególnych samorządów lokalnych. Obecnie pewne rozwiązania w tym obszarze proponowane są w ramach instrumentu ZIT³. Natomiast nadal w wielu strategiach rozwoju uchwalanych przez polskie samorzady nie dostrzega się potrzeby oraz możliwości współpracy kilku jednostek, aby sprostać wspólnym problemom. Dość często brakuje dokładnej analizy możliwości współpracy w ramach okolicznych jednostek samorządowych. Jest to szczególnie istotne dla prawidłowego funkcjonowania i rozwoju obszarów funkcjonalnych. Ponadto, bardzo istotną funkcją wieloletnich programów inwestycyjnych powinno być zbadanie możliwości realizacji poszczególnych przedsięwzięć wspólnie z innymi podmiotami, przede wszystkim z jednostkami samorządu terytorialnego, które tworzą jeden obszar funkcjonalny (Sierak i Bitner, 2016). Również w innych opracowaniach rekomenduje się, aby przynajmniej miasta na prawach powiatu miały przygotowane strategie oparte na studium uwarunkowań i kierunków zagospodarowania przestrzennego w ramach miejskiego obszaru funkcjonalnego. Oznacza to planowanie i integrowanie działań wykraczających poza granice poszczególnych gmin (Noworól, 2016).

1.2.2. Uregulowania ustawowe dotyczące obszarów metropolitalnych oraz miejskich obszarów funkcjonalnych

Wydaje się, że głównym problemem powszechnej współpracy pomiędzy poszczególnymi gminami jest to, że niezależnie od ich wielkości, posiadają te same kompetencje, co czyni je sobie równoważnymi. Natomiast występują dość często rozbieżności interesów pomiędzy poszczególnymi gminami okalającymi miasto oraz pomiędzy nimi a miastem. Taka sytuacja nie skłania ich do zacieśniania współpracy.

Po 1989 roku starano się w różny sposób przezwyciężyć sztywne uregulowania podziału administracyjnego i podejmowano inicjatywy porządkujące współpracę pomiędzy jednostkami samorządu terytorialnego tworzącego aglomeracje miejskie. Na przestrzeni wielu lat

³ZIT - Zintegrowane Inwestycje Terytorialne. Więcej informacji na temat tego instrumentu znajduje się w sekcji 1.3.2.

pojawiały się propozycje, projekty oraz ustawy dotyczące obszarów metropolitalnych⁴. Warto zwrócić uwagę na następujące projekty i ustawy, które pojawiły się w ostatnich latach:

- ustawa z 9 października 2015 roku o związkach metropolitalnych,
- projekt ustawy z 2016 roku o Poznańskim Związku Metropolitalnym,
- projekt ustawy z 2017 roku o ustroju miasta stołecznego Warszawy,
- ustawa z 9 marca 2017 roku o związku metropolitalnym w województwie śląskim.

Odnosząc się do powyższego zestawienia należy podkreślić, że ustawa o związkach metropolitalnych nie miała żadnej mocy sprawczej, gdyż nie wydano do niej rozporządzeń, które umożliwiłyby jej wdrożenie, przez co ostatecznie utraciła moc prawną. Mocj prawnej, jak do tej pory, nabrała jedynie ustawa o związku metropolitalnym w województwie śląskim. Jest to jedyny jak dotąd obszar metropolitalny utworzony na mocy ustawy. Inaczej potoczyły się losy projektu ustawy o Poznańskim Związku Metropolitalnym. Został on ostatecznie odrzucony po pierwszym czytaniu w sejmie w dniu 9 marca 2017 roku, a decyzję tę uzasadniano, między innymi, powielaniem rozwiązań ustawy z 2015 roku o związku metropolitalnym. Z kolei projekt ustawy o ustroju miasta stołecznego Warszawy został wycofany z obrad sejmiku, między innymi, z powodu licznych protestów, a także niekorzystnych wyników referendum przeprowadzanych w kilku gminach wchodzących w jej skład (Kuć-Czajkowska, 2019).

Powyższe przykłady świadczą o tym, że odgórne uregulowania w formie ustawy dotyczące obszarów metropolitalnych w Polsce do tej pory nie znalazły uniwersalnego rozwiązania. Tylko w jednym przypadku indywidualna inicjatywa ustawodawcza zakończyła się powodzeniem.

1.2.3. Ranga miejskich obszarów funkcjonalnych w świetle krajowych dokumentów strategicznych

Pomimo braku systemowego podejścia i jednolitej formuły funkcjonowania miejskich obszarów funkcjonalnych, odwołania do nich znajdują się w wielu kluczowych dokumentach strategicznych Polski, między innymi takich jak:

⁴Obszerne informacje na ten temat można znaleźć w monografii Kuć-Czajkowskiej (2019).

- Koncepcja Przestrzennego Zagospodarowania Kraju 2030,
- Krajowa Polityka Miejska 2023,
- Strategia na rzecz Odpowiedzialnego Rozwoju do roku 2020 (z perspektywą do 2030 r.),
- Krajowa Strategia Rozwoju Regionalnego 2030.

Koncepcja Przestrzennego Zagospodarowania Kraju 2030 (KPZK 2030)

Jest to w zasadzie pierwszy dokument, który definiuje pojęcie miejskiego obszaru funkcjonalnego, a większość autorów zajmujących się tymi zagadnieniami uważa go za kluczowy (por. między innymi: (Szafranek, 2015; Kociuba, 2017b)).

W dokumencie KPZK 2030 (Rada Ministrów, 2012) miejski obszar funkcjonalny (MOF) rozumiany jest jako złożony z odrębnych jednostek administracyjnych ciągły przestrzennie układ osadniczy. Obejmuje on zwarty obszar miejski oraz powiązaną z nim funkcjonalnie strefę zurbanizowaną. Jak podkreśla Kociuba (2015), kluczowe znaczenie ma tutaj siła wewnętrznych powiązań pomiędzy miastem rdzeniowym a jego obszarem funkcjonalnym. Siła tych powiązań zależy szczególnie od: intensywności dojazdów do pracy, wielkości przepływu towarów i usług o zróżnicowanym charakterze, wzajemnych relacji pomiędzy rynkiem pracy a rynkiem mieszkaniowym, infrastruktury technicznej, a także struktury przyrodniczej. Zgodnie z KPZK 2030, jednostki te mogą składać się zarówno z gmin miejskich, wiejskich, jak i z miejsko-wiejskich. Ich typologia odzwierciedla rolę, jaką poszczególne grupy ośrodków miejskich odgrywają w systemie osadniczym kraju i w głównej mierze bazuje na wielkości tych ośrodków. W ramach tego dokumentu rozróżniono cztery podtypy miejskich obszarów funkcjonalnych:

- **obszary funkcjonalne ośrodków wojewódzkich** – zaliczane są do nich obszary powyżej 300 tys. mieszkańców. W ramach tego podziału część z nich, ze względu na pełnioną rolę w systemie osadniczym kraju oraz ich znaczenie społeczno-gospodarcze w procesach rozwojowych, uzyskuje status ośrodków metropolitalnych (Warszawa, Aglomeracja Górnośląska (główny ośrodek Katowice), Kraków, Łódź, Trójmiasto, Poznań, Wrocław, Bydgoszcz z Toruniem, Szczecin oraz Lublin). Natomiast pozostałe obszary funkcjonalne ośrodków wojewódzkich są zaliczane do ośrodków o znaczeniu krajowym (Opole, Zielona Góra i Gorzów Wielkopolski, Kielce, Olsztyn). Dla wszystkich miast wojewódzkich istnieje obowiązek delimitacji obszarów funk-

cjonalnych oraz przygotowania planu zagospodarowania, a także strategii rozwoju dla całego swojego obszaru funkcjonalnego.

- **obszary funkcjonalne ośrodków regionalnych** – liczba mieszkańców zawiera się w przedziale 100–300 tys. Pełnią one kluczową rolę w życiu gospodarczym, społecznym i administracyjnym poszczególnych regionów (Częstochowa, Radom, Bielsko-Biała, Rybnik, Płock, Elbląg, Wałbrzych, Włocławek, Tarnów, Kalisz z Ostrowem Wielkopolskim, Koszalin, Legnica, Grudziądz, Słupsk). Wyznaczenie obszarów funkcjonalnych ośrodków regionalnych nie obliguje do tworzenia odrębnych planów zagospodarowania ani strategii rozwoju.
- **obszary funkcjonalne ośrodków subregionalnych** – są wyznaczane przez samorząd wojewódzki wokół miast liczących 50–100 tys. mieszkańców. W skład tej grupy wchodzi miasta wraz z ich obszarami funkcjonalnymi pełniące rolę ośrodków koncentrujących funkcje społeczno-gospodarcze oraz będące miejscem dostarczania usług publicznych ważnych dla ich mieszkańców. Obszary funkcjonalne ośrodków subregionalnych mogą być wyznaczane w planach zagospodarowania przestrzennego województw. Mogą też być tworzone, na zasadzie dowolności, plany zagospodarowania takich obszarów.
- **ośrodki lokalne** – miasta liczące poniżej 50 tys. mieszkańców. Miasta te nie mają swojego zaplecza funkcjonalnego w postaci wyznaczonego obszaru funkcjonalnego. Mają one jednak szansę na koncentrację funkcji gospodarczych oraz służą obszarom wiejskim swoim zapleczem usługowym. Szczególną rolę w ramach tego podziału odgrywają niektóre miasta powiatowe.

Krajowa Polityka Miejska 2023 (KPM 2023)

Bezpośrednią podstawą opracowania Krajowej Polityki Miejskiej 2023 był dokument Założenia Krajowej Polityki Miejskiej do roku 2020⁵. KPM 2023 określa planowane działania administracji rządowej dotyczące polityki miejskiej. Zgodnie z ustawą z dnia 6 grudnia 2006 r. o zasadach prowadzenia polityki rozwoju, uwzględnia ona cele i kierunki ujęte w średniookresowej strategii rozwoju kraju oraz krajowej strategii rozwoju regionalnego. Polityka miejska jest kierowana do wszystkich polskich miast wraz z ich obszarami funkcjonalnymi. Pojmowanie obszarów funkcjonalnych miast jest tożsame z ich definicją zawartą

⁵Przyjęty przez Radę Ministrów w dniu 16 lipca 2013 r.

w KPZK 2030. Jest to wyrazem zintegrowanego podejścia terytorialnego⁶ w wymiarze miejskim. Ukierunkowanie polityki miejskiej zarówno w stronę miast, ale także w stronę ich obszarów funkcjonalnych, wynika z dostrzeżenia siły powiązań pomiędzy nimi. Polityka miejska realizując zintegrowane podejście terytorialne odnosi się zatem do obszarów funkcjonalnych, a nie terytoriów wyznaczonych przez granice administracyjne. Daje to możliwość lepszego wykorzystania ukrytych w nich lub niewłaściwie dotychczas wykorzystanych potencjałów i zasobów. Dlatego też polityka miejska oddziałuje nie tylko na gminy miejskie oraz miasta na prawach powiatu, ale również na gminy miejsko-wiejskie i wiejskie, które mogą stanowić część obszarów funkcjonalnych (MIiR, 2015a). Jak nadmieniał Szlachta (2017), strategia kieruje swoją uwagę na potrzebę istotnej zmiany jakości zarządzania obszarami miejskimi, poprzez stymulowanie współpracy pomiędzy jednostkami samorządu terytorialnego, gdyż konkurencja pomiędzy sąsiadującymi gminami o inwestorów, mieszkańców czy też dochody wpływa niekorzystnie na działanie całego obszaru funkcjonalnego.

Strategia na rzecz Odpowiedzialnego Rozwoju do roku 2020 (z perspektywą do 2030 r.)

Strategia na rzecz Odpowiedzialnego Rozwoju (SOR) jest aktualizacją średniookresowej strategii rozwoju kraju – Strategii Rozwoju Kraju 2020,⁷. SOR jest strategicznym instrumentem zarządzania polityką rozwoju realizowaną przez instytucje państwa. Przedstawia ona cele do realizacji w horyzoncie czasowym do 2020 oraz 2030 roku. Celem strategii w obszarze polityki miejskiej jest tworzenie warunków dla zrównoważonego rozwoju ośrodków miejskich oraz wypracowanie mechanizmów współpracy w ich obszarach funkcjonalnych. Wymiar miejski strategii koncentruje się na wspieraniu rozwoju wszystkich miast, niezależnie od ich wielkości, położenia czy też pełnionych funkcji społecznych, gospodarczych i administracyjnych. W tym kontekście szczególny nacisk poświęcony został trzem obszarom oddziaływania, w tym największym aglomeracjom, miastom średnim tracącym funkcje społeczno-gospodarcze oraz obszarom zagrożonym trwałą marginalizacją. Zwrócono przy tym uwagę na pełniejsze wykorzystanie potencjału największych polskich

⁶Więcej informacji na temat idei zintegrowanego podejścia terytorialnego na poziomie europejskim znajduje się w podrozdziale 1.3.

⁷Strategia Rozwoju Kraju 2020, przyjęta uchwałą Rady Ministrów z dnia 25 września 2012 r., zgodnie z wymogami ustawy z dnia 6 grudnia 2006 r. o zasadach prowadzenia polityki rozwoju (Dz. U. z 2016 r. poz. 383, 1250, 1948 i 1954 oraz z 2017 r. poz. 5). Strategia uwzględnia główne elementy struktury wymagane przez tę ustawę dla tego typu dokumentów (art. 9 i art. 12a.).

aglomeracji. Wskazano, że polityka miejska koncentrować się ma na wspieraniu współpracy i koordynacji działań prorozwojowych w obszarach funkcjonalnych aglomeracji. Zaznaczono również, że w ich skład wchodzi zarówno administracyjne obszary miejskie, jak i wiejskie. Podkreślono także spójność działań przewidzianych do realizacji z priorytetami wskazanymi w Krajowej Polityce Miejskiej 2023 (Morawiecki, 2017).

Krajowa Strategia Rozwoju Regionalnego 2030 (KSRR 2030)

Krajowa Strategia Rozwoju Regionalnego 2030 (MIiR, 2019) rozwija postanowienia Strategii na rzecz Odpowiedzialnego Rozwoju do roku 2020 (z perspektywą do 2030 roku). KSRR jest podstawowym dokumentem strategicznym polityki regionalnej państwa w perspektywie do 2030 roku. Zastąpiła ona wcześniej obowiązującą Krajową Strategię Rozwoju Regionalnego 2010–2020 (MRR, 2010). Obecnie obowiązujący dokument podtrzymuje rangę i znaczenie miejskich obszarów funkcjonalnych w polityce rozwoju. Podkreśla nadal znaczenie obszarów funkcjonalnych miast wojewódzkich, kładąc jednak również nacisk na rozwój obszarów funkcjonalnych średnich miast tracących funkcje społecznie i gospodarcze. KSRR 2030 wskazuje, że miasta duże i średnie są głównymi centrami aktywności społecznej i gospodarczej, a ich pozycja wpływa bezpośrednio na obszary powiązane z nimi funkcjonalnie. Podkreśla się również, że pożądana jest współpraca miast w obszarach funkcjonalnych w ramach wspólnych projektów dotyczących m.in. transportu publicznego, ochrony środowiska, czy też gospodarki odpadami oraz wykorzystania nowych technologii. W dokumencie zwraca się uwagę na potrzebę uporządkowania pojęć dotyczących obszarów funkcjonalnych, obszarów strategicznej interwencji, a także obszarów problemowych m.in. na nieuregulowane zależności między nimi. Więcej informacji na temat problemów z nakładaniem się tych definicji i trudności w ich klasyfikacji, a także propozycji ich uporządkowania można znaleźć w opracowaniu Śleszyńskiego i in. (2017).

1.3. Wymiar terytorialny na poziomie Unii Europejskiej

1.3.1. Znaczenie terytorium i obszarów miejskich wykraczających poza granice administracyjne

Zmiany widoczne w ustawach i dokumentach strategicznych w Polsce w podejściu do postrzegania terytorium, w tym zwrócenie uwagi na znaczenie miejskich obszarów funkcjonalnych, ma w dużej mierze swoje źródło w zmianach zachodzących na poziomie Unii Europejskiej.

Trwająca od kilkunastu lat ewolucja w kształtowaniu polityki rozwoju na poziomie Wspólnoty przyczyniła się do zmiany kierunków i podejścia do programowania interwencji publicznej (Churski, 2018a; Słupińska, 2013; Zaucha i in., 2015). Kluczowym elementem tych zmian jest dołączenie trzeciego wymiaru spójności tzw. spójności terytorialnej, która na mocy Traktatu Lizbońskiego została wprowadzona i jest traktowana jako równorzędna z wymiarem gospodarczym i społecznym (Kaczmarek i Bul, 2014). Dlatego też, jak zostało podkreślone w opracowaniu Śleszyńskiego (2017), jednym z narzędzi obecnej polityki spójności jest określenie obszarów, wobec których niezbędne jest prowadzenie specjalnej, a często zindywidualizowanej polityki rozwoju. Jest to odejście od wcześniejszego paradygmatu opartego na wyrównywaniu różnic w poziomie rozwoju regionów oraz na podejściu sektorowym do podejmowanych działań. Obecnie wcześniejsze rozwiązanie zastępowane jest podejściem zintegrowanym terytorialnie (ang. *place-based approach*), które uwzględnia różnorodność uwarunkowań poszczególnych obszarów (MliR, 2015b). W tym nowym paradygmacie, terytorium do którego kierowana jest interwencja, nie jest postrzegane przez pryzmat wyznaczonych granic administracyjnych. Kluczowym czynnikiem w tej kwestii jest identyfikacja wspólnych potencjałów oraz wyzwań, jakie stoją przed tak określonym obszarem (Kaczmarek i Bul, 2014).

Jak wskazuje wielu autorów zajmujących się tą dziedziną, m.in. Churski (2018a), Kociuba (2017b), podejście to zostało szczegółowo opisane w raporcie Barcy (2009). Podkreślono w nim przede wszystkim potrzebę wykorzystania czynników endogenicznych (wewnętrznych zasobów) danego terytorium, nie zapominając jednak o czynnikach egzogenicznych,

które chociażby w zakresie innowacji wprowadzanej z „zewnątrz” mają istotne znaczenie dla rozwoju określonego terytorium. Prowadzenie skutecznej polityki rozwoju polega zatem na odpowiednim, łącznym wykorzystaniu czynników wewnętrznych oraz zewnętrznych. Jak podkreśla Słodowa-Hełpa (2014), kluczowe znaczenie ma również budowanie odpowiednich systemów zarządzania oraz sieci współpracy pomiędzy różnymi jednostkami samorządu i ich władzami, ale również partnerami społecznymi, przedsiębiorstwami, czy też bankami oraz podmiotami społecznymi i gospodarczymi społeczeństwa obywatelskiego.

Na potencjał związany ze współpracą jako wartością dodaną wskazuje również Agenda Terytorialna Unii Europejskiej 2020 (European Council, 2011), w której mówi się o stosowaniu zintegrowanego podejścia do rozwoju miast. Natomiast już w Karcie Lipskiej (European Council, 2007) podjęto tematykę zrównoważonego rozwoju miast europejskich. Wskazywano przede wszystkim na ich kluczowe znaczenie jako głównych ośrodków rozwoju regionów miejskich. Zwrócono w niej również uwagę na konieczność opracowania strategii i zharmonizowania działań przekraczających granice administracyjne miast, włączając w ten proces ich obszary funkcjonalne (Heffner i Klemens, 2016). Takie podejście ma również swoje odniesienie w Strategii Europa 2020, która podkreśla znaczenie konkurencyjności dla rozwoju społeczno-gospodarczego. Szczególną rolę mają w tym względzie odgrywają miasta duże z ich obszarami funkcjonalnymi, które przyczyniają się do wzrostu efektywności, konkurencyjności, a także innowacyjności w różnorodnych wymiarach przestrzennych (Szlachta, 2017).

Powyższe założenia mają swoje odzwierciedlenie w kształtowaniu polityki Unii Europejskiej na lata 2014–2020. Wzmocnienie wymiaru miejskiego polityki spójności i większe włączenie miast w jej programowanie realizowane jest poprzez kilka mechanizmów takich jak: zrównoważony rozwój obszarów miejskich, innowacyjne działania na rzecz zrównoważonego rozwoju obszarów miejskich, sieć na rzecz rozwoju obszarów miejskich czy też rozwój lokalny kierowany przez społeczność (MIiR, 2015b).

1.3.2. Zintegrowane Inwestycje Terytorialne

Wpisując się w nowy paradygmat w ramach polityki spójności Wspólnoty Europejskiej w perspektywie finansowej Unii Europejskiej na lata 2014–2020, Zintegrowane Inwesty-

cje Terytorialne (ZIT)⁸ były jednym z instrumentów interwencji, który powstał, przede wszystkim, z myślą o miastach i ich obszarach funkcjonalnych⁹. Na wykorzystanie tej formy wsparcia w ramach mechanizmu zrównoważonego rozwoju obszarów miejskich zdecydowało się czternaście państw UE. Poza tą formą, zrównoważony rozwój obszarów miejskich może być realizowany poprzez odrębną wielotematyczną oś priorytetową, dedykowany program operacyjny bądź też inne programy operacyjne. Wprowadzenie tego mechanizmu jest obligatoryjne, a każde z państw członkowskich zobowiązane jest do przekazania co najmniej 5% środków pochodzących z Europejskiego Funduszu Rozwoju Regionalnego na jego realizację (MIiR, 2015b). Niektóre kraje zastosowały dodatkowo inne instrumenty. Czechy, poza ZIT-ami, skorzystały dodatkowo z innej formy wsparcia w postaci Zintegrowanego Planu Zagospodarowania Terenu (ang. *Integrated Area Development Plan*)¹⁰. Główną ideą przyświecającą realizacji ZIT-ów w perspektywie 2014–2020 było zapewnienie długoterminowej współpracy samorządów lokalnych położonych na danym obszarze funkcjonalnym. Było to realizowane poprzez wspólne projekty, mające na celu podejmowanie działań ułatwiających rozwiązywanie w głównej mierze społeczno-ekonomicznych, demograficznych czy też środowiskowych problemów (Truskołaski i in., 2017). Polska należała do państw, które jedynie poprzez Zintegrowane Inwestycje Terytorialne realizują zrównoważony rozwój obszarów miejskich. Uregulowania prawne i wytyczne dotyczące ZIT-ów na poziomie naszego kraju zostały zapisane w ustawie¹¹ oraz kilku dodatkowych dokumentach¹². Określone są w nich między innymi szczegółowe warunki realizacji ZIT-ów w perspektywie finansowej 2014–2020 na terenie naszego kraju. Wdrożeniu tego instru-

⁸Pojęcie ZIT w niniejszej rozprawie doktorskiej utożsamiane będzie z instrumentem polityki Unii Europejskiej realizowanym w perspektywie finansowej 2014–2020.

⁹Podstawy prawne dotyczące tego instrumentu znajdują się w rozporządzeniach Parlamentu Europejskiego: Rozporządzenie Parlamentu Europejskiego i Rady UE nr 1303/2013 z 17 grudnia 2013 r. (przede wszystkim artykuł 36), Rozporządzenie Parlamentu Europejskiego i Rady UE nr 1301/2013 z 17 grudnia 2013 r. (przede wszystkim artykuł 7), Rozporządzenie Parlamentu Europejskiego i Rady UE nr 1304/2013 z 17 grudnia 2013 r. (artykuł 12). Komisja Europejska określiła dodatkowe wytyczne w tym zakresie: Wytyczne Komisji Europejskiej dla Państw Członkowskich dotyczące Zrównoważonego Rozwoju Obszarów Miejskich z dnia 18 maja 2015.

¹⁰Więcej informacji na temat implementacji ZIT-ów przez poszczególne kraje członkowskie można znaleźć w opracowaniu Binka i in. (2016).

¹¹Ustawa z 11 lipca 2014 r. o realizacji programów w zakresie polityki spójności finansowanych w perspektywie finansowej 2014–2020 (artykuł 30).

¹²Umowa Partnerstwa (21 maja 2014 r.) – ZIT-om poświęcona jest część rozdziału „Instrumenty rozwoju terytorialnego” (punkt 3.1.2); Zasady realizacji zintegrowanych inwestycji terytorialnych w Polsce – wytyczne Ministerstwa Rozwoju Regionalnego wydane w lipcu 2013 r. stanowiły podstawowy dokument dotyczący ZIT-ów w Polsce przed przyjęciem Umowy Partnerstwa oraz ustawy z 11 lipca 2014 r. o realizacji programów w zakresie polityki spójności finansowanych w perspektywie finansowej 2014–2020; Wytyczne Ministra Infrastruktury i Rozwoju w zakresie trybów wyboru projektów na lata 2014–2020 z 31 marca 2015 r.

mentu w Polsce przyświecały trzy kluczowe cele wynikające z Umowy Partnerstwa (MIiR, 2017):

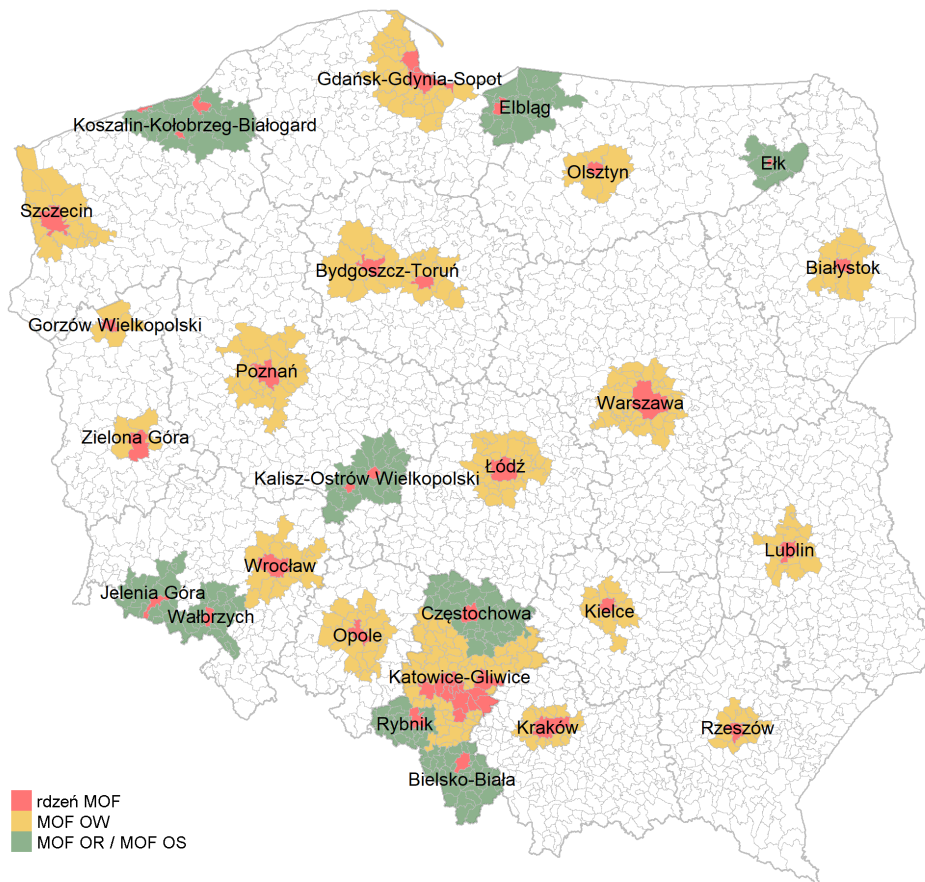
- promowanie partnerskiego modelu współpracy różnych jednostek administracyjnych na miejskich obszarach funkcjonalnych,
- zwiększenie efektywności podejmowanych interwencji poprzez realizację zintegrowanych projektów odpowiadających w sposób kompleksowy na potrzeby i problemy miast i obszarów powiązanych z nimi funkcjonalnie,
- zwiększenie wpływu miast i obszarów powiązanych z nimi funkcjonalnie na kształt i sposób realizacji działań wspieranych na ich obszarze w ramach polityki spójności.

Uruchomienie instrumentu ZIT związane było z przygotowaniem strategii dla poszczególnych obszarów funkcjonalnych objętych tym instrumentem, a także sformalizowaniem współpracy i ustanowieniem odpowiedniej jej formy w ramach obszaru objętego ZIT, w którego skład wchodzi jednostki samorządu terytorialnego. Natomiast działania wymienione w strategii realizowane były poprzez projekty finansowane z kilku osi priorytetowych i programów operacyjnych (Kociuba, 2018b; Szafranek, 2017b). Związek ZIT mógł być realizowany w formie porozumienia samorządów, stowarzyszenia czy też związku międzygminnego. W tym zakresie przeważają dwa modele zarządzania terytorialnego, tzw. model „doraźny”, w którym to specjalnie utworzono obszary ZIT w celu wydatkowania środków przeznaczonych na ten instrument oraz model „współpracy”, który opierał się na wcześniej podjętej współpracy przez jednostki terytorialne tworzące ZIT (Kociuba, 2017a). Instrument ZIT był realizowany obligatoryjnie w 17 miejskich obszarach funkcjonalnych ośrodków wojewódzkich¹³ (MOF OW), w których skład wchodzi 348 gmin¹⁴. Dodatkowo, na podstawie decyzji władz wojewódzkich, ZIT-y realizowane były również w 9 miejskich obszarach funkcjonalnych ośrodków regionalnych (MOF OR) i subregionalnych (MOF OS) w 5 województwach obejmując łącznie 181 gmin (Kociuba, 2018b). Rozmieszczenie poszczególnych obszarów wsparcia instrumentem ZIT znajduje się na rysunku 1.1.

Jak pisze Broniszewski (2017), zintegrowane za pomocą tego instrumentu poszczególne miejskie obszary funkcjonalne mają możliwość pobudzenia i wykorzystania swoich endo-

¹³Miasta Toruń oraz Bydgoszcz wraz z pozostałymi wybranymi jednostkami samorządu terytorialnego tworzą jeden wspólny miejski obszar funkcjonalny.

¹⁴Informacja na temat sposobu wyznaczenia miejskich obszarów funkcjonalnych w kontekście instrumentu ZIT znajduje się w sekcji 1.4.



Rysunek 1.1: Miejskie obszary funkcjonalne objęte instrumentem ZIT w Polsce w perspektywie finansowej Unii Europejskiej 2014–2020

Objaśnienia: Na wykresie mapowym kolorem czerwonym oznaczono nazwy miast tworzących rdzenie poszczególnych miejskich obszarów funkcjonalnych (za wyjątkiem MOF OW Subregionu Centralnego, gdzie oprócz Katowic i Gliwic rdzeń tworzy jeszcze kolejne 21 miast). Kolorem żółtym oznaczone zostały pozostałe gminy wchodzące w skład 17 miejskich obszarów funkcjonalnych ośrodków wojewódzkich. Z kolei kolorem zielonym oznaczono pozostałe gminy wchodzące w skład 14 miejskich obszarów funkcjonalnych ośrodków regionalnych i subregionalnych. Nazwy obszarów funkcjonalnych oznaczonych na rysunku 1.1 znajdują się aneksie w tabeli A.1

Źródło: Opracowanie własne na podstawie Kociuby (2018b).

genicznych potencjałów, które są charakterystyczne dla tak wyznaczonego terytorium. Ponadto różnią się od pozostałych obszarów. Jednocześnie, jak podkreśla Szafranek (2015), można zakładać, że miejskie obszary funkcjonalne poprzez wykorzystanie instrumentu ZIT staną się faktycznymi biegunami rozwoju regionów. Wydaje się jednak, że aby móc

to rzetelnie ocenić, potrzebna jest wiarygodna informacja statystyczna dotycząca tak wyznaczonego terytorium¹⁵.

Również w obecnej perspektywie finansowej Unii Europejskiej na lata 2021–2027 forma wsparcia w ramach tego instrumentu została podtrzymana. Planowane jest w niej zwiększenie liczby miejskich obszarów funkcjonalnych w ramach ZIT-ów, poprzez objęcie nimi również miast średnich.

1.3.3. Charakterystyka podziału statystycznego funkcjonującego w ramach Wspólnoty

Unia Europejska od dawna prowadzi działania, aby pozyskiwać odpowiednią informację statystyczną w wymiarze terytorialnym, która jest kluczowa chociażby w kontekście prowadzenia skutecznej polityki rozwoju. Funkcjonujący w ramach Wspólnoty Urząd Statystyczny Unii Europejskiej (Eurostat) przy współpracy z krajowymi urzędami statystycznymi państw członkowskich opracował zharmonizowany sposób podziału terytorialnego całego obszaru Wspólnoty. Jest on oparty na wspólnych założeniach metodologicznych, a jego celem jest między innymi uzyskanie porównywalnych danych statystycznych. Jak zaznacza Szlachta (2017), jest to istotne chociażby dla monitorowania oraz uzyskania wysokiej jakości informacji o procesach społeczno-gospodarczych zachodzących na poszczególnych terenach UE oraz ustaleniu obszarów problemowych, które później mogą stać się celem interwencji w ramach europejskiej polityki spójności. Podstawowym standardem geograficznym służącym do statystycznego podziału państw członkowskich Unii Europejskiej na trzy poziomy regionalne o określonej liczbie ludności jest Klasyfikacja Jednostek Terytorialnych do Celów Statystycznych – NUTS¹⁶. Podstawą tego podziału nie są jednak miasta i ich obszary funkcjonalne. Natomiast istnieją jeszcze inne unijne typologie terytorialne stosowane przez Eurostat. Ostatnia ich aktualizacja w ramach tzw.

¹⁵Więcej informacji na temat delimitacji miejskich obszarów funkcjonalnych oraz potrzeb informacyjnych dla tak określonych obszarów znajduje się w dalszej części pracy.

¹⁶Nazwa pochodzi z języka francuskiego: *Nomenclature des Unités territoriales statistiques*, po angielsku stosuje się określenie: *Classification of Territorial Units for Statistics*, a także *Nomenclature of Territorial Units for Statistics*.

inicjatywy „Tercet” odbyła się poprzez rozporządzenie Parlamentu Europejskiego i Rady Unii Europejskiej w 2017 roku¹⁷ i wprowadziła następujące typologie:

- oparta na siatce o poziomie rozdzielczości wynoszącym 1 km², przy pomocy której wyznacza się ośrodki miejskie, klastry miejskie, komórki, siatki obszarów wiejskich,
- ustalana na poziomie lokalnych jednostek administracyjnych (ang. *Local Administrative Units - abbr. LAU*). W Polsce w ramach tej typologii wykorzystuje się podział na gminy określając stopień urbanizacji (ang. *DEGURBA - Degree of Urbanization*), funkcjonalne obszary miejskie (ang. *Functional Urban Areas - abbr. FUA*) oraz obszary przybrzeżne,
- ustalone na poziomie jednostek NUTS 3, tj. typologia miejsko-wiejska, typologia metropolitalna oraz typologia przybrzeżna.

Szersze informacje na temat poszczególnych unijnych typologii terytorialnych dostępne są w publikacji Eurostatu (2019). Warto podkreślić, iż istotną typologią w kontekście tematu rozprawy doktorskiej są funkcjonalne obszary miejskie, które składają się z miasta i jego strefy dojazdów do pracy¹⁸.

Projekt Urban Audit

Potrzeba korzystania z typologii miast i obszarów powiązanych z nimi funkcjonalnie związana jest między innymi z uruchomionym pod koniec lat dziewięćdziesiątych projektem *Urban Audit*. Jak zaznaczają Kaczmarek i Bul (2014), jednym z istotniejszych celów tego programu jest gromadzenie danych nie tylko dla miast, ale również dla ich stref podmiejskich. Jest on jednocześnie odpowiedzią na zwiększający się popyt na obiektywną oraz porównywalną informację statystyczną w tych układach przestrzennych.

Początkowo w projekcie brały udział miasta liczące co najmniej 100 tys. mieszkańców. Z czasem próg ten został obniżony do 50 tys. Obecnie w projekcie uczestniczą miasta z 28 krajów członkowskich UE oraz Norwegii, Szwajcarii, a także Turcji. Wskazane obszary objęte projektem wyznaczane są na podstawie ustalonych wspólnych wytycznych metodologicznych.

¹⁷Rozporządzenia Parlamentu Europejskiego i Rady Unii Europejskiej 2017/2391 z dnia 12 grudnia 2017 r. zmieniającego rozporządzenie (WE) nr 1059/2003 w odniesieniu do typologii terytorialnych (Tercet) (Dz. Urz. UE L 350 z 29.12.2017).

¹⁸Szczegółowy opis przyjętej metodologii wyznaczania FUA, a także obszaru miasta znajduje się w opracowaniu Eurostatu (2017).

Dane statystyczne w ramach tego projektu zbierane są w kolejnych jego edycjach. Pierwsza pilotażowa faza rozpoczęła się w maju 1998 roku. Ostatnio zrealizowana została siódma faza projektu, za lata 2016–2018. W ramach ósmej fazy programu zbierane są informacje na czterech poziomach: funkcjonalnych obszarów miejskich (FUA), miast, grupy miast, a także obszarów wewnątrzmijskich (Eurostat, 2017).

Należy jednakże podkreślić, że zharmonizowana metodologia wyznaczania funkcjonalnych obszarów miejskich stosowana w ramach projektu Urban Audit kontrastuje z podejściem przyjętym do implementacji instrumentu ZIT. Dlatego też, warto przyjrzeć się problematyce delimitacji tego typu obszarów.

1.4. Delimitacja miejskich obszarów funkcjonalnych

Jak pisze Serafin (2018), złożoność problemu związanego z określeniem zasięgu oddziaływania miasta wynika „z wielowymiarowości istniejących relacji między miastem i jego otoczeniem o charakterze sektorowym, rzeczowym, czasowym i przestrzennym”. Jednocześnie, jak już wcześniej wykazano, rosnące znaczenie tak określanych obszarów w polityce krajowej oraz europejskiej jest na tyle istotne, aby przyjrzeć się bliżej temu zagadnieniu. Jak podkreśla Woźniak (2015), wyznaczanie obszarów funkcjonalnych powinno być oparte na analizie obiektywnych cech przestrzennych oraz występujących powiązań funkcjonalnych na danym obszarze, który nie musi być tożsamy z przyjętymi granicami administracyjnymi. Dlatego też, jak pisze Bul i Walaszek (2015), należy przyjąć odpowiednie kryteria delimitacyjne. Przyjęte wartości wskaźników umożliwiają określenie stopnia powiązań funkcjonalno-przestrzennych z jednostką, która stanowi punkt odniesienia. W literaturze przedmiotu jest wiele publikacji poruszających zagadnienie delimitacji miejskich obszarów funkcjonalnych (Śleszyński, 2013a; Heffner i Gibas, 2013; Herbst i Wójcik, 2013). Obszerny ich przegląd w kontekście Polski, można odnaleźć między innymi w pracy Sudra (2018). Jak wspomina autor tego opracowania, delimitacja może być rozpatrywana w dwóch wymiarach: badawczym (poznawczym) oraz praktycznym (aplikacyjnym). Na potrzeby niniejszej pracy dokonano przeglądu metod delimitacji w wymiarze aplikacyjnym. Jak podkreślono, wyznaczanie tych obszarów w takim ujęciu jest realizowane szczególnie na potrzeby kształtowania polityki przestrzennej, zarządzania określonymi obszarami, a także w kontekście dostarczania odpowiedniej informacji statystycznej.

Poniżej zostaną pokrótce scharakteryzowane i przedstawione na odpowiednim wykresie mapowym wybrane delimitacje miejskich obszarów funkcjonalnych, które wynikają z podjętej współpracy międzynarodowej, czy też mają swoje umocowanie w regulacjach krajowych. Pierwszy rodzaj delimitacji sprowadza się do wyznaczenia tych obszarów w ramach opisanego już wcześniej programu Urban Audit. Początkowo podczas realizacji pierwszych edycji tego programu w Polsce brakowało informacji statystycznych o dojazdach do pracy. W tej sytuacji, do wyznaczania tzw. szerszych stref miejskich (ang. *Larger Urban Zone - LUZ*) poszczególnych miast włączonych do projektu, przyjęto jeden lub więcej pierścieni gmin, czy też powiatów okalających miasto, w zależności od wielkości miasta wyrażonej liczbą mieszkańców¹⁹. Natomiast obecnie, zgodnie z przyjętym standardem określonym przez Eurostat, delimitacja tych stref odbywa się na podstawie dojazdów do pracy. Podstawową jednostką delimitacji jest gmina. Funkcjonalne obszary miejskie – FUA (wcześniej określane mianem szerszych stref miejskich – LUZ) – to obszary, z których co najmniej 15% i więcej populacji dojeżdża do pracy do miasta. Są one wyznaczane na podstawie informacji zawartych w opracowaniu „Przepływy ludności związane z zatrudnieniem w 2006 r.” przygotowane w Ośrodku Statystyki Miast w Urzędzie Statystycznym w Poznaniu²⁰.

Kolejną propozycję delimitacji, która w swoich założeniach miała obejmować jedynie miasta wojewódzkie i ich obszary, opracował w 2012 roku Śleszyński (2013b). Praca została przeprowadzona na zamówienie ówczesnego Ministerstwa Rozwoju Regionalnego. Jej wyniki opublikowano w dokumencie *Kryteria delimitacji miejskich obszarów funkcjonalnych ośrodków wojewódzkich*. Jak zauważa autor tego opracowania, potrzeba wyznaczenia tego typu obszarów wynikała z uregulowań zawartych w Koncepcji Przestrzennego Zagospodarowania Kraju 2030 (Rada Ministrów, 2012). Dlatego też w badaniu założono przygotowanie podejścia na tyle uniwersalnego, aby dało się je zastosować dla wszystkich miast wojewódzkich, a które dodatkowo uwzględniałoby specyfikę każdego z nich. Ponadto chodziło o przyjęcie takich założeń metodologicznych i metodycznych, aby nawiązywały one między innymi do prac związanych z wyznaczaniem obszarów LUZ wykonanych w ramach programu Urban Audit. Podstawową jednostką delimitacji była gmina, co ma swoje

¹⁹Więcej informacji na temat przyjętej wówczas metodologii delimitacji można znaleźć w pracy Młodak (2006), czy też na temat jednej z kolejnych modyfikacji procesu delimitacji w opracowaniu Zuzańskiej-Zyśko (2017)

²⁰Szerszy opis wytycznych dla uczestników projektu w zakresie wyznaczania FUA, a także obszaru miasta znajduje się w pracy Eurostatu (2017).

uzasadnienie w wymiarze merytoryczno-praktycznym. Odbyto również kilka spotkań konsultacyjnych z przedstawicielami samorządów. Następnie po wykonaniu serii testów opracowano ostateczny wykaz siedmiu wskaźników przydzielonych do jednej z trzech grup służących do przeprowadzenia procesu delimitacji. Jedną grupę stanowiły wskaźniki funkcjonalne. Przykładowo, do konstrukcji jednego z nich skorzystano z informacji o dojazdach do pracy. Kolejną grupę tworzyły wskaźniki społeczno-gospodarcze, w których zastosowano między innymi dane o pracujących w zawodach pozarolniczych, czy też podmiotach gospodarczych. Trzecia grupa składała się z mierników morfologicznych takich jak gęstość zaludnienia oraz liczba mieszkań oddanych do użytku²¹. Jak zaznacza Śleszyński (2013a), wykonana delimitacja obszarów funkcjonalnych miast wojewódzkich miała z założenia charakter uniwersalny. Wyznaczone w ten sposób obszary można porównywać pomiędzy sobą, gdyż opierają się na tych samych założeniach metodycznych. Jednak, jak podkreśla sam autor opracowania, narzuca to pewne ograniczenia w przypadku zindywidualizowanych potrzeb zmian granic tych obszarów. Dlatego też sugeruje, że w wymiarze praktycznym ich terytorium powinno być ostatecznie ustalane na poziomie samorządów czy też środowisk lokalnych.

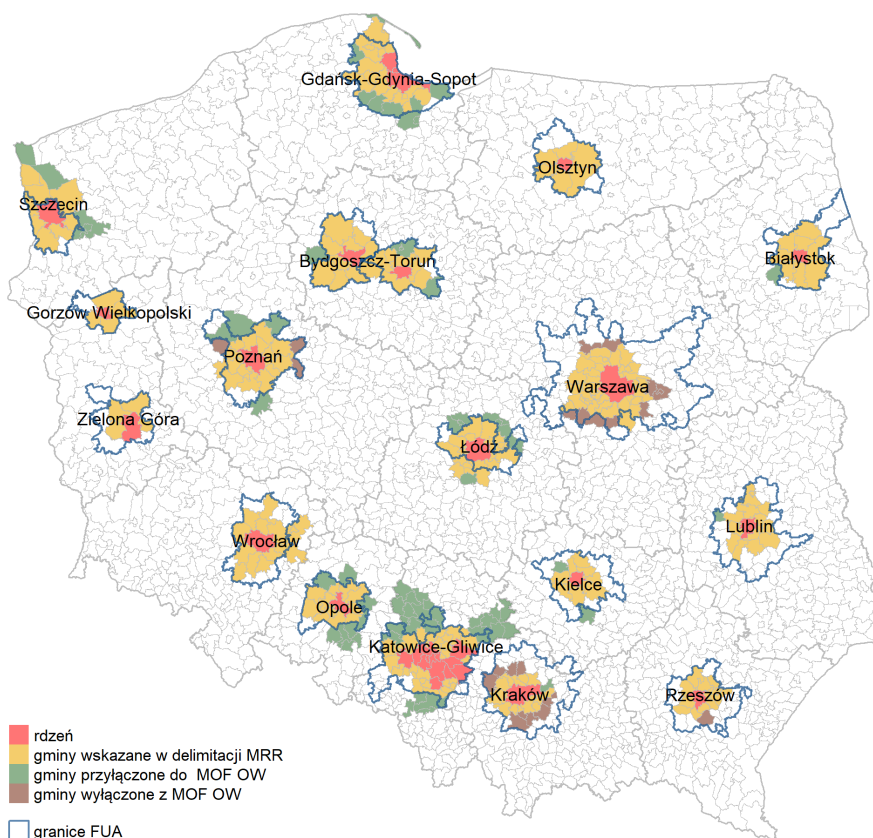
Taki też kierunek działań w pewnym wymiarze został podjęty w celu realizacji instrumentu ZIT. Podstawą do wyznaczenia obszarów wsparcia, był wspomniany już wcześniej dokument *Kryteria delimitacji miejskich obszarów funkcjonalnych ośrodków wojewódzkich* (Śleszyński, 2013b). Jak wskazano w Umowie partnerstwa (MIiR, 2017) przynajmniej połowa gmin tworzących poszczególne obszary funkcjonalne miast wojewódzkich objęte ZIT-ami, musi być tożsama z gminami, które zostały wyznaczone w ramach delimitacji przeprowadzonej przez Śleszyńskiego (2013b). Zmiany w ich składzie w odniesieniu do dokumentu wyjściowego mogą wynikać ze szczególnych powiązań funkcjonalnych, których nie dało się uchwycić w bazowym opracowaniu, a także ze względu na ustalenia wynikające z regionalnych planów i strategii rozwoju. Takie wytyczne do wyznaczania obszarów ZIT dawały samorządom województw względnie dużo swobody jeżeli chodzi o ustalanie ich ostatecznego kształtu. Skategoryzowania ostatecznie zastosowanych podejść do wyznaczania obszarów ZIT, w odniesieniu do poszczególnych miast wojewódzkich, dokonała Kociuba (2017a), dzieląc je na kilka wariantów:

²¹Szczegółowy opis przyjętej procedury delimitacyjnej oraz wykaz wskaźników znajduje się w pracy Śleszyńskiego (2013a).

- ministerialny (9 obszarów ZIT): nie dokonano żadnych korekt, bądź były one na poziomej +/- jednej gminy w stosunku do wytycznych Ministerstwa Rozwoju Regionalnego,
- współpracy (3 obszary ZIT): podstawą ich delimitacji była wcześniejsza współpraca jednostek samorządu terytorialnego w ramach istniejących stowarzyszeń,
- strategiczno-planistyczny (2 obszary ZIT): podstawą określenia tych obszarów była delimitacja przeprowadzona w strategiach rozwoju województw; wzięto też pod uwagę uzgodnienia jednostek samorządu terytorialnego współpracujących lub deklarujących współpracę w ramach stowarzyszeń,
- ekspercki (2 obszary ZIT): władze województw zleciły opracowanie odpowiednich ekspertyz lub same dokonały analiz w celu wyznaczenia obszaru wsparcia,
- doraźny (2 obszary ZIT): powołano stowarzyszenia jedynie w celu realizacji instrumentu ZIT.

Rysunek 1.2 przedstawia różnice pomiędzy delimitacją opisaną w pracy Śleszyńskiego (2013b), a ostatecznym podziałem zatwierdzonym przez poszczególne samorządy województw na potrzeby realizacji Zintegrowanych Inwestycji Terytorialnych. Dodatkowo naniesione są także obecnie granice miejskich obszarów funkcjonalnych, w ramach projektu Urban Audit.

Jak widać na rysunku 1.2, podstawową jednostką podziału we wszystkich przytoczonych delimitacjach jest gmina. Jednak w zależności od przyjętego sposobu ich wyznaczania obszary te różnią się zasięgiem. Dodatkowo należy również mieć na uwadze dynamikę zachodzących zmian społeczno-gospodarczych oraz funkcjonalnych na terenach tak wyznaczonych i bezpośrednio do nich przyległych. Mogą się one przełożyć na wartości wskaźników, które decydują o zaliczeniu danej jednostki do obszaru funkcjonalnego.



Rysunek 1.2: Porównanie miejskich obszarów funkcjonalnych wyznaczonych dla miast wojewódzkich w ramach programu Urban Audit do realizacji instrumentu ZIT

Źródło: Opracowanie własne.

Tak jak w przypadku delimitacji obszarów wsparcia w ramach ZIT-ów widać, że trudno jest przyjąć jedno obowiązujące wszystkich kryteria delimitacji, niezależnie od specyficznych wewnętrznych uwarunkowań poszczególnych obszarów. Wydaje się też, że w tym względzie odgrywają rolę pewne polityczne czynniki czy też historyczne powiązania na danych obszarach. Mając powyższe na uwadze dużym wyzwaniem oraz kluczową potrzebą staje się zapewnienie odpowiedniego pokrycia informacyjnego dla tak wyznaczanych obszarów, które dodatkowo mogą się również zmieniać w czasie. Istotne wydaje się także posiadanie odpowiednich informacji statystycznych w celu określenia związków funkcjonalnych pomiędzy poszczególnymi obszarami.

1.5. Informacja statystyczna w kontekście miejskich obszarów funkcjonalnych

1.5.1. Zgłaszane potrzeby informacyjne

W środowisku naukowym, szczególnie w ostatnich latach, zgłaszana jest potrzeba zapewnienia odpowiedniej infrastruktury statystycznej dla miejskich obszarów funkcjonalnych, w tym dla obszarów metropolitalnych. Postulat ten podnoszony jest zarówno w kontekście ich precyzyjnego wyznaczenia, jak również diagnozy oraz monitoringu zmian zachodzących na tak wyznaczonych obszarach. Wskazywana jest potrzeba opracowania zbioru zharmonizowanych wskaźników, który pozwoliłby na porównania tych obszarów zarówno na poziomie krajowym jak i międzynarodowym. Podkreśla się również, że badanie poziomu rozwoju gospodarczego tych obszarów nie jest tożsame z analizami prowadzonymi dla wyższych poziomów agregacji przestrzennej (województw, podregionów), dla których zakres dostępności odpowiednich danych jest znacznie szerszy. Dodatkowym problemem jest brak spójnego podejścia w delimitacji tych obszarów. Wiele różnych podmiotów i instytucji dokonuje niezależnych delimitacji, w których granice wyróżnionych jednostek różnie przebiegają. Należy jednak podkreślić, że podejmowane próby wynikają z rosnącego popytu na wiarygodne i aktualne dane statystyczne na niskim poziomie agregacji. Bowiem jak wiadomo, jedynie wysokiej jakości informacje mogą zapewnić skuteczne zarządzanie daną jednostką terytorialną, w tym również tak wyznaczonymi obszarami (Batóg i in., 2015; Kaczmarek, 2015; Młodak, 2012). Jak zaznacza Rogalińska (2017), poza dostarczaniem przez statystykę publiczną danych w podziale na jednostki administracyjne oraz statystyczne, prowadzone są również prace badawcze dotyczące delimitacji oraz monitorowania obszarów funkcjonalnych, w tym obszarów metropolitalnych. Jednocześnie autorka podkreśla, że należy prowadzić pogłębione analizy i badania w tym zakresie przy współudziale wielu interesariuszy. Zaznacza ponadto, że potrzebne jest tu wykorzystanie nowych źródeł informacji oraz rozwój innowacyjnych i zaawansowanych metod badawczo-analitycznych, do których zalicza między innymi metody statystyki małych obszarów.

Również w opracowaniu *Wskaźniki strategiczne na poziomie miejskich obszarów funkcjonalnych* (Truskolaski i in., 2016), będącym trzecim raportem przygotowanym w ramach

badania dotyczącego *Monitoringu strategicznego miejskich obszarów funkcjonalnych realizujących Zintegrowane Inwestycje Terytorialne w Polsce*, wskazano na istotne braki w pokryciu informacyjnym. Jako powód wskazano brak regulacji prawnych, które umocowałyby działalność miejskich obszarów funkcjonalnych, przykładowo w ramach rozporządzenia Rady Ministrów. Wobec braku podstaw prawnych statystyka publiczna nie dostarcza informacji koniecznych do prowadzenia badań i analiz tych obszarów, szczególnie na potrzeby ich monitoringu strategicznego. Dlatego też, podczas przygotowywania strategii ZIT-ów przedstawione w nich analizy społeczno-gospodarcze opierały się czasami na mniej wiarygodnych źródłach. Zauważono również brak ujednoczonych wskaźników strategicznych w ramach tych samych priorytetów inwestycyjnych realizowanych przez różne obszary ZIT. Były one oparte na wielu różnych źródłach danych, w tym często na oszacowaniach eksperckich. Brak harmonizacji w tym zakresie nie daje możliwości przeprowadzenia analiz porównawczych i możliwości monitorowania zmian pomiędzy poszczególnymi obszarami ZIT.

Warto również zaznaczyć, jak podaje Szafranek (2017a), że 46% związków ZIT jako jeden z celów priorytetowych w swoich strategiach wskazało poprawę sytuacji na rynku pracy. Szersza analiza kierunków planowanej poprawy sytuacji na rynku pracy, w tym poprzez wspieranie samozatrudnienia, przedsiębiorczości oraz tworzenia przedsiębiorstw i nowych miejsc pracy, a także zwiększenie zatrudnienia osób bezrobotnych znajduje się w opracowaniu Szafranek i Kociuby (2018). Autorzy wskazują jednocześnie na mniejsze znaczenie w strategiach działań adresowanych do osób bezrobotnych. Wynika to najprawdopodobniej z założenia, że w większości miejskich obszarów funkcjonalnych zjawisko bezrobocia jest stosunkowo niewielkie, a nacisk należy położyć na działania dostosowujące zasoby pracy do zmieniających się warunków na rynku pracy.

Wydaje się jednak, że dokładne zbadanie sytuacji wyjściowej w poszczególnych obszarach ZIT, jak i ich cykliczny monitoring to działania, które należy pilnie podjąć. Również w dokumentach rządowych wskazuje się na potrzebę zapewnienia odpowiedniej informacji statystycznej na poziomie miejskich obszarów funkcjonalnych, w tym informacji na temat rynku pracy.

Jednym z pierwszych dokumentów poświęconych tym zagadnieniom są Założenia Krajowej Polityki Miejskiej do roku 2020 (MRR, 2012). Wskazano w nim na duże zapotrzebowanie na informację statystyczną w obrębie miejskich obszarów funkcjonalnych, gdyż zakładano

że „stworzenie kompleksowego systemu monitoringu i ewaluacji zjawisk w miastach i miejskich obszarach funkcjonalnych doprowadzi do tego, że wszelkie aspekty rozwoju miast staną się stałym elementem prac analitycznych prowadzonych na poziomie krajowym w ramach krajowego obserwatorium terytorialnego”.

W ogłoszonej Krajowej Polityce Miejskiej 2023 (MliR, 2015a) w odniesieniu do monitorowania realizacji polityki miejskiej podkreślono, że minister właściwy ds. rozwoju regionalnego, we współpracy z Głównym Urzędem Statystycznym, ma rozwijać i poświęcać więcej uwagi działaniom, które mają na celu między innymi zwiększenie możliwości cyklicznego badania zjawisk i występujących tendencji w granicach miejskich obszarów funkcjonalnych, biorąc pod uwagę również dynamikę ich rozwoju. Wskazano także konkretną bazę STRATEG, której funkcjonalność, w miarę możliwości, powinna być uzupełniana o dane dotyczące tego typu obszarów, w tym wyznaczonych w ramach ZIT-ów. W dokumencie zaznaczono również, że „konkurencyjność i zdolność do rozwoju głównych ośrodków miejskich budowana będzie także przez stopniowe rozwijanie zarządzania całymi obszarami funkcjonalnymi, co przyczyniać się będzie do synergicznego wykorzystania potencjałów całych obszarów zurbanizowanych i lokalnych rynków pracy”. Wspomniano w niej również, że miasta odgrywają istotną rolę w tworzeniu nowych miejsc pracy. Zatem odpowiednia baza informacyjna na temat rynku pracy, w ujęciu obszarów funkcjonalnych, wydaje się niezwykle istotna chociażby w celu monitorowania prowadzonej w tym zakresie polityki miejskiej czy też diagnozy sytuacji na lokalnym rynku pracy obszaru funkcjonalnego.

W ogłoszonej w 2017 roku strategii SOR (Morawiecki, 2017) odniesiono się do rozwoju instrumentu ZIT. Jednym z projektów strategicznych, który został podany w dokumencie szczegółowej dyskusji jest projekt Zintegrowane Inwestycje Terytorialne PLUS. Został on podzielony na dwa etapy. Pierwszy z nich miał w głównej mierze polegać na zaplanowanym do roku 2020 bardziej efektywnym i sprawnym wdrażaniu obecnego instrumentu ZIT. Natomiast po 2020 roku działania skoncentrowane mają być na wypracowaniu zasad realizacji ZIT 2020 PLUS. Mają one bazować na funduszach unijnych oraz na rozwiązaniach krajowych umożliwiających sprawną współpracę jednostek samorządu terytorialnego, a także realizację wspólnych strategii rozwojowych w układzie miejskich obszarów funkcjonalnych. Rozwiązania te mają dotyczyć między innymi kreowania rozwoju metropolitalnego czy też poprawy zarządzania, ale także prowadzenia analiz rozwojowych w ramach miejskich obszarów funkcjonalnych. W przytoczonej strategii (Morawiecki, 2017) w ramach działań

dotyczących poprawy warunków rozwojowych polskich miast, jako jedno z kluczowych działań, które mają być realizowane już do 2020 roku wymienia się „Usprawnienie systemu monitorowania i diagnozowania sytuacji społeczno-gospodarczej i przestrzennej na poziomie miejskich obszarów funkcjonalnych, w tym dostosowanie metodologii statystyki publicznej, rozbudowa lokalnych, regionalnych i krajowych centrów wiedzy na temat obszarów miejskich”. W strategii tej podniesiona jest również kwestia wzmocnienia zdolności ośrodków miejskich do tworzenia miejsc pracy. Podkreślono jednak, iż aby to osiągnąć, niezbędna jest poprawa zdolności tych obszarów do kreowania wzrostu zatrudnienia (Morawiecki, 2017). W dokumencie poruszono ponadto kwestię tworzenia miejsc pracy oraz kreowania wzrostu zatrudnienia, co wiąże się jednoznacznie z koniecznością posiadania dostępności do odpowiedniej informacji na poziomie, między innymi, miejskich obszarów funkcjonalnych.

W KSRR 2030 (MliR, 2019), ogłoszonej w 2019 roku, podkreśla się potrzebę produkcji odpowiednich informacji statystycznych umożliwiających monitorowanie rozwoju miast i obszarów powiązanych funkcjonalnie. Również w tym dokumencie wskazuje się, że planowane jest wykorzystywanie w szerszym zakresie lokalnych potencjałów prorozwojowych, w tym między innymi dostępnych zasobów pracy na obszarach miejskich, a także wspieranie przedsiębiorczości w ośrodkach miejskich, szczególnie w ośrodkach wojewódzkich posiadających odpowiedni potencjał zarówno naukowy jak i gospodarczy. Podkreśla się także, że dotychczas to największe aglomeracje najbardziej skorzystały z działań prowadzonych w ramach polityki regionalnej. Pierwotnie zakładano, że impulsy rozwojowe będą się rozprzestrzeniały na otaczające je obszary słabiej rozwinięte. Stwierdzono jednak, że oddziaływanie to było niewystarczające, co przyczyniło się do wzrostu dysproporcji zarówno międzyregionalnych, jak i wewnątrzregionalnych, a utrzymujące się zróżnicowanie jest między innymi skutkiem niezrównoważenia popytu i podaży na lokalnych rynkach pracy. Dlatego też w nowej KSRR 2030 obszary te nie będą stanowiły obszarów strategicznej interwencji państwa. Nadal jednak będą korzystały ze wsparcia z poziomu krajowego oraz regionalnego na warunkach ogólnych. Podkreślono jednak, że jedną ze strategicznych inicjatyw istotnych dla realizacji celów KSRR jest projekt Zintegrowane Inwestycje Terytorialne PLUS (MliR, 2019). Również ten dokument podkreśla konieczność priorytetowego traktowania obszarów znajdujących się poza strefą oddziaływa aglomeracji, między innymi w kwestiach dotyczących rynku pracy. To wskazuje również, że warto podjąć

działania analityczne, aby zweryfikować tezę o dużych dysproporcjach w obszarze rynku pracy pomiędzy aglomeracjami, a pozostałymi obszarami. Do realizacji tego celu należy uwzględnić informacje o wszystkich podmiotach na rynku pracy, a więc o osobach pracujących, bezrobotnych oraz biernych zawodowo, z których przynajmniej część może być aktywizowana.

1.5.2. Zakres informacji dostarczanej przez służby statystyki publicznej

Obecnie pewien zakres zmiennych na poziomie miejskich obszarów funkcjonalnych dostępny jest w dwóch różnych układach, w zależności od przyjętego sposobu ich delimitacji. Jeden z nich wynika z uzgodnień międzynarodowych, a dotyczy realizacji omówionego wcześniej projektu Urban Audit. Drugi natomiast odnosi się do obszarów wyznaczonych w celu realizacji instrumentu ZIT.

Zakres wymaganej informacji dla różnych podziałów przestrzennych przyjętych w projekcie Urban Audit, w tym dla miejskich obszarów funkcjonalnych, podlega w każdej rundzie projektu pewnym modyfikacjom. Obecny zakres zmiennych w ramach projektu został określony w opracowaniu *Methodological manual on city statistics* (Eurostat, 2017). Natomiast szeroki wachlarz danych zebranych dla miast oraz miejskich obszarów funkcjonalnych w dziedzinie demografii, edukacji, warunków życia, środowiska, a także rynku pracy w ramach poszczególnych edycji projektu znajduje się w bazie Eurostatu, do której dostęp możliwy jest poprzez stronę internetową²². Tam też, jak wynika z informacji umieszczonej na stronie internetowej Głównego Urzędu Statystycznego opisującej projekt²³, zawarte są wszystkie informacje przekazane przez GUS w ramach projektu. Analizując zakres udostępnionej informacji dla miejskich obszarów funkcjonalnych biorących udział w projekcie ze strony polskiej, szczególnie w obszarze rynku pracy, zauważyć można brak dostępnej informacji chociażby o liczbie osób pracujących w podziale na płeć i określone grupy wieku dla poszczególnych lat badania. Natomiast informacja w zakresie liczby osób bezrobotnych dostępna jest dla zawężonego zakresu lat w podziale na płeć, jednak zgodnie z prawną definicją osoby bezrobotnej obowiązującą w Polsce (ustawa z dnia 20 kwietnia

²²<http://ec.europa.eu/eurostat/web/cities/data/database>

²³<https://stat.gov.pl/statystyka-regionalna/badania-regionalne/urban-audit-250/>

2004 r. o promocji zatrudnienia i instytucjach rynku pracy, Dz.U. 2004 nr 99, poz. 1001), co utrudnia porównania międzynarodowe w tym zakresie.

Podstawowy zakres informacji statystycznej dla obszarów wyznaczonych w celu realizacji instrumentu ZIT dostępny jest obecnie jako ujednoczone²⁴ opracowanie Vademecum Samorządowca dla poszczególnych obszarów²⁵. Znaleźć tam można syntetyczną informację statystyczną dotyczącą obszarów takich jak demografia, mieszkalnictwo, wychowanie i edukacja, kultura, sport, zdrowie, pomoc społeczna, środowisko oraz dochody i wydatki budżetów gmin należących do ZIT-ów. Broszura zawiera również podstawowe informacje o strukturze podmiotów prowadzących działalność gospodarczą oraz informacje o liczbie osób bezrobotnych (zarejestrowanych w urzędach pracy) oraz pracujących (według siedziby przedsiębiorstwa i z wyłączeniem pewnych kategorii pracujących). Informacje te nie odnoszą się jednak do tej samej populacji. Uzyskiwane są bowiem na podstawie różnych źródeł, które charakteryzują się innymi założeniami metodologicznymi, co utrudnia ich łączną interpretację. Dlatego też, warto przyjrzeć się dostępnym źródłom danych na temat rynku pracy w celu zdiagnozowania przyczyn prezentowania tych informacji w takich ujęciach.

Podobnie sytuacja wygląda w opracowaniach poszczególnych strategii ZIT, których przygotowanie było obligatoryjne, aby mogły one funkcjonować na wyznaczonych obszarach. Wszystkie strategie zawierające m.in. wszechstronny zestaw informacji statystycznych na ich temat, zostały udostępnione również poprzez bazę STRATEG²⁶, która zawiera też wskaźniki monitorujące realizację poszczególnych strategii. Również w tym przypadku zakres informacji na temat rynku pracy jest zróżnicowany, szczególnie w kontekście diagnozy stanu na tych obszarach. Przykładowo, poszczególne strategie zawierały informację o liczbie osób bezrobotnych w podziale na płeć i wybrane grupy wieku, jednak wyłącznie bazując na bezrobociu rejestrowanym (ZIT, 2015b, 2016a). Prezentowany był również wskaźnik bezrobocia zawierający informację na poziomie gmin tworzących obszar funkcjonalny (ZIT, 2016b). Także dane o pracujących na poziomie poszczególnych gmin obszaru funkcjonalnego mają pewne ograniczenia informacyjne wynikające z przyjętych założeń metodologicznych. Podnoszony jest również problem bierności zawodowej (ZIT, 2015a),

²⁴W początkowym okresie jak pisze Batóg i in. (2015) nie były one opracowywane przez wszystkie regionalne urzędy statystyczne. Dodatkowo nie zawierały tego samego zakresu danych, co utrudniało ich porównywanie.

²⁵<https://stat.gov.pl/statystyka-regionalna/statystyczne-vademecum-samorzadowca/>

²⁶<https://strateg.stat.gov.pl/>

natomiast brak jest odpowiednich informacji opisujących to zjawisko. Zawarte tam informacje na temat liczby osób pracujących odnoszą się jedynie do pracujących w podmiotach zatrudniających więcej niż 9 osób i jest to informacja związana z siedzibą firmy, a nie faktyczną liczbą zatrudnionych zamieszkałych w danym obszarze.

Również wskaźniki monitorujące osiągnięcie celów priorytetowych w części strategii wyznaczone są na podstawie informacji, których wyliczenie wymaga danych o liczbie pracujących, czy też bezrobotnych. Przykładem może tu być jeden ze wskaźników monitorujących określony w strategii rozwoju Rzeszowskiego Obszaru Metropolitalnego – współczynnik feminizacji, tj. liczba pracujących kobiet przypadająca na 100 mężczyzn w grupie wiekowej 20–34 lata (Batóg i in., 2015).

Występujący brak pokrycia informacyjnego dla wybranych wskaźników na temat rynku pracy w ramach projektu Urban Audit, czy też bardzo podstawowa informacja z tego obszaru zawarta w biuletynie Vademecum Samorządowca, a także zróżnicowany zakres informacyjny na temat rynku pracy w poszczególnych strategiach ZIT jest odzwierciedleniem dostępnych informacji dla tych poziomów agregacji przestrzennej. W związku z powyższym kolejny rozdział pracy poświęcony został analizie dostępnych zasobów informacyjnych w tym zakresie oraz ocenie ich użyteczności.

1.6. Podsumowanie

Analiza dokumentów strategicznych oraz ustaw odnoszących się do miejskich obszarów funkcjonalnych, a także prowadzona dyskusja w tym zakresie, zarówno w literaturze światowej jak i krajowej, świadczy o dużej wadze tak definiowanych obszarów. Nabierają one również dużego znaczenia w kontekście realizowanej polityki spójności, w ramach której do obszarów tych kierowane są środki na rozwój poprzez takie instrumenty jak Zintegrowane Inwestycje Terytorialne. Realizowany od końca lat 90-tych projekt Urban Audit również wskazuje na duże znaczenie obszarów funkcjonalnych miast. Są one jednym z poziomów agregacji przestrzennej, dla których zbierane są w ramach tego projektu dane statystyczne.

Przytoczone przykłady różnych typów delimitacji i dynamicznie zmieniająca się w czasie sytuacja w tych układach wskazują na trudności w dostarczeniu odpowiedniej jakości

informacji statystycznych odnoszących się do tych obszarów. Jednocześnie wyraźnie zauważalne jest rosnące zapotrzebowanie na tego typu dane zgłaszane zarówno ze strony środowiska naukowego, jak i w coraz większej liczbie krajowych dokumentów strategicznych. Diagnoza pokrycia informacyjnego odnośnie tych obszarów wskazuje natomiast na deficyt danych, szczególnie w kontekście zmiennych dotyczących rynku pracy. Przyjęto założenie, że docelowymi na tym poziomie agregacji przestrzennej pożądanymi przez odbiorców danych wskaźnikami z obszaru rynku pracy mogą być stopa bezrobocia, wskaźnik zatrudnienia oraz współczynnik aktywności ekonomicznej, w podziale na płeć oraz cztery grupy wieku (tj. 15–29, 30–39, 40–49, 50 i więcej). Dlatego ważnym elementem prac badawczych podjętych przez autora, mających charakter teoretyczny była analiza obecnie dostępnego zasobu informacyjnego w tym zakresie. Jej wyniki przedstawiono w kolejnym rozdziale. Jej celem będzie analiza dostępności w wybranych źródłach danych informacji o liczbie osób pracujących, bezrobotnych oraz biernych zawodowo w wyżej wymienionych przekrojach, które służą do wyliczenia wytypowanych wskaźników. Założono również, iż na podstawie przeprowadzonej kwerendy zgłaszanego zapotrzebowania, weryfikowana będzie przede wszystkim analiza wybranych źródeł danych w kontekście możliwości dostarczenia na ich podstawie informacji dla miejskich obszarów funkcjonalnych wyznaczonych do realizacji instrumentu ZIT w latach 2014–2020, w tym szczególnie dla miejskich obszarów funkcjonalnych miast wojewódzkich.

Rozdział 2

Wybrane źródła danych w obszarze ryнку pracy

Niniejszy rozdział poświęcony jest omówieniu kluczowych źródeł danych¹ dotyczących rynku pracy, które w całości lub częściowo dotyczą populacji osób pracujących, bezrobotnych oraz biernych zawodowo. Należy mieć jednak na uwadze, że dostępne w nich informacje różnią się pomiędzy sobą sposobem ich pozyskiwania, definicjami poszczególnych kategorii rynku pracy oraz rodzajem jednostek podlegających badaniu. Dotarcie do pożądaných informacji na temat rynku pracy jest możliwe poprzez badanie gospodarstw domowych, przedsiębiorstw lub też osób ujętych w rejestrach administracyjnych², takich jak chociażby rejestr osób bezrobotnych. Ważnym źródłem są również narodowe spisy powszechne, które poza szerokim spektrum tematów odnoszących się do demografii, niepełnosprawności czy mieszkań poddają kompleksowej analizie zagadnienia z zakresu rynku pracy.

¹Przyjęto w pracy, iż oprócz pojęcia **źródła danych** zamiennie stosowane będą określenia: źródła, zasoby danych, zasoby informacyjne, zbiory informacyjne.

²Przyjęto w pracy, iż oprócz pojęcia **rejestry administracyjne** zamiennie stosowane będą określenia: administracyjne źródła danych, administracyjne zbiory danych, systemy informacyjne administracji publicznej.

Biorąc pod uwagę powyższe informacje oraz rozważane kategorie zasobów przyjęto podział na badania prowadzone przez Główny Urząd Statystyczny poprzez gospodarstwa domowe, przedsiębiorstwa i jednostki sfery budżetowej, administracyjne źródła danych oraz spisy powszechne. Z każdej z wyżej wskazanych kategorii wybrano konkretne badania, które są istotne z punktu widzenia prowadzonej w ramach rozprawy analizy i dokonano ich oceny pod kątem ich wykorzystania w toku dalszych prac. Część rozdziału jest także poświęcona ocenie użyteczności poszczególnych zbiorów informacyjnych oraz zawiera analizę porównawczą pomiędzy źródłami danych w odniesieniu do dwóch bardzo ważnych kategorii rynku pracy tj. osób pracujących i bezrobotnych.

2.1. Badania prowadzone poprzez gospodarstwa domowe

Najważniejszym badaniem dostarczającym szerokiego zakresu danych o sytuacji na polskim rynku pracy jest Badanie Aktywności Ekonomicznej Ludności (BAEL), a podstawową jednostką badania jest w nim gospodarstwo domowe.

2.1.1. Podstawowe informacje

Badanie Aktywności Ekonomicznej Ludności jest jednym z kluczowych badań prowadzonych przez Główny Urząd Statystyczny w celu bieżącej oceny stopnia wykorzystania zasobów rynku pracy. Dostarcza ono informacji o rozmiarach i strukturze populacji osób pracujących, bezrobotnych oraz biernych zawodowo. Uwzględnia się w nim cechy demograficzno-społeczne i zawodowe ankietowanych osób oraz podstawowe charakterystyki miejsca pracy. Pozyskiwane dane służą do monitorowania tempa i kierunków zmian zachodzących w poziomie aktywności ekonomicznej ludności. Podstawą tego badania są obowiązujące regulacje prawne na poziomie zarówno krajowym, jak i międzynarodowym³.

Pierwszy raz badanie to zostało przeprowadzone w maju 1992 roku. Jego wdrożenie było ściśle związane z okresem transformacji systemowej i zapotrzebowaniem na dane pozwalające na uzupełnienie luki informacyjnej dotyczącej charakterystyki nowych zjawisk na

³Więcej informacji na ten temat napisano w publikacji GUS (2018a) oraz w artykule Zgierskiej (2017).

ryнку pracy⁴. Wraz z przystąpieniem Polski do Unii Europejskiej dane z tego badania stały się podstawą do opracowywania wielu kluczowych wskaźników ujętych w różnych strategiach i programach operacyjnych na poziomie zarówno europejskim, jak i krajowym. Warto też podkreślić, że jego wyniki wykorzystywane są przez wiele instytucji międzynarodowych.

2.1.2. Podstawowe kategorie i mierniki z obszaru rynku pracy - ujęcie definicyjne

Kluczowym kryterium podziału ludności w BAEL jest fakt wykonywania, posiadania lub poszukiwania pracy. Na tej podstawie wydziela się grupy osób, które definiuje się jako pracujące lub bezrobotne. Pozostała ludność niezakwalifikowana do powyższych kategorii zalicza się do osób biernych zawodowo. W ten sposób każda z badanych osób może być zaliczona tylko do jednej z trzech grup. Natomiast źródłem obowiązujących w badaniu definicji tych kategorii są zalecenia Międzynarodowej Organizacji Pracy (MOP) zawierające definicje przyjęte podczas XIII Międzynarodowej Konferencji Statystyków Pracy w październiku 1982 roku, które od tamtej pory uległy tylko niewielkim modyfikacjom. Jak wskazuje Wilak (2017), uściślenie rekomendacji MOP nastąpiło w aktach prawnych Komisji Europejskiej i są one obecnie stosowane przez kraje członkowskie Unii Europejskiej w badaniach siły roboczej (ang. *Labour Force Survey - LFS*). Zgodnie z przyjętymi określeniami (GUS, 2018a):

Pracujący to wszystkie osoby w wieku 15 lat i więcej, które w okresie badanego tygodnia:

- wykonywały przez co najmniej 1 godzinę pracę przynoszącą zarobek lub dochód tzn. były zatrudnione w charakterze pracownika najemnego, pracowały we własnym (lub dzierżawionym) gospodarstwie rolnym lub prowadziły własną działalność gospodarczą poza rolnictwem, pomagały (bez wynagrodzenia) w prowadzeniu rodzinnego gospodarstwa rolnego lub rodzinnej działalności gospodarczej poza rolnictwem,
- miały pracę, ale jej nie wykonywały:
 - z powodu choroby, urlopu macierzyńskiego, urlopu rodzicielskiego lub wypoczynkowego,

⁴Więcej informacji na temat przesłanek wdrożenia BAEL znajduje się w publikacji Witkowskiego (2017).

- z innych powodów, przy czym długość przerwy w pracy wynosiła:
 - do 3 miesięcy,
 - powyżej 3 miesięcy, ale osoby te były pracownikami najemnymi i w tym czasie otrzymywały co najmniej 50% dotychczasowego wynagrodzenia.

Do pracujących – zgodnie z międzynarodowymi standardami – zaliczani są również uczniowie, z którymi zakłady pracy lub osoby fizyczne zawarły umowę o naukę zawodu lub przyuczenie do określonej pracy, jeżeli otrzymywali wynagrodzenie.

Bezrobotni – osoby w wieku 15–74 lata, które spełniają jednocześnie trzy warunki:

- w okresie badanego tygodnia nie były osobami pracującymi,
- aktywnie poszukiwały pracy, tzn. podjęły konkretne działania w ciągu 4 tygodni (wliczając jako ostatni – tydzień badany), aby znaleźć pracę,
- były gotowe (zdolne) podjąć pracę w ciągu dwóch tygodni następujących po tygodniu badanym.

Do bezrobotnych zalicza się także osoby, które nie poszukiwały pracy, ponieważ miały pracę załatwioną i oczekiwały na jej rozpoczęcie przez okres nie dłuższy niż 3 miesiące oraz były gotowe tę pracę podjąć.

Ludność bierna zawodowo, tzn. pozostająca poza siłą roboczą. Są to wszystkie osoby w wieku 15 lat i więcej, które w badanym tygodniu:

- nie pracowały, nie miały pracy i jej nie poszukiwały,
- nie pracowały, poszukiwały pracy, ale nie były zdolne (gotowe) do jej podjęcia w ciągu dwóch tygodni następujących po tygodniu badanym,
- nie pracowały i nie poszukiwały pracy, ponieważ miały pracę załatwioną i oczekiwały na jej rozpoczęcie w okresie:
 - dłuższym niż trzy miesiące,
 - do 3 miesięcy, ale nie były gotowe tej pracy podjąć.

Bazując na zdefiniowanych kategoriach rynku pracy można powyższe zbiorowości opisywać za pomocą różnych wskaźników monitorujących sytuację na rynku pracy. Podstawowymi miarami służącymi do tego celu, które są również estymowane w ramach niniejszej rozprawy doktorskiej, są (GUS, 2019c):

- **współczynnik aktywności ekonomicznej** – procentowy udział aktywnych zawodowo danej kategorii w ogólnej liczbie ludności danej kategorii;
- **wskaźnik zatrudnienia** – procentowy udział pracujących danej kategorii w ogólnej liczbie ludności danej kategorii;
- **stopa bezrobocia** – procentowy udział bezrobotnych danej kategorii w liczbie aktywnych zawodowo danej kategorii.

2.1.3. Metodologia badania

Badanie Aktywności Ekonomicznej Ludności jest badaniem częściowym, prowadzonym kwartalnie metodą reprezentacyjną. Realizowane jest ono na próbie losowej umożliwiającej uogólnianie wyników badania na populację generalną. Stosowana jest w nim metoda obserwacji ciągłej. Dzięki temu wyniki badania obrazują sytuację na rynku pracy w okresie całego kwartału, a tym samym i sytuację w całym roku. Populację generalną stanowi ludność Polski w wieku 15 lat i więcej zamieszkała w gospodarstwach domowych⁵. Próby kwortalne są wybierane za pomocą losowania dwustopniowego. Jednostkami losowania pierwszego stopnia, poza sporadycznymi wyjątkami, są rejony statystyczne w miastach, a na wsi - obwody spisowe. Natomiast jednostkami losowania drugiego stopnia są mieszkania. Losowanie jednostek pierwszego oraz drugiego stopnia przeprowadza się na podstawie Operatu do Badań Społecznych, który zawiera między innymi wykaz terytorialnych jednostek statystycznych oraz wykaz adresowy mieszkań w układzie poszczególnych rejonów i obwodów spisowych. Dobór prób kwartalnych odbywa się zgodnie z przyjętym schematem rotacyjnym⁶. Zaimplementowana metoda estymacji w BAEL polega na wyliczeniu odpowiednich mnożników uogólniających (wag). Proces ten przebiega w trzech etapach. W pierwszym oblicza się wagi pierwotne, które są odwrotnościami prawdopodobieństw wyboru do próby poszczególnych mieszkań. Następnie wylicza się wagi wtórne, które korygują wagi pierwotne na podstawie informacji o niezrealizowanych wywiadach w mieszkaniach. Ostatnim krokiem do pozyskania wag finalnych jest dostosowanie wag wtórnych do bieżących szacunków demograficznych (GUS, 2018d). Działania te są tożsame z pro-

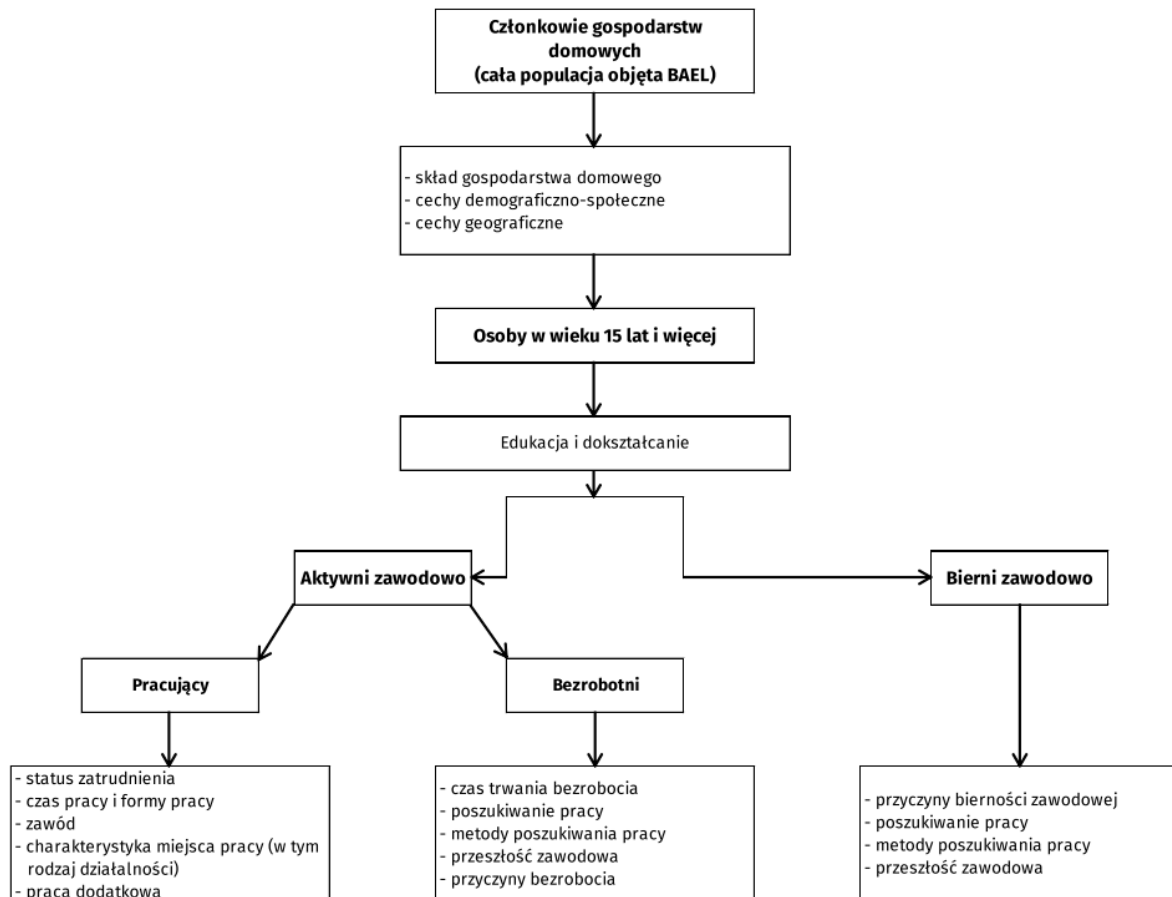
⁵Należy zaznaczyć, iż przyjęte ograniczenie populacji generalnej powoduje, iż liczba ludności w wieku 15 lat i więcej uzyskiwana po uogólnieniu wyników BAEL jest mniejsza od szacunków uzyskiwanych na podstawie bieżących bilansów ludności (która obejmuje między innymi również osoby bezdomne, czy też przebywające w gospodarstwach zbiorowego zakwaterowania).

⁶Jest on niezmienny od drugiego kwartału 1993 r. Obecnie jego szczegóły ujęte są w Karcie rotacji na lata 2016–2021.

cesem dwukrokowej kalibracji wag (Särndal i Lundström, 2005), co daje podstawy, aby estymator stosowany do otrzymania wyników w BAEL określać mianem kalibracyjnego. Tak też będzie on określany w niniejszej rozprawie. Szerszy opis przyjętej metodologii badania, w tym uogólniania jego wyników, znaleźć można w publikacjach GUS (2018a; 2018d), Wilaka (2017) oraz Popińskiego (2006).

2.1.4. Zakres informacyjny oraz ocena jakości badania

BAEL dostarcza obszernych danych, które umożliwiają ocenę bieżącej sytuacji na rynku pracy w cyklu kwartalnym. Poniższy schemat przedstawia ogólny zakres pozyskiwanych informacji dla poszczególnych zbiorowości wyodrębnianych w badaniu (GUS, 2018d).



Rysunek 2.1: Zakres informacji pozyskiwanych w BAEL dla poszczególnych zbiorowości

Źródło: Opracowanie na podstawie publikacji (GUS, 2018d).

Na podstawie uzyskanych w danym kwartale informacji w zakresie przedstawionym w powyższym schemacie (por.2.1) przeprowadza się proces estymacji, a jego wyniki następnie są publikowane w opracowaniu Aktywność ekonomiczna ludności Polski (GUS, 2018d). Dodatkowo w publikacji za ostatni kwartał roku dostępne są dane średnioroczne. Wyniki badania udostępniane są także w Dziejzinowej Bazie Wiedzy Rynek Pracy⁷, Banku Danych Lokalnych (BDL)⁸, bazie Strateg⁹ czy też Banku Danych Makroekonomicznych¹⁰. Nieodzownym elementem dotyczącym pozyskanych w badaniu informacji jest ocena ich jakości. Jednym z jej komponentów jest kompletność badania określana stopniem realizacji wylosowanej próby mieszkań. Podlega ona stałemu monitoringowi miesięcznemu oraz kwartalnemu na poziomie wojewódzkim i krajowym. Analizując ten element jakości brane są pod uwagę następujące wskaźniki: wskaźniki odpowiedzi, wskaźniki braku odpowiedzi, wskaźniki odmów. Dodatkowo analizowane są również przyczyny braku realizacji wywiadu w mieszkaniu. Wartości wskaźników odpowiedzi oraz jednostkowych braków odpowiedzi są uwzględniane w obliczaniu wag podczas uogólniania wyników badania na populację generalną. Warto zaznaczyć, że zbyt niski próg wskaźników odpowiedzi może wpłynąć negatywnie na precyzję wyników badania, która jest kluczowym składnikiem oceny szacowanych parametrów. Zaleca się w BAEL weryfikację jakości uzyskanych szacunków poprzez wyliczenie dla nich względnych błędów standardowych (ang. *Relative Standard Error* – RSE), które w literaturze przedmiotu określane są również jako współczynniki zmienności (ang. *Coefficient of Variation* – CV). Podczas interpretacji uzyskanych wyników BAEL należy mieć świadomość, że gdy konkretne oszacowanie było otrzymane na podstawie nielicznej reprezentacji w próbie może być ono obarczone dużym błędem losowym. W efekcie może się to przełożyć na niską precyzję estymacji. W takiej sytuacji informacje uzyskane na tej podstawie nie uprawniają badaczy do wyciągania wiążących wniosków o rozkładzie danej cechy w populacji. Do takich przypadków może dojść szczególnie podczas szacowania wyników dla niższych poziomów agregacji przestrzennej oraz grup osób określonych na podstawie wielu cech, tj. dla szczegółowo zdefiniowanych domen (GUS, 2018d). Więcej informacji na temat sposobu wyliczania wskaźników kompletności oraz ich wartości dla badania w 2017 roku znajduje się w zeszycie metodologicznym pt. Badanie Aktywności Ekonomicznej Ludności (GUS, 2018d). Natomiast zastosowane podejście do oceny ich

⁷Por. <http://swaid.stat.gov.pl/SitePagesDBW/RynekPracy.aspx>

⁸Por. <https://bdl.stat.gov.pl/BDL/start>

⁹Por. <https://strateg.stat.gov.pl/#/>

¹⁰Por. <https://bdm.stat.gov.pl/>

precyzji za pomocą metody bootstrap zostanie opisane w rozdziale czwartym niniejszej rozprawy.

W odniesieniu do publikowanych danych na podstawie BAEL europejska regulacja dotycząca badania siły roboczej zawiera następujące progi precyzji¹¹:

- dla grupy osób bezrobotnych, reprezentujących 5% ludności w wieku produkcyjnym, względny błąd standardowy przeciętnych rocznych oszacowań na poziomie NUTS 2¹² i nie może przekraczać 8% omawianej grupy ludności,
- dla grupy stanowiącej 5% ludności w wieku produkcyjnym, względny błąd standardowy na szczeblu krajowym dla oceny zmian między dwoma kolejnymi kwartałami nie może przekraczać 2% omawianej grupy ludności.

Opracowane wyniki badania po ich uogólnieniu podlegają też jeszcze dodatkowym kryteriom, które ograniczają ich udostępnianie tj. (GUS, 2018d):

- oszacowania udostępniane są z dokładnością do 1 tysiąca osób,
- przyjmuje się, iż oszacowania wynoszące poniżej 5 tys. osób uznawane są za mało wiarygodne,
- w przypadku uzyskania oszacowań poniżej 15 tys. osób zaleca się ostrożność w posługiwaniu się takimi danymi.

Należy mieć jednak na uwadze fakt, że powyższe wytyczne odnoszą się do danych z badania na poziomie kraju czy też województw, gdyż zgodnie z założeniem właśnie na tych poziomach agregacji przestrzennej udostępnia się wyniki badania w cyklu kwartalnym¹³. Powodem takiej sytuacji jest brak możliwości zapewnienia odpowiedniej jakości szacunku parametrów dla bardziej szczegółowych domen.

¹¹Rozporządzenia Rady nr 577/98 z dnia 9 marca 1998 r. w sprawie organizacji badania prób losowych dotyczącego siły roboczej we Wspólnocie.

¹²Do końca 2017 roku, poziom ten odpowiadał w Polsce podziałowi na województwa. Obecnie w wyniku czwartej rewizji NUTS 2016, wydzielono dodatkowy region z województwa mazowieckiego obejmujący Warszawę z jej obszerną otuliną.

¹³Od 2017 roku, ze względu na dokonaną modyfikację alokacji próby, dane średnioroczne udostępniane są również na poziomie miast wojewódzkich oraz dla obszaru województwa bez miasta wojewódzkiego. Otrzymujemy jednak w tych układach niepełny zakres informacji, gdyż występują tam w niektórych przypadkach braki danych w udostępnianych zestawieniach ze względu na zbyt duży błąd szacunku - <https://bdl.stat.gov.pl/BDL/dane/podgrup/temat>. Ten dodatkowy poziom dostępności nie wynika jednak z określonych wymogów międzynarodowych, a jedynie zgłaszanych potrzeb krajowych.

2.2. Administracyjne źródła danych

Obecnie służby statystyki publicznej w wielu krajach powszechnie pozyskują dane statystyczne z rejestrów administracyjnych (Wallgren i Wallgren, 2014). Jest to w głównej mierze związane z dążeniem do obniżenia kosztów prowadzenia badań oraz zmniejszenia obciążeń respondentów biorących w nich udział, a także ma na celu dostarczenie dodatkowych danych statystycznych na możliwie niskim poziomie agregacji przestrzennej. Również w Polsce rośnie znaczenie administracyjnych źródeł danych (AZD) w kontekście ich zastosowań¹⁴. Opis wybranych praktycznych obszarów zastosowań oraz potencjalnych możliwości ich wykorzystania można znaleźć między innymi w następujących opracowaniach (Chłoń-Domińczak i Ptak-Chmielewska, 2020; Filas-Przybył i Klimanek, 2019; Ryczkowski i Stopiński, 2018; Dehnel i Kowalewski, 2015; Dehnel, 2015; Roszka, 2012; Gołata i Dehnel, 2012; Józefowski i Młodak, 2009). Warto mieć jednak na uwadze fakt, że rejestry administracyjne są utrzymywane przez ich gestorów w innych celach niż statystyczne, dlatego ich użycie wymaga wcześniejszej oceny jakości tych zasobów (Daas i in., 2009; Józefowski i Rynarzewska-Pietrzak, 2010), a następnie przeprowadzenia całego procesu przekształcenia rejestrów administracyjnych w rejestry statystyczne¹⁵. Jest to duże wyzwanie metodologiczne i programistyczne, a sposób wykorzystania oraz metodyka przetwarzania poszczególnych zasobów zależy przede wszystkim od potrzeb statystycznych jakie mają one zaspokoić.

Jak wskazuje Piórek (2019), wykorzystanie rejestrów administracyjnych do zaspokojenia potrzeb informacyjnych na temat rynku pracy rekomendowane jest przez Eurostat oraz Międzynarodową Organizację Pracy (ang. *International Labour Organization* – ILO). Niewątpliwą zaletą tych zbiorów informacyjnych jest przede wszystkim zmniejszenie obciążeń sprawozdawców obowiązkami statystycznymi oraz ograniczenie kosztów badań. Dają one również szansę na pozyskanie nowych informacji, i co jest istotne w odniesieniu do tematyki rozprawy doktorskiej, umożliwiają prezentowanie danych na niższych poziomach agregacji przestrzennej oraz dla bardziej szczegółowo zdefiniowanych domen. Piórek (2019)

¹⁴Syntetyczny opis wykorzystania rejestrów urzędowych i systemów informacyjnych administracji publicznej w statystyce publicznej znajduje się pod adresem internetowym: <https://stat.gov.pl/badania-statystyczne/przekazywanie-danych-z-systemow/wykorzystanie-rejestrow-urzedowych-i-systemow-informacyjnych-administracji-publicznej-w-statystyce-publicznej>.

¹⁵Rejestr statystyczny jest rejestrem tworzonym do celów statystycznych. Powstaje on zazwyczaj poprzez przekształcanie danych z rejestrów i/lub innych administracyjnych źródeł danych - https://ec.europa.eu/eurostat/cros/content/statistical-register_en.

podkreśla jednak, iż korzystając ze źródeł administracyjnych należy mieć świadomość, iż gestorzy na podstawie utrzymywanych rejestrów realizują swoje cele, a nie cele statystyki publicznej. Istotną niedogodnością jest również brak pełnej spójności definicyjnej pojęć i zmiennych występujących w rejestrach z tymi, które funkcjonują w statystyce publicznej. Należy być też świadomym, iż w przypadku decyzji o wykorzystaniu rejestrów administracyjnych należy opracować metody ich transformacji i integracji danych, aby można je wykorzystać w celu zaspokajania potrzeb realizowanych przez statystykę publiczną. Należy się również liczyć z koniecznością dostosowania metodyki badań prowadzonych na bazie rejestrów administracyjnych do pojawiających się zmian prawnych, a odnoszących się do poszczególnych gestorów administracyjnych zbiorów danych i gromadzonych w tych zasobach informacji.

2.2.1. Rynek pracy w świetle rejestrów administracyjnych - prace eksperymentalne

Niniejszy podpunkt zawiera syntetyczny opis sposobów pozyskania informacji statystycznych pochodzących z rejestrów administracyjnych w odniesieniu do osób pracujących i bezrobotnych w układach rozważanych w rozprawie doktorskiej¹⁶ tj. liczby osób pracujących oraz bezrobotnych. Podejście to zostało wypracowane w ramach pracy badawczej, której wyniki szczegółowo opisano w pracy Śliwickiego i współautorów (2018)¹⁷.

Podstawowymi źródłami danych administracyjnych realizowanego projektu były zasoby:

- Zakładu Ubezpieczeń Społecznych (systemów KS i ZUS – Centralny Rejestr Płatników Składek, Centralny Rejestr Ubezpieczonych);
- Kasy Rolniczego Ubezpieczenia Społecznego (systemu nSIU).

Powyższe administracyjne zbiory danych zawierają informacje o osobach pracujących, a zbiory ZUS dodatkowo dane o osobach bezrobotnych, które pochodzą z powiatowych urzędów pracy. Dane w nich zawarte umożliwiają podział według płci, wieku czy też miejsca zamieszkania osób należących do tych dwóch kategorii. W celu uzupełnienia bra-

¹⁶W ramach resortu statystyki publicznej Urząd Statystyczny w Bydgoszczy posiada bogate doświadczenie w dziedzinie wykorzystania rejestrów administracyjnych dotyczących rynku pracy zob. Gierańczyk i Ryczkowski (2018); Zwara i in. (2014).

¹⁷Sam projekt miał szerszy zakres i dotyczył również kwestii dostarczenia informacji o pracujących według sekcji PKD oraz miar wynagrodzeń brutto.

kujących informacji korzystano również z rejestru Ministerstwa Finansów – Centralnego Rejestru Podmiotów Krajowej Ewidencji Podatników (CRP KEP) oraz Bazy Jednostek Statystycznych (BJS) prowadzonej przez GUS.

Brakujące wartości w docelowej bazie zawierającej dane pochodzące z rejestrów zostały uzupełnione za pomocą metody imputacji dedukcyjnej. Skorzystano również z technik imputacji opartej na rozkładzie prawdopodobieństwa oraz imputacji wielokrotnej¹⁸. Przyjęte rozwiązania pozwoliły na uzupełnienie braków na poziomie danych jednostkowych, a także agregatów. Wykorzystano do tego celu inne zmienne znajdujące się w już utworzonej bazie na podstawie danych z ZUS oraz KRUS, bądź też dostępnych zmiennych w innych systemach informacyjnych administracji publicznej (tj. CRP KEP oraz BJS). Konieczna była również korekta błędnych wartości (spoza dopuszczalnego zakresu oraz nietypowych), a także usunięcie występujących powtórzeń. Ostatecznie wypracowana metodyka przygotowania zbiorów danych, w tym cały proces ich harmonizacji (proces konieczny do przeprowadzenia prawidłowej integracji różnych źródeł danych¹⁹), umożliwiła prezentację danych i wskaźników odnoszących się do osób pracujących i bezrobotnych w zakładanych w projekcie przekrojach i dodatkowo według miejsca ich zamieszkania (Śliwicki i in., 2018).

Rysunek 2.2 obrazuje schemat procesu tworzenia docelowej bazy danych dla kategorii osób pracujących i bezrobotnych, która stanowiła podstawę do obliczenia poszczególnych wskaźników²⁰.

Warto zaznaczyć, iż uzyskane dane pozbawione są błędów losowych charakterystycznych dla badań reprezentacyjnych, gdyż informacje pochodzące z rejestrów administracyjnych można traktować jako badanie pełne. Z kolei błędy o charakterze nielosowym odnoszące się do badanych zbiorowości zostały w znacznej mierze usunięte w toku przeprowadzonych

¹⁸Podstawy matematyczne zastosowanych metod znaleźć można w publikacji Wesołowskiego i Tarczyńskiego (2016).

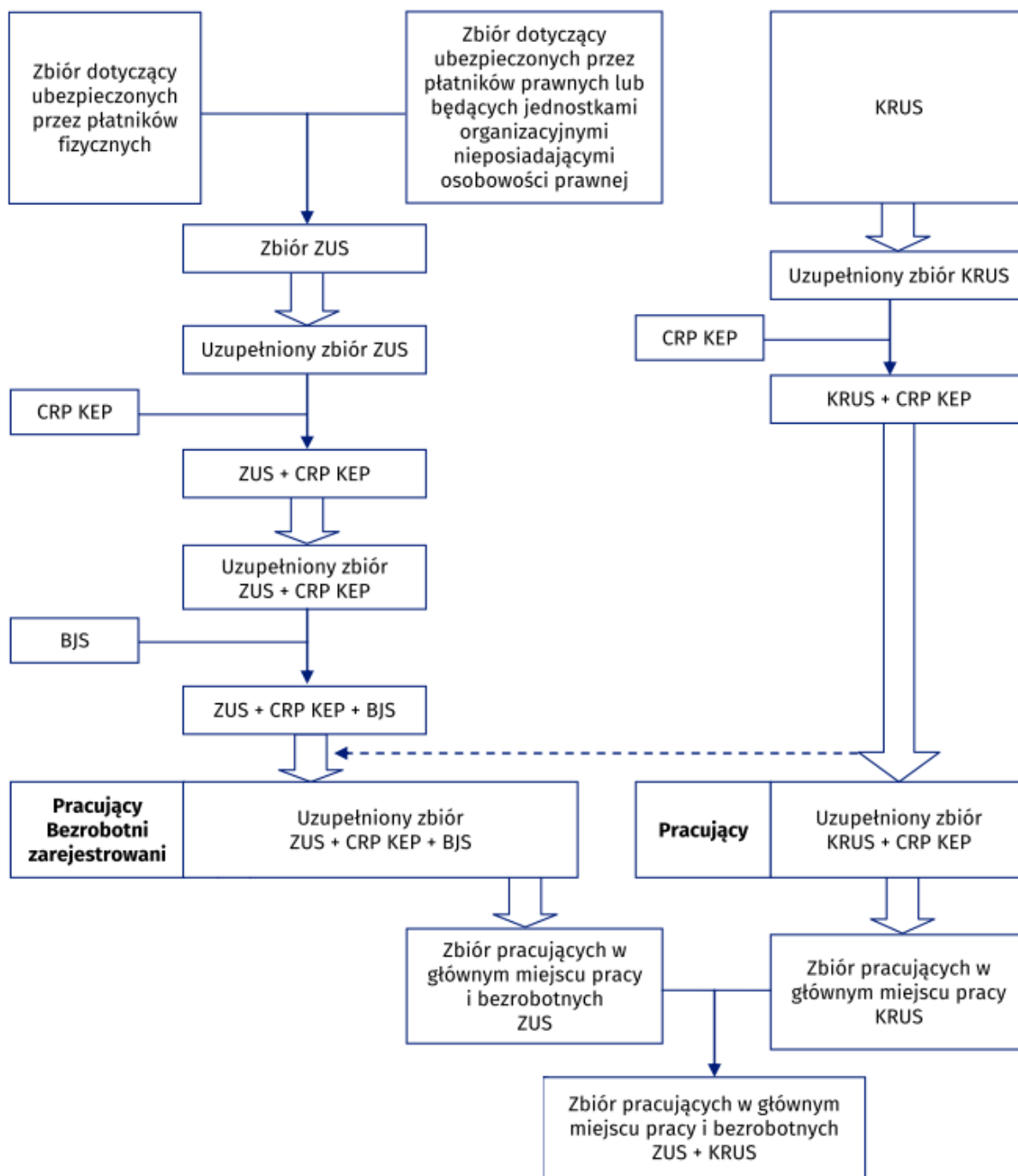
¹⁹Szerzej o etapach procesu harmonizacji por. Roszka (2013).

²⁰Szczegółowy proces przygotowania zbiorów do ostatecznych naliczeń docelowych zmiennych znajduje się w załączniku 34 omawianej pracy badawczej.

prac²¹. Przedstawione wyniki mają jednak wymiar eksperymentalny, a metodyka ich pozyskiwania jest doskonała w Urzędzie Statystycznym w Bydgoszczy²².

²¹Przyjęty szczegółowy opis rozwiązań, również w odniesieniu do błędów nielosowych, przedstawiony jest w szerszym wymiarze w raporcie końcowym z przeprowadzonej pracy badawczej oraz dołączonych do niej załącznikach.

²²Wybrane wyniki prac można znaleźć na stronie internetowej <https://esrp.stat.gov.pl/ESRP/>. Z kolei opis wdrożonego podejścia do wyznaczania liczby osób pracujących w gospodarce narodowej według źródeł administracyjnych znajduje się w zeszycie metodologicznym (GUS, 2020c).



Rysunek 2.2: Przebieg procedury przygotowującej zbiory liczby osób pracujących oraz bezrobotnych w głównym miejscu pracy na podstawie AZD

Źródło: Śliwicki i in. (2018).

Zgodnie z wypracowaną metodyką w ramach projektu zostały również przygotowane przez pracowników Urzędu Statystycznego w Bydgoszczy tabele z danymi zagregowanymi na poziomie gmin dla dwóch kategorii rynku pracy, tj. osób pracujących i bezrobotnych w podziale na płeć oraz przyjęte grupy wieku dla poszczególnych kwartałów 2017 roku. Po ich

zintegrowaniu w jeden zbiór reprezentujący dane średnioroczne i dla wyższych poziomów agregacji przestrzennej otrzymano bazę danych, którą wykorzystano w przeprowadzonym w ramach pracy doktorskiej badaniu empirycznym. W rozprawie jest ona określana jako *Zintegrowany rejestr osób pracujących i bezrobotnych - ZROPiB*²³.

2.2.2. Definicje podstawowych kategorii rynku pracy

Na potrzeby realizacji niniejszej pracy badawczej dostosowano²⁴ analizowaną zbiorowość osób pracujących do obowiązującej w statystyce publicznej definicji osób pracujących w gospodarce narodowej, która znajduje zastosowanie w badaniach przedsiębiorstw i jednostek sfery budżetowej. Jest ona jednak różna od obowiązującej w badaniu BAEL co jest istotne z punktu widzenia możliwości wykorzystania pozyskanych w oparciu o nią informacji.

Do osób pracujących zaliczono (GUS, 2018c):

- 1) osoby zatrudnione na podstawie stosunku pracy (umowa o pracę, powołanie, mianowanie, wybór lub stosunek służbowy);
- 2) pracodawców i pracujących na własny rachunek, czyli:
 - a) właścicieli, współwłaścicieli i dzierżawców gospodarstw indywidualnych w rolnictwie (łącznie z pomagającymi członkami ich rodzin),
 - b) właścicieli i współwłaścicieli (łącznie z pomagającymi członkami ich rodzin; z wyłączeniem wspólników spółek, którzy nie pracują w spółce) podmiotów prowadzących działalność gospodarczą poza gospodarstwami indywidualnymi w rolnictwie,
 - c) inne osoby pracujące na własny rachunek, np. osoby wykonujące wolne zawody;
- 3) osoby wykonujące pracę nakładczą;
- 4) agentów (łącznie z pomagającymi członkami ich rodzin oraz osobami zatrudnionymi przez agentów);
- 5) członków spółdzielni produkcji rolniczej (rolniczych spółdzielni produkcyjnych i spółdzielni powstałych na ich bazie oraz spółdzielni kółek rolniczych);

²³Wybrane wyniki prac wykorzystujące to źródło danych były prezentowane przez autora niniejszej rozprawy między innymi na konferencji „230 Lat Statystyki Publicznej w Polsce” – <https://konferencja230.stat.gov.pl>.

²⁴Poza nieistotnymi z punktu widzenia rozprawy różnicami opisanymi w raporcie Śliwickiego i in. (2018) oraz GUS (2020c).

6) duchownych pełniących obowiązki duszpasterskie.

Drugą rozważaną kategorią rynku pracy były osoby bezrobotne. W tym przypadku przyjęto definicję osoby bezrobotnej uregulowaną zapisami ustawy o promocji zatrudnienia i instytucjach rynku pracy²⁵. Takie podejście było możliwe dzięki temu, że zarówno rejestr ZUS jak i badanie prowadzone w statystyce publicznej (Badanie bezrobocia rejestrowanego²⁶), wykorzystują tę samą definicję osoby bezrobotnej. Zgodnie z nią, do osób bezrobotnych zalicza się te osoby, które ukończyły 18 lat i nie osiągnęły wieku emerytalnego, a są niezatrudnione i nie wykonują innej pracy zarobkowej. Jednocześnie są zdolne i gotowe do podjęcia zatrudnienia w pełnym wymiarze czasu pracy. Kluczowy jest tutaj też fakt zarejestrowania we właściwym dla miejsca zameldowania (stałego lub czasowego) powiatowym urzędzie pracy oraz poszukiwania zatrudnienia lub innej pracy zarobkowej. Dokładny zakres włączeń osób w poczet tej zbiorowości uregulowany jest w ustawie z 2004 roku.

2.3. Badanie bezrobocia rejestrowanego

Główny Urząd Statystyczny od wielu lat korzysta z administracyjnych źródeł danych dostarczając statystyk na temat bezrobocia rejestrowanego. Badanie bezrobocia rejestrowanego (BR²⁷) wykorzystuje sprawozdania MRPiPS-01 wraz z załącznikami, MRPiPS-02 wraz z załącznikiem oraz MRPiPS-07, które są dostarczane przez powiatowe i wojewódzkie urzędy pracy. Jak wspomniano w poprzednim punkcie, definicja osób objętych badaniem bezrobocia rejestrowanego jest taka sama jak przytoczona w punkcie 2.2.2. Badanie realizowane jest we współpracy z Ministerstwem Rodziny, Pracy i Polityki Społecznej (MRPiPS). Pozyskane informacje umożliwiają dostarczenie danych o liczbie i strukturze osób bezrobotnych między innymi według (Rada Ministrów, 2016):

- cech demograficzno-społecznych (wiek, płeć, poziom wykształcenia);
- miejsca zamieszkania (miasto, wieś);

²⁵Ustawa z dnia 20 kwietnia 2004 r. o promocji zatrudnienia i instytucjach rynku pracy (Dz.U. 2004 nr 99, poz. 1001).

²⁶Więcej informacji o tym badaniu znajduje się w punkcie 2.3.

²⁷Skrót przyjęty w niniejszej rozprawie stosowany zamiennie do określenia badania pt. *Bezrobocie rejestrowane*.

- pozostawania w szczególnej sytuacji na rynku pracy (np. osoby do 30 roku życia; bezrobotni długotrwale, osoby powyżej 50. roku życia, korzystający ze świadczeń z pomocy społecznej).

Istotną wartością opracowywanych informacji jest to, iż dostępne są one na poziomie powiatów, a w pewnej części również na poziomie gmin. Warto też podkreślić, że podstawowe statystyki dotyczące wielkości i natężenia bezrobocia rejestrowanego dostępne są z częstotliwością miesięczną. Następnie uzupełniane są one kwartalnie lub półrocznie o dodatkowe informacje takie jak cechy demograficzno-społeczne oraz charakterystykę sytuacji zawodowej poprzedzającej rejestrację osoby jako bezrobotnej (GUS, 2019c).

Publikacja na temat bezrobocia rejestrowanego zatytułowana *Bezrobocie rejestrowane* (GUS, 2018b) wydawana jest przez GUS co kwartał. Informacje z tego zakresu udostępniane są również w Banku Danych Lokalnych, bazie Strateg czy też w Banku Danych Makroekonomicznych.

Odnosząc się do celu rozprawy doktorskiej, do realizacji której potrzebne są informacje o wszystkich kategoriach rynku pracy, dane z tego badania dostępne są jedynie dla jednej z analizowanych kategorii tj. osób bezrobotnych i w dodatku nie we wszystkich zdefiniowanych w rozprawie przekrojach. Informacja o liczbie osób bezrobotnych w podziale na płeć z uwzględnieniem ustalonych grup wieku (innych niż przyjęte w niniejszej rozprawie) jest dostępna kwartalnie jedynie na poziomie powiatów. Z kolei informacja na poziomie gmin jest udostępniana jedynie w podziale na płeć lub grupy wieku w cyklu półrocznym.

Warto również w tym miejscu podkreślić, że zmiany w prawie krajowym wpływają na zmiany populacji osób, które mogą być uznane za bezrobotnych. Z tego powodu dane te tylko w ograniczonym stopniu mogą służyć do porównań międzynarodowych (GUS, 2019c).

2.4. Badania przedsiębiorstw i jednostek sfery budżetowej

Badania przedsiębiorstw i jednostek sfery budżetowej prowadzone są w formie badań pełnych lub reprezentacyjnych. Realizowane są z częstotliwością miesięczną, kwartalną,

roczną albo jako badania cykliczne co 2 lub 4 lata²⁸. Z punktu widzenia celu rozprawy doktorskiej najistotniejszą rolę odgrywa badanie pt. *Pracujący w gospodarce narodowej*.

2.4.1. Pracujący w gospodarce narodowej

Pracujący w gospodarce narodowej (PwGN²⁹) jest badaniem stałym realizowanym raz w roku. Operat do badania, utworzony jest na podstawie Bazy Jednostek Statystycznych i jest aktualizowany w oparciu o dane pochodzące z Krajowego Rejestru Urzędowego Podmiotów Gospodarki Narodowej (REGON). Operat stanowią jednostki wybierane do badania z BJS na podstawie wybranych cech takich jak: klasa wielkości przedsiębiorstwa mierzona liczbą pracujących, przeważający rodzaj działalności, forma własności itp. Na podstawie operatu tworzone są kartoteki do poszczególnych badań, według których to badanie jest realizowane. Zawierają one listę jednostek zobligowanych do przekazania danych na formularzach statystycznych zgodnie z zakresem podmiotowym poszczególnych badań.

Źródłami danych w tym przypadku są *Sprawozdanie o zatrudnieniu i wynagrodzeniach* (Z-03), *Sprawozdanie o pracujących, wynagrodzeniach i czasie pracy* (Z-06), *Meldunek o działalności gospodarczej* (DG-1), *Sprawozdanie o działalności gospodarczej przedsiębiorstw* (SP-3)³⁰. Realizowane są one w formie badań pełnych lub reprezentacyjnych o częstotliwości miesięcznej, kwartalnej lub rocznej. Uzupełnienie powyższych badań stanowią informacje pozyskane o osobach pracujących z Powszechnego Spisu Rolnego 2010 oraz z rejestrów prowadzonych przez Ministerstwo Obrony Narodowej oraz Ministerstwo Spraw Wewnętrznych i Administracji, Instytut Statystyki Kościoła Katolickiego, Ankiety Wyznaniowej GUS, a także szacunków GUS. Dodatkowo wykorzystuje się również opracowane na podstawie danych administracyjnych szacunki liczby osób pracujących w fundacjach, stowarzyszeniach i innych organizacjach społecznych (GUS, 2020b).

²⁸Więcej informacji na temat zakresu prowadzonych badań znaleźć można w opracowaniach GUS (2018e; 2020d).

²⁹Skrót przyjęty w niniejszej rozprawie stosowany zamiennie do określenia badania *Pracujący w gospodarce narodowej*.

³⁰Więcej informacji na temat wymienionych badań znaleźć można w pracach GUS (2019d; 2019e; 2020d).

Dane na temat pracujących w całej gospodarce narodowej³¹ dostępne są jedynie z częstotliwością roczną według stanu na dzień 31 grudnia. Jednocześnie w badaniu tym w odniesieniu do analizowanych w rozprawie doktorskiej cech i przekrojów dostępna jest tylko informacja o pracujących do poziomu województw w podziale na płeć. Warto zaznaczyć, że w tym badaniu osoby pracujące klasyfikowane są według faktycznego miejsca wykonywania pracy, a nie według miejsca zamieszkania.

2.5. Spisy powszechne

Jak pisze Gołata (2018): „Spis ludności jest najstarszym badaniem statystycznym o powszechnym zasięgu, przeprowadzanym w zasadzie we wszystkich państwach świata. Stąd jest zrozumiałe, że badanie to bywa różnie definiowane i w różny sposób przeprowadzane. Zarówno metody przeprowadzania spisów, jak i związane z nimi tradycje i zwyczaje były kształtowane przez wieki i ulegały wielu zmianom.”. W Polsce metodologia jego przeprowadzania również przez lata ewoluowała. Szczególnie istotne zmiany nastąpiły w Narodowym Spisie Powszechnym Ludności i Mieszkań 2011 (NSP 2011)³² kiedy to odstąpiono od podejścia tradycyjnego opartego jedynie na informacjach pozyskanych od rachmistrzów spisowych zbieranych na formularzach w formie papierowej, na rzecz tzw. metody mieszanej, tj. w ramach której informacje były pozyskane z rejestrów administracyjnych oraz badania reprezentacyjnego.

2.5.1. NSP 2011

Spis zrealizowany w 2011 roku składał się z badania pełnego opartego na rejestrach administracyjnych oraz badania reprezentacyjnego, które swoim zakresem obejmowało również zmienne dotyczące rynku pracy. Badanie częściowe dostarczyło pożądanych informacji o respondentach, których nie można było uzyskać na podstawie rejestrów i systemów informacyjnych. Próba losowa wynosiła około 20% mieszkań w kraju. Jednostką losowa-

³¹Należy zaznaczyć, iż dane te zawierają również informacje o osobach pracujących w podmiotach gospodarczych o liczbie pracujących do 9 osób, (które są objęte badaniem na formularzu SP-3) oraz osobach duchownych i pracujących w jednostkach budżetowych prowadzących działalność w zakresie obrony narodowej i bezpieczeństwa publicznego.

³²Do spisów powszechnych zalicza się także spisy rolne, które w swoim zakresie również podejmują tematykę pracujących w rolnictwie. Jednak ze względu na cel realizowanej pracy zostały one pominięte jako potencjalne źródło informacji.

nia było mieszkanie, a ich wykaz został przygotowany w postaci odpowiedniego operatu losowania z rozbudowanym warstwowaniem. Zastosowany został schemat losowania jednostopniowego, który zakładał odpowiednią alokację próby w poszczególnych powiatach (GUS, 2014)³³. Zastosowanie metody mieszanej miało przede wszystkim zmniejszyć koszty spisu oraz obciążenie respondentów, przy jednoczesnym zachowaniu dobrej jakości wyników.

Opracowując publikacje z NSP 2011, wskazywano iż interpretując jego wyniki oparte na badaniu reprezentacyjnym zaleca się przyjęcie następujących progów precyzji, które świadczą o jakości uzyskanych wyników estymacji (GUS, 2014):

- $CV \leq 10\%$ - akceptowalny poziom błędu szacunku;
- $10\% < CV \leq 20\%$ - powinny być traktowane z ostrożnością, jako mało precyzyjne;
- $CV > 20\%$ - przekroczenie tego progu uznaje się za wyraźny sygnał niewiarygodności uzyskanych oszacowań.

Ostatecznie wykorzystanie estymatora kalibracyjnego³⁴ do uogólniania wyników z badania reprezentacyjnego towarzyszącego spisowi umożliwiło uzyskanie informacji spisowej jedynie do poziomu powiatu. Jednocześnie należy podkreślić, że taki poziom agregacji nie był możliwy dla wszystkich kategorii zbieranych zmiennych. Tak było m.in. w przypadku zmiennych dotyczących rynku pracy. Dane o pracujących, bezrobotnych oraz biernych zawodowo były dostarczone na poziomie powiatów jedynie w podziale na płeć. Charakter domen, w ramach których dokonywano szacunku, uzależniony był w głównej mierze od zrealizowanej wielkości próby w danym przekroju, która przełożyła się na wielkość błędu szacunku. Przyjęcie zbyt szczegółowego przekroju prowadziło do błędów na nieakceptowalnym poziomie.

2.5.2. NSP 2021

Ostatni spis powszechny został przeprowadzony w 2021 roku (NSP 2021). W odniesieniu do osób objął swoim zakresem Polaków mieszkających w Polsce oraz cudzoziemców mieszkających w naszym kraju na stałe lub przebywających czasowo, w dniu 31.03.2021 roku³⁵.

³³Szersze informacje na temat metodologii NSP 2011, w tym sposobu przeprowadzenia badania reprezentacyjnego i oceny jego jakości, można znaleźć w publikacjach GUS (2012; 2013).

³⁴Szczegóły zastosowanego podejścia opisane są w artykule Szymkowiaka (2014).

³⁵Nie jest przy tym istotne czy osoby te posiadają zameldowanie w Polsce.

Dodatkowo obejmował również Polaków przebywających za granicą bez względu na okres przebywania, którzy nie wymeldowali się z pobytu stałego w Polsce w związku z wyjazdem na stałe za granicę. Zakres spisu obejmował również osoby bezdomne. Składał się on w całości z badania pełnego i nie zaplanowano osobnego badania reprezentacyjnego jak to miało miejsce w NSP 2011. Główny nacisk został położony na samospis internetowy, ale informacje były również pozyskiwane poprzez wywiad telefoniczny oraz rachmistrzów terenowych.

Jeden z modułów ankiety spisowej odnosił się do aktywności ekonomicznej ludności i na jego podstawie możliwe będzie pozyskanie informacji między innymi o liczbie osób pracujących, bezrobotnych oraz biernych zawodowo w przewidzianych przekrojach (GUS, 2021). Obecnie trwają prace nad opracowaniem danych z NSP 2021, którego wyniki sukcesywnie mają się pojawiać od początku 2022 roku.

Oceniając spis NSP 2011 z punktu widzenia celu badania przyjętego w pracy można stwierdzić, że jako źródło danych nie stwarza on możliwości dostarczenia informacji na temat rynku pracy w przekroju miejskich obszarów funkcjonalnych. Wynika to przede wszystkim z faktu, że dane takie publikowane były na poziomie powiatów. Ważne jest również to, że analizy prowadzone w pracy dotyczyły 2017 r. W związku z tym w charakterze danych wejściowych do estymacji typu SPREE bardziej zasadne było uwzględnienie danych administracyjnych, których okres referencyjnych odnosił się również do 2017 r.

Z kolei spis 2021 jest obecnie na etapie opracowywania wyników i dopiero po ich opublikowaniu i ocenie jakości uzyskanych wyników będzie można ocenić to źródło pod kątem możliwości jego wykorzystania w odniesieniu do miejskich obszarów funkcjonalnych.

2.6. Możliwości i ograniczenia poszczególnych źródeł danych - analiza porównawcza

Na podstawie przeprowadzonej kwerendy opisanych źródeł danych, można zauważyć, iż każde z nich ma swoje ograniczenia z punktu widzenia wymaganej w pracy dostępności pożądanych zmiennych na poziomie miejskich obszarów funkcjonalnych. Jest to spowo-

dowane albo brakiem zgodności definicyjnej z wymogami międzynarodowymi bądź też poziomem dostępności danych.

Głównym i najbardziej aktualnym źródłem danych o wszystkich podstawowych kategoriach rynku pracy (pracujący, bezrobotni, bierni zawodowo) jest Badanie Aktywności Ekonomicznej Ludności. Zawiera ono szeroki wachlarz informacji w różnym ujęciu (demograficzno-społecznym czy ekonomicznym) na temat rynku pracy. Definicje analizowanych kategorii rynku pracy są również zgodne z definicjami zalecanymi przez Międzynarodową Organizację Pracy. Bardzo ważne i niezwykle użyteczne jest to, iż informacje pochodzące z tego źródła udostępniane są w cyklu kwartalnym i rocznym w publikacjach, informacjach sygnałnych oraz bazach danych. Jednak kluczowym ograniczeniem użyteczności BAEL-u jest brak informacji dla nietypowych przekrojów terytorialnych, takich jak miejskie obszary funkcjonalne. Wynika to z reprezentacyjnego charakteru badania, które ze względu na założoną wielkość próby i przyjęty schemat losowania nie jest w stanie zapewnić wiarygodnych oszacowań dla bardziej szczegółowych podziałów analizowanej populacji. Przyczyną tej sytuacji jest za mała liczba obserwacji dla jednostek wyróżnionych w zbyt szczegółowych przekrojach co prowadzi do dużych błędów szacunków.

Ewentualna modyfikacja badania w taki sposób, aby uzyskać szacunki na niższym poziomie agregacji wiązałyby się ze znacznym zwiększeniem jego kosztów oraz obciążeń respondentów. Poza tym, ze względu na zastosowany w BAEL-u panel rotacyjny, uzyskanie rzetelnych wyników możliwe byłoby dopiero po około 2 latach³⁶. Wiadomo jednak, jak ważne są słowa przytaczane często przez Johna Rao, zaczerpnięte z publikacji Fullera (1999), że „użytkownicy danych statystycznych oczekują zawsze więcej, niż zostało to ustalone na etapie projektowania badania”. Rzeczą niezwykle trudną, a w zasadzie niemożliwą, jest dokonywanie częstych zmian w schemacie losowania próby, aby dostosować go do dynamicznie zmieniających się potrzeb użytkowników informacji statystycznej.

Drugim analizowanym i jednocześnie niezwykle obiecującym źródłem informacji o osobach pracujących i bezrobotnych o charakterze innowacyjnym, jest *Zintegrowany rejestr osób pracujących i bezrobotnych*. Jego niewątpliwą zaletą jest możliwość dostarczenia informacji na temat wybranych charakterystyk rynku pracy nawet na poziomie gminy. Co więcej, zakładając że metodyka badania zostanie zgodnie z założeniami w przyszłości udoskonalona, informacje dostępne będą na jeszcze niższym poziomie agregacji. Na uwagę zasługuje

³⁶Przykład takiej sytuacji został przytoczony podczas opisu BAEL.

również fakt, że dane znajdujące się w tym zbiorze odnoszą się do miejsca zamieszkania osób pracujących. Stanowi to niewątpliwie dużą zaletę zwłaszcza w porównaniu z wynikami badania *Pracujący w gospodarce narodowej*, gdzie informacje o osobach pracujących podawane są według faktycznego miejsca wykonywania pracy. Również informacje o osobach bezrobotnych w ZROPiB są ujmowane według miejsca zamieszkania, co ułatwia wspólne analizowanie tych dwóch kategorii zmiennych. Definicja osoby bezrobotnej jest tożsama z obowiązującą w badaniu *Bezrobocie rejestrowane* prowadzonym przez GUS.

Warto podkreślić, że pomimo niewątpliwych zalet tego źródła informacji, jakimi są możliwość pozyskiwania bardzo dobrej jakości informacji o częstotliwości kwartalnej i brak błędów spotykanych w badaniach reprezentacyjnych o charakterze losowym, ma ono pewne ograniczenia. Jednym z nich jest nieuwzględnianie osób pracujących w szarej strefie³⁷, które zaliczane są do pracujących w BAEL. Dlatego również w przypadku tego źródła danych, pomimo dostarczania informacji łącznej o dwóch kategoriach rynku pracy, definicje zbiorowości osób pracujących oraz bezrobotnych nie są zgodne ze stosowanymi przez Międzynarodową Organizację Pracy.

Różnice pojawiają się też w definicji osób pracujących stosowanej w badaniu *Pracujący w gospodarce narodowej*, w której osoby pracujące są ujmowane według faktycznego miejsca pracy, a nie miejsca zamieszkania, co jeszcze bardziej odbiega od definicji zalecanej przez MOP.

Problem rozbieżności definicyjnych, który nie dotyczy tylko wyżej opisanych źródeł danych, jest jednym z najpoważniejszych i najtrudniejszych do przezwyciężenia. Różnice w definicjach mogą prowadzić do nieprawidłowej interpretacji wyników prowadzonych analiz (Markowicz, 2015, 2017). W BAEL do kategorii pracujący nie są zaliczane osoby pracujące i mieszkające w hotelach pracowniczych oraz pracujące za granicą na rzecz polskich pracodawców, jak to ma miejsce w badaniu *Pracujący w gospodarce narodowej*, w którym uzyskujemy dane na koniec okresu (stan na 31 XII). Natomiast BAEL przedstawia dane średnie dla badanego okresu. Więcej informacji na temat różnic między tymi badaniami można znaleźć w publikacjach GUS (2019c, 2018e).

Należy również zwrócić uwagę na to, że informacja o liczbie osób pracujących uzyskiwana w badaniu *Pracujący w gospodarce narodowej* powstaje na podstawie danych zbieranych

³⁷Więcej informacji o tej kategorii pracujących znaleźć można w publikacji GUS (2019a).

w ramach kilku badań prowadzonych przez statystykę publiczną, z których część ma charakter reprezentacyjny, a część pełny³⁸. Może to wpłynąć na zakres wykorzystania danych. O ograniczonej użyteczności źródła, zwłaszcza z punktu widzenia niniejszej pracy, świadczy również fakt, że dostarcza ono informacji jedynie na temat jednej kategorii rynku pracy tj. osób pracujących. Warto ponadto raz jeszcze podkreślić, iż w porównaniu do analizowanych w rozprawie doktorskiej przekrojów (miejskie obszary funkcjonalne w podziale na płeć i grupy wieku) dane z tego badania dostępne są na koniec każdego roku i to jedynie do poziomu województw w podziale na płeć. Trudności z dostępem do odpowiedniej informacji w tych przekrojach podnoszone są między innymi w opracowaniu Kłosowskiej i in. (2013).

Kolejnym rozważanym źródłem informacji na temat rynku pracy, które należy ocenić pod kątem jego użyteczności w odniesieniu do realizowanego celu rozprawy doktorskiej, jest badanie *Bezrobocie rejestrowane*. W źródle tym informacje o osobach bezrobotnych dostępne są nawet na poziomie gmin w podziale na niektóre cechy demograficzno-społeczne. Ze względu na dostępność danych na niskim poziomie agregacji przestrzennej informacja ta jest często wykorzystywana przez władze szczebla samorządowego i krajowego. Ważne jest również, że dane dotyczące wielkości bezrobocia rejestrowanego dostępne są co miesiąc, a w szerszym wymiarze cech, dodatkowo kwartalnie lub półrocznie. Dlatego też, dane te wykorzystywane są często również do prowadzenia pogłębionych analiz (por. Bieszk-Stolorz (2017); Bieszk-Stolorz i Dmytrów (2019); Kwiatkowski i Kwiatkowska (2020)). Jednak pomimo wymienionych zalet tego źródła nie jest ono wolne od ograniczeń. W odróżnieniu od spisów czy BAEL-a, badanie to dostarcza informacji tylko o jednej kategorii rynku pracy tj. osobach bezrobotnych. Dodatkowo ze względu na przyjętą i przytoczoną wcześniej definicję osoby bezrobotnej zarejestrowanej w urzędzie pracy, nie powinno się jej wykorzystywać do porównań międzynarodowych. Jest ona bowiem ściśle związana z regulacjami prawnymi obowiązującymi w naszym kraju. Więcej szczegółów na temat różnic pomiędzy obowiązującą w tym źródle definicją, a przyjętą w BAEL i wynikających z tego konsekwencji, można znaleźć w opracowaniach Głuszczyk i Raszkowskiego (2016), Madras-Kobus i Rogowskiego (2013) oraz Janukowicza (2010).

³⁸Więcej informacji na temat oceny jakości poszczególnych badań z tego obszaru znajduje się w publikacji GUS (2020b).

Ostatnimi analizowanymi w niniejszym podrozdziale źródłami danych ważnymi z punktu widzenia możliwości wykorzystania w pracy doktorskiej są spisy powszechne NSP 2011 i 2021. Niewątpliwą zaletą NSP 2011 jest to, że dostarcza informacji o sytuacji na rynku pracy w podziale na jej podstawowe kategorie zgodne z definicjami zalecanymi przez Międzynarodową Organizację Pracy³⁹. Głównym ograniczeniem natomiast jest bardzo niska częstotliwość badania. Wpływa to niewątpliwie na jego malejącą wartość informacyjną w okresach między spisowych. Dodatkowo, w NSP 2011 brak jest wielu kluczowych informacji na temat sytuacji na rynku pracy poniżej poziomu powiatu ze względu na przyjęty sposób pozyskiwania tych danych (badanie reprezentacyjne)⁴⁰. Z kolei dane ze spisu 2021 nie są obecnie jeszcze dostępne i na moment pisanie dysertacji trudno jest kompleksowo ocenić możliwości ich wykorzystania. Istotne jednak jest również to, że zarówno NSP 2011 jak i NSP 2022 mają inne okresy referencyjne niż przyjęty w pracy 2017 rok. Ma to niewątpliwie wpływ na aktualność pozyskanych z nich informacji w odniesieniu do przyjętego w niniejszej rozprawie okresu.

Warto na koniec tej części analizy porównawczej podkreślić, że obecnie jedynie dwa źródła, tj. BAEL oraz NSP 2011, zawierają łączną informację o wszystkich kategoriach rynku pracy tj. pracujących, bezrobotnych oraz biernych zawodowo. Wydaje się to kluczowe w zakresie kompleksowej analizy rynku pracy, opartej na jednolitej metodologii wyznaczania poszczególnych kategorii rynku pracy. Dodatkowo należy podkreślić, iż w żadnym innym źródle, poza tymi dwoma, nie można uzyskać informacji o osobach biernych zawodowo.

Charakteryzując zasoby ludzkie z punktu widzenia rynku pracy, większość autorów koncentruje się przede wszystkim na zasobach osób pracujących i bezrobotnych oraz odpowiadających tym zbiorowościom wskaźnikach poświęcając tym samym niewiele uwagi osobom biernym zawodowo. Osoby pozostające poza rynkiem pracy są jednak bardzo zróżnicowaną zbiorowością i z pewnością wartą głębszej analizy. Zbiorowość ta obejmuje zarówno osoby, które na rynek pracy dopiero wejdą, jak i te, które definitywnie zakończyły karierę zawodową. Do pierwszej podzbiorowości należą głównie osoby młode, które jeszcze nie rozpoczęły poszukiwania pracy oraz osoby, które przez jakiś czas pozostają nieaktywne, np. z powodu macierzyństwa. Jednostki należące do tej podzbiorowości mogą jednak w krót-

³⁹Pomimo przyjęcia tych samych definicji dla kategorii rynku pracy w BAEL oraz NSP 2011 istnieją pomiędzy nimi pewne różnice (Dolecka i Raczkiewicz, 2014; Golata i Dehnel, 2013).

⁴⁰Wcześniejsze rundy spisów ludności dostarczały informacji z tego obszaru do poziomu gmin.

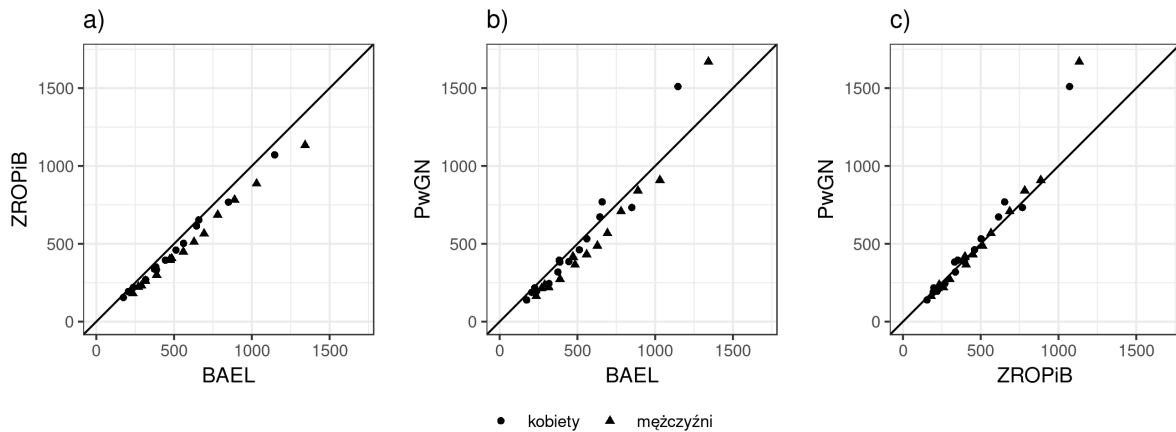
kim czasie przejść do jednej z pozostałych kategorii rynku pracy zmieniając ich liczebność (GUS, 2019c). Dlatego też informacja o tej kategorii osób nie powinna pozostać poza zakresem zainteresowań badaczy.

Podsumowując powyższe rozważania można stwierdzić, że żadne z dostępnych w Polsce i omówionych źródeł danych nie dostarcza pożądaných informacji o poszczególnych kategoriach rynku pracy w odniesieniu do miejskich obszarów funkcjonalnych w podziale na płeć oraz grupy wieku i zgodnie z definicjami zalecanymi przez MOP. Uniemożliwia to prowadzenie pogłębionych analiz na temat rynku pracy dla nietypowych przekrojów terytorialnych jakimi niewątpliwie są miejskie obszary funkcjonalne czy też obszary specjalnych stref ekonomicznych⁴¹. Biorąc jednak pod uwagę zgłaszane przez odbiorców danych zapotrzebowanie na tego rodzaju informacje warto jest podjąć próbę oszacowania liczebności odpowiednich kategorii charakteryzujących rynek pracy na tak określonym poziomie agregacji przestrzennej.

W uzupełnieniu do przeprowadzonej analizy porównawczej na wykresach rozrzutu przedstawiono porównanie wyników omówionych powyżej badań, według wybranych kategorii osób na rynku pracy (por. rysunki 2.3 i 2.4). Pozwala to na prześledzenie nie tylko różnic, ale i podobieństw w poziomach zmiennych⁴². Ze względu na ograniczenia w zakresie poziomu dostępności danych w ramach poszczególnych źródeł analiza porównawcza dokonana została na poziomie województw w przekroju płci.

⁴¹W tym drugim przypadku obszar przez nie zajmowany istotnie różni się pomiędzy poszczególnymi strefami i trudno go utożsamiać z poziomem gmin, czy też powiatów. Również zasięg ich oddziaływania nie mieści się w żadnych granicach wyznaczonych przez podział administracyjny. Jednak pomimo ograniczeń w dostępie do adekwatnych danych starano się prowadzić analizy w tym obszarze (Józefowski i Szymkowiak, 2015; Cicha-Nazarczuk i Nazarczuk, 2017; Józefowski i Młodak, 2017).

⁴²Ponieważ informacje o osobach biernych zawodowo dostępne są jedynie w dwóch źródłach danych tj. w BAEL oraz NSP 2011 do porównań wzięto dwie pozostałe kategorie rynku pracy tj. osoby pracujące oraz bezrobotne. Natomiast porównanie pomiędzy BAEL oraz NSP 2011 ze względu na okres jaki minął od spisu zostało pominięte w analizie porównawczej.



Rysunek 2.3: Porównanie liczby osób pracujących w podziale na płeć na poziomie województw w 2017 roku pomiędzy poszczególnymi źródłami danych

Źródło: Opracowanie własne.

Największa zgodność pomiędzy źródłami danych w odniesieniu do liczby osób pracujących widoczna jest w przypadku Badania Aktywności Ekonomicznej Ludności oraz *Zintegrowanego rejestru osób pracujących i bezrobotnych* (por. rysunek 2.3a). Pomimo przytaczanych wcześniej różnic definicyjnych pomiędzy tymi źródłami oba ujmują zbiorowość pracujących według ich miejsca zamieszkania, co przekłada się na ich dość dużą zgodność. Warto w tym miejscu zauważyć, że poziom szacunków liczby pracujących w przypadku ZROPIB jest nieco niższy niż w BAEL-u.

Jedną z przyczyn widocznych na wykresie różnic może być to, iż w badaniu BAEL pytanie o status na rynku pracy nie odnosi się do stanu formalno-prawnego, ale faktycznego. Z tego powodu osoby pracujące w szarej strefie są zaliczane do pracujących. Dodatkowo do pracujących w badaniu BAEL zalicza się uczniów, studentów oraz emerytów, którzy pracowali zarobkowo w okresie podlegającym obserwacji przynajmniej przez jedną godzinę, co może wyjaśniać wyższą liczbę osób pracujących wykazywaną w tym badaniu. Z kolei ograniczeniem w tym badaniu, zwłaszcza w porównaniu z ZROPIB, jest uwzględnienie w nim jedynie osób w gospodarstwach domowych. Należy mieć również na uwadze, że co prawda oba źródła ujmują pracujących według miejsca zamieszkania, jednak w BAEL-u ta informacja ma charakter deklaracyjny, a w ZROPIB jest to stan formalno-prawny. To może mieć znaczenie w przypadku dużych aglomeracji, gdzie informacje uzyskane w badaniu BAEL mogą odbiegać od stanu formalno-prawnego.

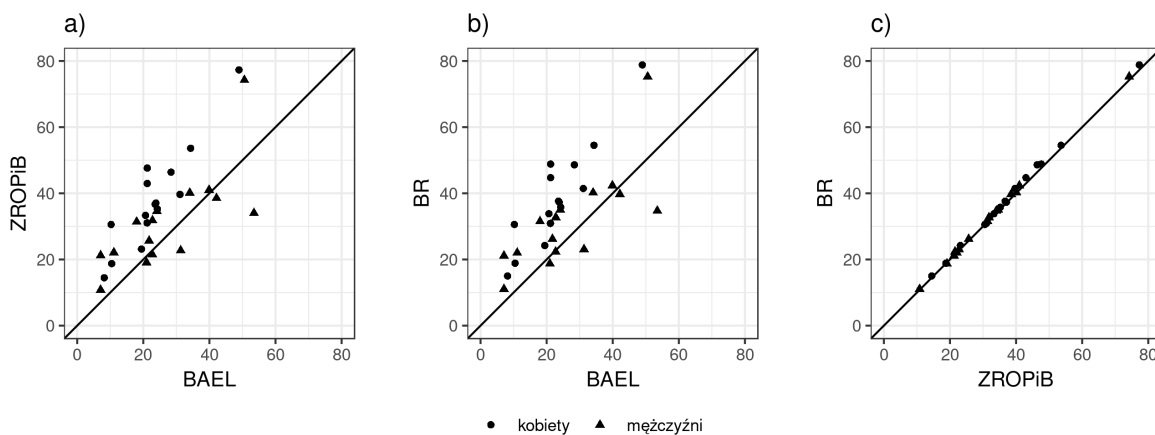
Kolejny wykres (por. rysunek 2.3b) przedstawia porównanie liczby osób pracujących według badania Pracujący w gospodarce narodowej oraz na podstawie Badania Aktywności Ekonomicznej Ludności. W tym przypadku widać większe rozbieżności pomiędzy analizowanymi źródłami w postaci przeszacowania liczby osób pracujących w badaniu PwGN w porównaniu do BAEL, co jest widoczne szczególnie w przypadku dwóch obserwacji odstających. Są to obserwacje odnoszące się do województwa mazowieckiego w podziale na płeć. Można zakładać, że wynika to w głównej mierze z różnic definicyjnych osoby pracującej, jakie występują pomiędzy źródłami, w szczególności to, że w BAEL-u osoby te są ujmowane według miejsca zamieszkania, natomiast w badaniu PwGN według faktycznego miejsca wykonywania pracy. Można zakładać, iż województwo mazowieckie z Warszawą i jej najbliższym otoczeniem zawiera w swoim obszarze bardzo dużą liczbę firm, więc powiązanie jej z miejscem faktycznie wykonywanej pracy przez pracujących, zgodnie z definicją osoby pracującej w gospodarce narodowej, przekłada się na różnice pomiędzy analizowanymi źródłami. Wpływ na te różnice ma również to, że informacja o liczbie pracujących w badaniu PwGN podawana jest według stanu na koniec roku. Natomiast w BAEL porównywane dane mają charakter średnioroczny, co niweluje zjawisko występującej sezonowości.

Podobny układ widoczny jest na ostatnim wykresie (por. rysunek 2.3c), na którym co prawda można dostrzec większą zgodność pomiędzy źródłami w porównaniu do wykresu zamieszczonego na rysunku 2.3b, jednak nadal zauważalne są dwie obserwacje odstające odpowiadające pracującym mężczyznom oraz kobietom w województwie mazowieckim. W tym przypadku występuje co prawda zgodność co do samej definicji osób pracujących w *Zintegrowanym rejestrze osób pracujących i bezrobotnych* oraz w badaniu *Pracujący w gospodarce narodowej*, mamy tu jednak do czynienia z różnym sposobem w ujmowaniu zbiorowości. ZROPiB grupuje osoby pracujące w gospodarce narodowej według miejsca zamieszkania, podczas gdy drugie źródło według faktycznego miejsca wykonywania pracy. Różnice wynikają również ze stanu danych w obu badaniach. W ZROPiB są to dane średnioroczne, natomiast informacja o liczbie osób pracujących w badaniu PwGN podawana jest zgodnie ze stanem na koniec roku.

Należy przypuszczać, że przy przejściu na niższy poziom agregacji przestrzennej mielibyśmy do czynienia z większym zróżnicowaniem dotyczącym liczby osób pracujących. Można bowiem założyć, że większość osób pracujących wykonuje swoją pracę w granicach wo-

jewództwa, które zamieszkuje. Wówczas różnice w definicjach w tym zakresie się znoszą. Tymczasem, jak wskazują wyniki badań na poziomie gmin i powiatów (GUS, 2019b)⁴³, informacje są pod tym względem znacznie bardziej zróżnicowane, również w zależności od tego, czy są to informacje o charakterze średniokresowym, czy też jest to stan na koniec badanego okresu.

Największa zgodność w odniesieniu do liczby osób bezrobotnych widoczna jest pomiędzy BAEL-em, a ZROPiB (por. rysunek 2.4a). Zauważyć można, iż w większości przypadków liczba osób bezrobotnych według ZROPiB jest wyższa niż w BAEL, chociaż zdarzają się również sytuacje odwrotne. Niezależnie od różnic, widoczna jest wyraźna korelacja pomiędzy poziomami badanego zjawiska.



Rysunek 2.4: Porównanie liczby osób bezrobotnych w podziale na płeć na poziomie województw w 2017 roku pomiędzy poszczególnymi źródłami danych

Źródło: Opracowanie własne.

Kolejny rysunek 2.4b jest bardzo podobny do wykresu 2.4, co wynika z faktu, iż badanie bezrobocia rejestrowanego oraz ZROPiB jest zasilane informacjami pochodzącymi z powiatowych urzędów pracy.

Natomiast ostatni wykres (por. rysunek 2.4c) wskazuje niemalże na pełną zgodność pomiędzy rozważanymi źródłami, tj. ZROPiB oraz BR. Niewielkie różnice widoczne na ostatnim wykresie mogą wynikać z przesunięć w przekazywaniu informacji o liczbie osób bezrobotnych do ZUS przez powiatowe urzędy pracy. Ta sama przyczyna może powodować pewne różnice pomiędzy pierwszym i drugim wykresem (por. rysunek 2.4a i 2.4b). Inaczej niż w przypadku osób pracujących, informacje o liczbie osób bezrobotnych podawane są

⁴³Można także założyć, że uwaga ta odnosi się do miejskich obszarów funkcjonalnych.

we wszystkich porównywanych na wykresach źródłach danych jako dane średnioroczne. W przypadku rejestru ZROPiB oraz badania BR jest to średnia arytmetyczna ze stanów na koniec poszczególnych kwartałów. Natomiast w badaniu BAEL, zgodnie z przyjętą metodologią badania i opracowywania jego wyników, jest to średnia z szacunków kwartalnych.

2.7. Podsumowanie

Na podstawie przeprowadzonego przeglądu źródeł danych i dokonanych porównań można wskazać, że najpełniejszym, najbardziej aktualnym i zgodnym z wymogami definicji międzynarodowych źródłem danych o kategoriach rynku pracy jest cyklicznie prowadzone Badanie Aktywności Ekonomicznej Ludności. Uwzględnia ono faktyczną sytuację osób na rynku pracy, niezależnie od ich statusu formalno-prawnego. Nie dostarcza ono jednak informacji statystycznej na poziomie rozważanych w pracy obszarów funkcjonalnych miast.

Z kolei *Zintegrowany rejestr osób pracujących i bezrobotnych* daje możliwość uzyskania na tym poziomie agregacji przestrzennej informacji dla dwóch kategorii rynku pracy, tj. pracujących i bezrobotnych. Jednakże i w tym przypadku informacja ta jest nieadekwatna, gdyż definicje osób pracujących oraz bezrobotnych odbiegają od określonych przez Międzynarodową Organizację Pracy. W rejestrze ZROPiB brak jest również informacji o osobach biernych zawodowo.

Pomimo niewątpliwych różnic pomiędzy BAEL i ZROPiB, istnieje duża zgodność pomiędzy szacunkami liczby osób pracujących i bezrobotnych, jakich dostarczają oba źródła. Zależności te można wykorzystać w celu poprawy jakości szacunku najważniejszych charakterystyk dotyczących rynku pracy stosując wybrane techniki estymacji pośredniej. Metodom tym poświęcony zostanie kolejny rozdział pracy. Zawiera on wprowadzenie do problematyki zagadnienia statystyki małych obszarów oraz teorii estymacji typu SPREE.

Rozdział 3

Estymacja typu SPREE

Rozdział trzeci poświęcono charakterystyce estymacji typu SPREE. Pierwsza jego część stanowi teoretyczne wprowadzenie do zagadnień związanych z estymacją pośrednią z uwzględnieniem metod oceny jakości szacunku parametrów¹. Opisano w niej również tak zwane **podejście procesowe** do estymacji pośredniej. Część druga poświęcona jest estymacji typu SPREE. W pierwszej kolejności przybliżono jej ideę. Następnie opisano ją z punktu widzenia jednego z podejść, które można zastosować w ramach szerokiej rodziny metod statystyki małych obszarów dla uzyskania szacunków liczby osób pracujących, bezrobotnych oraz biernych zawodowo. Oszacowania te są niezbędne z punktu widzenia wyznaczenia najważniejszych wskaźników opisujących sytuację na rynku pracy tj. stopy bezrobocia, współczynnika aktywności zawodowej oraz wskaźnika zatrudnienia. W dalszym ciągu przedstawiono ewolucję metod estymacji typu SPREE, począwszy od podejścia klasycznego, aż po obecnie rozwijane podejścia w ujęciu modelowym. W kolejnej części rozdziału zaprezentowano metody estymacji będące rozszerzeniem estymacji typu SPREE. Na końcu omówiono wybrane przykłady praktycznego zastosowania estymatorów typu SPREE, zaprezentowane w literaturze przedmiotu.

¹W niniejszej rozprawie doktorskiej termin „estymacja pośrednia” – zamiennie używany jest z takimi terminami jak: statystyka małych obszarów, estymacja dla małych domen, techniki SMO.

3.1. Wprowadzenie do estymacji pośredniej

3.1.1. Problematyka statystyki małych obszarów

Rozwój alternatywnych podejść do szacunku parametrów, w tym metod estymacji pośredniej, jest w głównej mierze spowodowany ograniczeniami, jakie napotyka się w badaniach prowadzonych metodą reprezentacyjną. Metoda ta jest powszechnie wykorzystywana w statystyce publicznej, w badaniach częściowych w celu uzyskania oszacowań podstawowych parametrów analizowanych zmiennych na poziomie całego kraju lub województw. Zastosowanie metody reprezentacyjnej w miejsce badania pełnego pozwala zarówno na znaczne obniżenie kosztów badania, jak i zmniejszenie obciążeń respondentów.

W metodzie reprezentacyjnej wykorzystywany jest często klasyczny estymator bezpośredni Horvitz–Thompsona (HT). Gdy szacujemy wartości globalne analizowanych zmiennych, przyjmuje on następującą postać:

$$\hat{Y}_d^{HT} = \sum_{j \in s_d} d_j y_j, \quad (3.1)$$

gdzie:

- s_d to część próby s wylosowanej z domeny d ,
- y_j to wartość badanej zmiennej dla jednostki j ,
- d_j to waga pierwotna, będąca odwrotnością prawdopodobieństwa wylosowania do próby jednostki j .

W badaniach reprezentacyjnych pojawia się problem braków odpowiedzi. Mają one istotny wpływ na jakość szacunków i stąd też muszą być uwzględnione w procesie estymacji. Jednym z możliwych podejść zaproponowanych w teorii, jak i stosowanych w praktyce badawczej, jest dokonanie korekty pierwotnych, przyjętych w badaniu wag poprzez zastąpienie ich wagami kalibracyjnymi².

Otrzymany w ten sposób estymator kalibracyjny, który będzie również wykorzystany w rozprawie doktorskiej do szacunków parametrów, wyraża się wzorem:

²Alternatywą dla kalibracji mogą być wybrane metody imputacji.

$$\hat{Y}_d^{CAL} = \sum_{j \in s_d} w_j y_j, \quad (3.2)$$

gdzie:

- w_j to waga kalibracyjna dla jednostki j uwzględniająca wszystkie modyfikacje wagi pierwotnej.

Głównym ograniczeniem tego estymatora jest to, że w procesie estymacji wykorzystywane są jedynie obserwacje wylosowane do próby z danej domeny³. Takie rozwiązanie bywa wystarczające w przypadku subpopulacji, które są licznie reprezentowane w próbie. Dlatego też estymator kalibracyjny stosowany jest często w badaniach reprezentacyjnych w celu zapewnienia wysokiej jakości oszacowań na poziomie krajowym, a czasem również regionalnym tzn. na poziomie województwa. Ważną cechą tego estymatora jest jego asymptotyczna nieobciążoność. Objawia się ona tym, iż wraz ze wzrostem liczby obserwacji w próbie w danej podgrupie maleje obciążenie otrzymywanych dla niej szacunków. Należy mieć jednak na uwadze, że w przypadku wysokiej wartości wskaźnika braków odpowiedzi w wylosowanej próbie w danym przekroju, estymator ten charakteryzuje się istotnym obciążeniem niezależnie od wielkości zrealizowanej próby. Wpływa to niewątpliwie na jakość uzyskiwanej wynikowej informacji statystycznej. Dodatkowo szacunki parametrów dla domen nielicznie reprezentowanych w próbie charakteryzują się niską precyzją, a w przypadku subpopulacji nie posiadających w próbie swoich reprezentantów nie ma możliwości przeprowadzenia jakiegokolwiek procesu estymacji.

Bazując na rozwiązaniach wypracowanych w metodzie reprezentacyjnej, jedną z możliwych alternatyw jest zwiększanie wielkości próby we wszystkich przewidzianych w analizie przekrojach. Wiąże się to jednak z istotnym zwiększeniem kosztów badania i obciążeniem respondentów, co w powiązaniu z rosnącym z roku na rok wskaźnikiem odmów odpowiedzi w prowadzonych badaniach nie jest satysfakcjonującym rozwiązaniem. Dodatkowo coraz częściej zgłaszane są potrzeby użytkowników dotyczące popytu na informację na niskim poziomie agregacji, który nie był uwzględniany na etapie projektowania badania reprezentacyjnego. Dlatego też poszukuje się alternatywnych metod estymacji dla pożądaných przez odbiorców informacji statystycznej przekrojów, które w mniejszym lub większym stopniu niwelują ograniczenia metody reprezentacyjnej.

³Zamiennie do słowa domena będą stosowane wyrażenia: subpopulacja, podgrupa, mały obszar.

Jednym z możliwych rozwiązań stosowanych w tego rodzaju wyzwaniach jest estymacja pośrednia zaproponowana w ramach statystyki małych obszarów (SMO). Obejmuje ona zbiór technik, które często umożliwiają uzyskiwanie wystarczająco rzetelnych i akceptowalnych szacunków parametrów zmiennych na takim poziomie agregacji, który okazuje się zbyt niski (lub zbyt szczegółowy), aby można było dla niego efektywnie skorzystać z metod estymacji bezpośredniej. Należy jednocześnie wyraźnie podkreślić za Dehnel (2003), że nie można utożsamiać technik statystyki małych obszarów jedynie z szacunkami dla niewielkich jednostek przestrzennych, takich jak na przykład powiat czy gmina. Termin „obszar” może również odnosić się do domen wyróżnionych na podstawie cech rzeczowych, takich jak na przykład płeć czy grupa wieku. Dlatego też, przy wykorzystaniu tych metod uzyskuje się szacunki dla określonych subpopulacji, które łączą w sobie zarówno podział terytorialny, jak i rzeczowy. Punktem wyjścia i jednocześnie odniesienia dla technik statystyki małych obszarów jest estymacja pośrednia. Jeśli wielkość próby dla danego obszaru\domeny uniemożliwia uzyskanie szacunków bezpośrednich o odpowiednim stopniu precyzji, wtedy taki obszar traktuje się jako „mały”. Jak pisze Molina i Rao (2015), w takiej sytuacji problem estymacji można rozwiązać wykorzystując informacje pomocnicze pozyskane z innych źródeł danych uwzględniając powiązania istniejące między zmienną badaną, a zmiennymi pomocniczymi. Takie rozwiązanie rekompensuje niedostateczną ilość informacji z próby w danej domenie. Jak podkreśla Luna (2016), techniki SMO, ze szczególnym uwzględnieniem estymacji typu SPREE, dostarczają użytecznego mechanizmu łączenia informacji z różnych źródeł i szacowania wybranych parametrów na podstawie danych zintegrowanych w ramach modelu statystycznego.

W związku z tym kluczową rolę w estymacji dla małych domen odgrywają informacje spoza próby, które możemy pozyskiwać z wielu zewnętrznych źródeł danych. Molina i Rao (2015) wskazują, że jednym z nich jest samo badanie reprezentacyjne, gdzie „moc” pożyczana jest z wyższego poziomu agregacji przestrzennej czy też rzeczowej, z domen tego samego poziomu licznie reprezentowanych w próbie bądź też z wcześniejszych edycji tego samego badania.

Niezwykle cennym źródłem zmiennych pomocniczych są spisy powszechne, a ich kluczową zaletą jest to, że definicje występujących w nim zmiennych są bardzo często zgodne z definicjami zmiennych w badaniu reprezentacyjnym. Dzięki temu można je zastosować jako zmienne objaśniające do modelu. Charakteryzując spisy pod kątem ich użyteczności w sta-

tystyce małych obszarów należy także wskazać na ich słabszą stronę. Istotną wadę spisów stanowi ich postępująca dezaktualizacja. Przekłada się ona na jakość oszacowań parametrów dla zmiennych pochodzących z bieżącego badania reprezentacyjnego, którego termin realizacji może być znacznie oddalony od ostatnio przeprowadzonego spisu powszechnego. Warto dodać ponadto, że w przypadku spisu przeprowadzonego w Polsce w 2011 r. informacje dotyczące części zmiennych pochodziły wyłącznie z badania reprezentacyjnego, a nie z badania pełnego. Wpłynęło to znacząco na zmniejszenie możliwości wykorzystania tych informacji jako źródła zmiennych pomocniczych. Dane te bowiem, w zależności od analizowanego przekroju terytorialnego czy rzeczowego, obarczone były większym lub mniejszym błędem losowym.

Szczególnie istotnym i wartościowym źródłem zmiennych pomocniczych, wykorzystywanych w modelach statystyki małych obszarów, są rejestry administracyjne. Jak już wspomniano w podrozdziale 2.2 należy podkreślić, że są to źródła danych, które nie są tworzone dla celów statystycznych. Prowadzi to niekiedy do trudności w ich bezpośrednim wykorzystaniu, wynikających przede wszystkim z różnic występujących w definicjach poszczególnych zmiennych. Często bowiem jest tak, że inne definicje stosowane są przez statystykę publiczną na potrzeby badań statystycznych, inne zaś obowiązują w rejestrach administracyjnych.

Ostatnią grupę zmiennych pomocniczych, która nabiera obecnie kluczowego znaczenia są alternatywne – dla badań statystycznych oraz rejestrów administracyjnych – źródła danych. Przykładami takich źródeł danych mogą być zbiory big data czy internetowe źródła danych. Rozważania na temat możliwości ich potencjalnego wykorzystania można znaleźć w wielu opracowaniach: Koebe i in. (2021); Szreder (2019, 2017, 2015); Schmid i in. (2017); Wagner i in. (2017); Groß i in. (2017); Beręsewicz i Szymkowiak (2015); Marchetti i in. (2015, 2016).

Należy jednocześnie zaznaczyć, wykorzystując rekomendacje Australijskiego Biura Statystycznego (Australian Bureau of Statistics, 2006), iż informacje pomocnicze, aby mogły być z powodzeniem wykorzystane do uzyskania wiarygodnych szacunków dla małych obszarów powinny:

- być dostępne dla całej populacji we wszystkich małych obszarach;

- zawierać informacje umożliwiające ich jednoznaczne powiązanie z poszczególnymi małymi obszarami;
- odnosić się do tego samego okresu co dane dla zmiennej badanej oraz innych zmiennych pomocniczych uwzględnionych w modelu;
- być silnie powiązane ze zmienną badaną tj., powinien to być związek uzasadniony merytorycznie;
- pozostawać w statystycznie istotnej relacji (najczęściej wyrażonej współczynnikiem korelacji) z bezpośrednimi szacunkami dla małych obszarów.

Informacje pomocnicze mogą być dostępne na poziomie obszarów jako dane zagregowane, dane jednostkowe lub mogą stanowić kombinację obu możliwości. Jak pisze Ghosh (2020), w zależności od rodzaju dostępnych informacji pomocniczych w SMO stosowane są w ramach tak zwanego podejścia modelowego dwa rodzaje modeli. Pierwszy z nich to tzw. modele zagregowane, które buduje się, gdy dane dotyczące zmiennej badanej, jak i zmiennych pomocniczych, są na poziomie małych obszarów. Drugi rodzaj modeli stanowią modele dla danych jednostkowych.

Jak podkreśla Rao (2015) w estymacji dla małych domen wyróżnia się z reguły dwie podstawowe grupy estymatorów: estymatory wykorzystywane w obszarze klasycznej metody reprezentacyjnej, tzn. oparte na schemacie losowania próby (ang. *design-based*) oraz estymatory wykorzystujące wskazane wyżej podejścia modelowe (ang. *model-based*). Do pierwszej grupy zalicza się przede wszystkim estymatory wykorzystujące wagi wynikające ze schematu losowania próby. Należą do tej grupy m.in. estymatory bezpośrednie, GREG, estymatory kalibracyjne, syntetyczne oraz złożone, stanowiące kombinację liniową estymatora bezpośredniego i syntetycznego.

Drugą grupę stanowią natomiast estymatory wykorzystujące odpowiednio skonstruowany model. Należą do niej przede wszystkim empiryczne estymatory bayesowskie (ang. *Empirical Bayes* – EB), hierarchiczne estymatory bayesowskie (ang. *Hierarchical Bayes* – HB) oraz empiryczne najlepsze liniowe nieobciążone predyktory (ang. *Empirical Best Linear Unbiased Predictor* – EBLUP).

Nie jest to jedyna klasyfikacja estymatorów, jaką można znaleźć w literaturze przedmiotu. Inną zaproponował Schaible (2013). Zgodnie z nią estymatory pośrednie można pogrupować biorąc pod uwagę zakres dostępu do informacji pomocniczych. Wyróżnić wówczas

można trzy rodzaje estymatorów pośrednich: pośrednie ze względu na domenę (co oznacza, że informacje pomocnicze pochodzą z innego obszaru, ale z tego samego okresu), pośrednie ze względu na czas (czyli informacje pomocnicze pochodzą z tego samego obszaru, ale z innego okresu) oraz estymatory pośrednie ze względu na domenę i czas.

Wybór odpowiedniej metody szacunku odgrywa w badaniu ważną rolę i przekłada się w sposób bezpośredni na jakość estymacji. Rola ta jest szczególnie istotna w sytuacji, kiedy dokonujemy szacunku na niskim poziomie agregacji przy wykorzystaniu metod estymacji pośredniej. Cenione jest tu zwłaszcza podejście modelowe, w którym wyboru konkretnego modelu dokonuje się w głównej mierze w zależności od dostępu do dobrej jakości zmiennych pomocniczych oraz celu badania. Warto w tym miejscu również podkreślić, iż wysoko cenioną i pożądaną własnością estymatorów pośrednich, którą warto uwzględnić przy ich wyborze, jest spójność oszacowań z oszacowaniami bezpośrednimi otrzymanymi na wyższym poziomie agregacji (najczęściej na poziomie kraju lub ewentualnie województw). Własność ta bywa zachowywana jedynie w przypadku nielicznej grupy estymatorów statystyki małych obszarów. Należą do niej przede wszystkim estymatory syntetyczne, a wśród nich estymatory typu SPREE (ang. *Structure Preserving Estimation* – estymatory zachowujące strukturę). Dla pozostałych estymatorów pośrednich po uzyskaniu ocen szacowanych parametrów powszechną i często stosowaną metodą zapewnienia spójności z oszacowaniami bezpośrednimi na wyższym poziomie agregacji jest raking. Pozwala on na uzyskanie wymaganej zgodności (ang. *benchmarking*) z ocenami estymatorów bezpośrednich. Nie jest to jednak rozwiązanie optymalne, gdyż trudno jest zmierzyć wpływ metody rakingu na ostateczną jakość uzyskanych szacunków parametrów.

Szerokie kompendium wiedzy dotyczącej metod estymacji pośredniej stanowią między innymi: książka Moliny i Rao (2015) oraz artykuł Ghosha (2020), a także publikacja Hidiroglou i współpracowników (2019). Z kolei ważną publikacją poświęconą użyteczności metod SMO i kierunków ich rozwoju stanowi artykuł Pfeffermana (2013).

3.1.2. Pomiar jakości oszacowań pośrednich

Istotnym elementem procesu estymacji pośredniej jest ocena jakości uzyskanych oszacowań. Analizując literaturę przedmiotu z zakresu metod estymacji pośredniej trudno jest wskazać jedną uniwersalną miarę. Dla kontrastu, w przypadku estymatora bezpośredniego

za taką można przyjąć ocenę jego wariancji. Przyjmuje się bowiem, że estymator bezpośredni jest nieobciążony, a zatem ocenie podlega jedynie jego efektywność. W przypadku estymacji pośredniej wariancja powinna stanowić pierwszy element oceny jakości szacunku wykorzystywany także przy weryfikacji empirycznej zakładanego zysku na precyzji. Zgodnie bowiem z teorią statystyki małych obszarów, w przypadku domen nielicznie reprezentowanych w próbie, stosując estymator pośredni w miejsce estymatora bezpośredniego zyskujemy na precyzji. Analizując jakość ocen parametrów uzyskanych z wykorzystaniem technik SMO ocenie należy również poddać obciążenie szacunku. Zgodnie z propozycją zgłoszoną w projekcie ESSnet on SAE w ramach pakietu WP6 (2012), ocena ta może przyjąć formę graficzną. Zakładamy wówczas, że oceny estymatorów pośrednich, jeśli są nieobciążone, powinny przyjmować taki sam poziom jak oceny estymatorów bezpośrednich w ramach przyjętego modelu. Można to zweryfikować konstruując diagram rozrzutu, w którym jedna oś dotyczy oszacowań bezpośrednich, natomiast druga oszacowań pośrednich. Wówczas punkty znajdujące się wzdłuż prostej o równaniu $y = x$ świadczą o nieobciążoności estymatora pośredniego. Z kolei im większe jest oddalenie punktu od tej prostej, tym większe obciążenie estymatora pośredniego. W ramach tego samego projektu oraz w artykule Browna i współautorów (2001) proponuje się również inne metody oceny jakości oszacowań, takie jak: badanie dobroci dopasowania (ang. *goodness-of-fit diagnostic*) czy też badanie stopnia pokrycia (ang. *coverage diagnostic*). Polecaną miarą oceny jakości oszacowań jest również MSE (ang. *mean square error*, pol. *błąd średniokwadratowy, średni błąd kwadratowy*). Jak jednak pisze Tzavidis i współautorzy (2018), problem wyboru sposobu wyznaczania MSE za pomocą podejścia opartego na modelu (ang. *model-based*) czy też opartego na schemacie losowania (ang. *design-based*) jest ciągle aktualny i wywołuje ożywioną debatę wśród badaczy. Każde z wyżej wymienionych podejść posiada swoje zalety i wady. Ich poprawność uwarunkowana jest zachowaniem określonych założeń, które w warunkach badań opartych na danych rzeczywistych nie zawsze są spełnione. Dlatego też wydaje się, że z praktycznego punktu widzenia, szczególnie w odniesieniu do krajowych urzędów statystycznych, ostateczna weryfikacja metod pośrednich powinna odbywać się w odniesieniu do uzyskanych wyników z badania pełnego, tzn. spisu powszechnego. Wówczas można zweryfikować rzeczywiste obciążenie estymatorów pośrednich odnosząc uzyskane wyniki do wiarygodnych danych spisowych. Warto również zaznaczyć, że podstawową miarą oceny jakości stosowaną w ramach statystyki publicznej jest współczynnik zmienności – CV , który służy nie tylko do oceny wyników uzyskanych

za pomocą estymatora bezpośredniego, w tym również kalibracyjnego, ale także ilorazowego. Ogólna postać wzoru na wartość współczynnika zmienności wyraża się poniższym wzorem:

$$CV = \frac{\sqrt{Var(\hat{Y})}}{\hat{Y}} \cdot 100\%. \quad (3.3)$$

Do wyznaczenia współczynnika CV konieczne jest uzyskanie ocen wariancji estymatora. W przypadku złożonych schematów losowania, dla których poziomu wariancji nie da się wyznaczyć bezpośrednio ze wzoru, stosuje się metody przybliżone. Jedną z nich – powszechnie stosowaną w praktyce – jest metoda bootstrapowa. Opis teoretycznych podstaw metody oraz przykład jej praktycznego zastosowania można znaleźć m.in. w artykule Młodaka (2014). Metoda bootstrapowa pozwala na ujednoliczoną estymację wariancji dla różnych wskaźników, bez wyprowadzania złożonych wzorów analitycznych⁴.

3.1.3. Wskazówki i rekomendacje dla procesu estymacji pośredniej

Jeżeli chodzi o wymiar praktyczny stosowania metod pośrednich, kluczowy wydaje się pakiet roboczy WP6 projektu ESSnet on SAE⁵ zawierający wskazówki i rekomendacje w zakresie implementacji tych metod w realiach statystyki publicznej⁶. Przedstawiono tam między innymi w jaki sposób procesowo podchodzić do etapowego wdrażania metod statystyki małych obszarów. Wskazano, że na wstępie należy dobrze rozpoznać potrzeby użytkowników oraz dostępne źródła danych. Następnie trzeba założyć pożądany poziom agregacji przestrzennej i/lub rzeczowej dla estymacji parametrów zmiennych z wykorzystaniem metod pośrednich, którego oczekuje odbiorca wynikowych danych statystycznych. Zaznaczono również, że zaleca się osiągnięcie poziomu docelowego stopniowo poczynając od mniej złożonych estymatorów oraz wyższego poziomu agregacji. Autorzy projektu re-

⁴Szerszy opis zastosowanego podejścia bootstrap w części empirycznej niniejszej rozprawy znajduje się w rozdziale czwartym.

⁵Zob. https://ec.europa.eu/eurostat/cros/content/sae-finished_en.

⁶Autor rozprawy doktorskiej był również członkiem międzynarodowego zespołu w ramach projektu ESSnet on SAE opracowującego raporty badawcze zawierające wytyczne i rekomendacje odnośnie stosowania technik SMO w statystyce publicznej, por. https://ec.europa.eu/eurostat/cros/system/files/Heerlen_Meeting_Minutes.pdf_en.

komendują rozpoczęcie estymacji pośredniej od przekrojów bardziej ogólnych i przechodzenie do poziomów bardziej szczegółowych, aż do osiągnięcia przekroju docelowego, na każdym z etapów oceniając jakość otrzymywanych wyników. Jeżeli wyniki uzyskane na etapach pośrednich są zadowalające, wówczas należy przejść do docelowego poziomu estymacji wybierając najmniej złożony estymator, który dostarcza optymalnych wyników pod względem jakości. Praktyczny przykład schematu zastosowania powyższych wytycznych, z pewnymi modyfikacjami w odniesieniu do estymacji stopy ubóstwa, został przedstawiony w artykule Tzavidisa i współautorów (2018).

W niniejszej pracy doktorskiej proces estymacji pośredniej był również realizowany zgodnie z powyższymi zaleceniami.

3.2. Idea estymacji typu SPREE

Jak już wcześniej wspomniano w podrozdziale 3.1 jedną z pożądanych własności estymatorów jest to, że ich oceny pozostają spójne z oszacowaniami bezpośrednimi na wyższym poziomie agregacji przestrzennej. Grupą estymatorów, które z założenia spełniają tę własność, są estymatory typu SPREE. Głównym i pierwotnym celem stosowania estymacji typu SPREE jest wypełnienie luki informacyjnej, która występuje w okresach międzyspisowych w wielu obszarach badawczych. Chodzi tu przede wszystkim o to, że estymatory te dają możliwość dostarczenia aktualnych ocen parametrów, zapobiegając tym samym dezaktualizacji informacji pochodzących ze spisu. Istotnym elementem tego typu estymacji pośredniej jest to, że jest ona przeprowadzana w układzie całej tabeli kontyngencji skonstruowanej do prezentacji danych spisowych, w celu aktualizacji jej zawartości dla okresów pospisowych. Podstawowe podejście polega na połączeniu znanych i wiarygodnych informacji pochodzących z aktualnych badań reprezentacyjnych i/lub bilansów struktur ludności, które stanowią wartości brzegowe tabeli kontyngencji, z danymi spisowymi w środku tabeli kontyngencji, które zdążyły się już zdezaktualizować. Należy zaznaczyć, iż przed rozpoczęciem procesu estymacji częstości w komórkach wewnątrz tabeli pochodzących ze spisu nie sumują się do wartości brzegowych pochodzących z aktualnego badania reprezentacyjnego. Dopiero zastosowanie estymatora typu SPREE prowadzi do pełnej zgodności między wartościami brzegowymi, a wnętrzem tabeli kontyngencji.

Jak podaje Molina i Rao (2015), estymatory typu SPREE mogą być wykorzystywane do szacowania wartości globalnych dla małych obszarów w okresach międzypisowych. Stanowią one zatem specjalną klasę estymatorów pozwalających na uzyskanie ocen parametrów, które zwykle dostępne są jedynie dla okresu spisowego, na niskim poziomie agregacji przestrzennej lub dla szczegółowych przekrojów określonych przez inne cechy.

Estymatory typu SPREE można również wykorzystać do konstrukcji tabel, które nie są zwykle dostarczane przez statystykę publiczną, z uwagi na zbyt duży błąd szacunku estymatora bezpośredniego dla przynajmniej niektórych komórek wewnątrz tabeli kontyngencji. Zastosowanie tego podejścia zostanie zaprezentowane w niniejszej rozprawie doktorskiej.

Jak wskazują Molina i Rao (2015), estymatory typu SPREE stanowią uogólnioną klasę estymatorów syntetycznych, które w pełni wykorzystują informacje wynikające z szacunków bezpośrednich. Szacowanymi wartościami globalnymi są tutaj wielkości znajdujące się w środku tabeli kontyngencji. Zatem tabelę, której wymiar uwzględnia badaną populację możemy oznaczyć – za Luną i współautorami (2014) – symbolem Y i zgodnie z zaleceniem Purcella i Kisha (1980) podzielić ją na dwie części. Pierwsza z nich to znajdująca się w środku tabeli kontyngencji struktura asocjacji (ang. *association structure*). Natomiast wartości brzegowe tabeli określane są mianem struktury alokacji (ang. *allocation structure*). Więcej informacji na ten temat znajduje się w podrozdziale 3.3.

W procesie estymacji bezpośredniej dla małych domen częstym problemem w odniesieniu do struktury asocjacji jest brak możliwości uzyskania oszacowań dla części komórek ze względu na brak reprezentantów w próbie, bądź też ich niewielką liczbę, która nie zapewnia odpowiedniej jakości szacunku estymatora bezpośredniego. Struktura alokacji może jednak zawierać wartości, które pozbawione są błędu losowego (np. struktury ludności), bądź też mogą być oszacowane z akceptowalnym poziomem błędu (niewielką wariancją) na podstawie bieżących badań reprezentacyjnych.

W podejściu SPREE, zaproponowanym przez Purcella i Kisha (1980)⁷, zachowując strukturę alokacji tabeli Y , w miejsce jej struktury asocjacji, jak podaje Luna (2016), wykorzystujemy strukturę asocjacji z tabeli zastępczej (ang. *proxy table*) oznaczaną jako X . Informacje znajdujące się w tabeli X mogą pochodzić z jednego bądź kilku spisów lub

⁷W pracy podejście to określane będzie mianem klasycznego ze względu na fakt, że Purcell i Kish jako pierwsi opisali ten rodzaj estymacji.

rejestrów administracyjnych, czy też kombinacji tych źródeł. Zakłada się również, że tak zwana tabela zastępcza X jest tego samego wymiaru co tabela Y . Określenie „tabela zastępcza” odnosi się do zmiennej zastępczej (ang. *proxy variable*). Została ona zdefiniowana przez Uptona i Cooka (2008) jako zmienna mierzalna, która jest używana w miejsce zmiennej, której nie można zmierzyć. Z kolei w podejściu wykorzystującym uogólnioną postać estymatora SPREE (zaproponowanego przez Zhanga i Chambersa (2004) oraz zastosowanego w niniejszej rozprawie doktorskiej) w procesie estymacji wykorzystywana jest również struktura asocjacji tabeli Y . Więcej informacji na ten temat znajduje się w podrozdziale 3.4.

Warto również zaznaczyć, że w przypadku danych spisowych tabela zastępcza zawiera najczęściej taki sam lub bardzo zbliżony co do definicji zestaw zmiennych, choć pochodzący z innego okresu (moment spisu). Z kolei w przypadku tabeli zastępczej opartej na rejestrach administracyjnych istnieje przynajmniej częściowa zgodność definicji. Uprawnia to z reguły badaczy do wykorzystywania takich informacji jako zmiennych pomocniczych w procesie estymacji pośredniej.

Z kolei odnosząc się do zakresu dostępnych wartości brzegowych, stanowiących strukturę alokacji tabeli Y , rozróżniamy podejście jednokrokowe⁸ (w którym struktura alokacji dostępna jest dla wierszy albo kolumn) oraz dwukrokowe (gdy struktura alokacji dostępna jest zarówno dla wierszy jak i kolumn). Dodatkowo w literaturze przedmiotu w odniesieniu do estymatorów typu SPREE, rozpatrywane są tabele w ujęciu dwuwymiarowym oraz trójwymiarowym. Więcej informacji na ten temat znajduje się w podrozdziale 3.4 oraz 3.5.

Warte podkreślenia jest również to, że estymatory SPREE, zarówno w przypadku klasycznego podejścia (Purcell i Kish, 1980), jak i jego uogólnienia zaproponowanego przez Zhanga i Chambersa (2004), do uzyskania ostatecznych wyników oszacowania docelowej struktury alokacji tabeli Y wykorzystują tzw. metodę iteracyjnego proporcjonalnego dopasowania (ang. *Iterative Proportional Fitting – IPF*). Jest to dość powszechnie znana z literatury przedmiotu technika korygowania liczebności w tabelach kontyngencji, tak aby dopasować je do znanych liczebności brzegowych (Deming i Stephan, 1940), czasami

⁸Podejście to jest poza zakresem rozważań niniejszej rozprawy doktorskiej m.in. ze względu na dostępność danych brzegowych na poziomie wierszy, jak i kolumn. Przykład podejścia jednokrokowego można znaleźć między innymi w artykule Józefowskiego i Szymkowiaka (2013).

także określana terminem *raking* (Stephan, 1942). Metoda ta jest wykorzystywana do aktualizacji liczebności, które w podejściu klasycznym opierają się jedynie na strukturze asocjacji tabeli X , a jej zastosowanie powoduje, że sumują się one do znanych wartości brzegowych tabeli Y . Natomiast w przypadku poszczególnych wariantów uogólnionego SPREE technika IPF jest wykorzystywana jako ostatni element procesu estymacji. Więcej informacji na ten temat znajduje się w podrozdziale 3.4 oraz 3.5.

3.3. Estymacja typu SPREE – ujęcie klasyczne

Celem przedstawienia idei dwukrokowego estymatora SPREE zaproponowanego przez Purcella i Kisha (1980) założmy, że estymacji podlegają liczebności $\{Y_{ak}\}$ w środku dwuwymiarowej tabeli kontyngencji Y , gdzie $a = 1 \dots A$ to mały obszar (np. województwo), zaś $k = 1, \dots, K$ to wariant zmiennej grupującej (np. pracujący, bezrobotni, bierni zawodowo). Przyjmijmy również, że znane są liczebności brzegowe wierszy $Y_{a.}$ i kolumn $Y_{.k}$ oraz wartości X_{ak} w tabeli zastępczej X , której wymiar jest taki sam jak tabeli Y tzn. ma A wierszy oraz K kolumn. Wartości X_{ak} mogą pochodzić ze spisu lub z rejestru administracyjnego.

Ideą klasycznego estymatora typu SPREE jest odpowiednie skorygowanie liczebności stanowiących strukturę asocjacji tabeli X , tak aby otrzymane w wyniku tej korekty wartości \hat{Y}_{ak}^{SPREE} były zgodne ze strukturą alokacji tabeli Y . Purcel i Kish (1980) zaproponowali, aby w tym celu zastosować iteracyjną procedurę IPF, którą opisać można za pomocą trzech kroków:

- 1) przeskalowujemy wiersze tabeli X wartościami brzegowymi wierszy tabeli Y :

$$\hat{Y}_{ak}^{(1)} = X_{ak} \frac{Y_{.k}}{X_{.k}}, \quad (3.4)$$

- 2) przeskalowujemy kolumny tabeli z uzyskanymi w poprzednim kroku oszacowaniami $\hat{Y}_{ak}^{(1)}$:

$$\hat{Y}_{ak}^{(2)} = \hat{Y}_{ak}^{(1)} \frac{Y_{a.}}{\hat{Y}_{a.}^{(1)}}, \quad (3.5)$$

- 3) przeskalowujemy wiersze tabeli z uzyskanymi w poprzednim kroku oszacowaniami $\hat{Y}_{ak}^{(2)}$:

$$\hat{Y}_{ak}^{(3)} = \hat{Y}_{ak}^{(2)} \frac{Y_{.k}}{\hat{Y}_{.k}^{(2)}}. \quad (3.6)$$

Kroki 2. i 3. powtarzane są naprzemiennie, przy czym w każdej kolejnej iteracji używa się tabeli otrzymanej w poprzedniej iteracji, aż do czasu uzyskania tabeli o pełnej zgodności z wartościami określonymi przez strukturę alokacji $Y_{.k}$ oraz $Y_{a.}$.

Zapewnienie zgodności oszacowań z wartościami brzegowymi tabeli Y nie jest istotne wyłącznie z arytmetycznego punktu widzenia, ale również jako ochrona przed błędną specyfikacją modelu. Z założenia pozwala to uniknąć niepożądanych odchyłeń szacunków opartych na modelu od odpowiadających im szacunków bezpośrednich na poziomach, na których te ostatnie są uważane za wiarygodne.

Purcell i Kisch (1980) wskazali również, że estymator SPREE da się wyrazić w formie modelu log-liniowego, w którym szukane liczebności Y_{ak} można wyznaczyć zgodnie z podejściem Bircha (1963) według następujących wzorów:

$$\log Y_{ak} = \alpha_0^Y + \alpha_a^Y + \alpha_k^Y + \alpha_{ak}^Y. \quad (3.7)$$

gdzie:

$$\alpha_0^Y = \frac{1}{AK} \sum_a \sum_k \log Y_{ak}, \quad (3.8)$$

$$\alpha_a^Y = \frac{1}{K} \sum_k \log Y_{ak} - \alpha_0^Y, \quad (3.9)$$

$$\alpha_k^Y = \frac{1}{A} \sum_a \log Y_{ak} - \alpha_0^Y, \quad (3.10)$$

$$\alpha_{ak}^Y = \log Y_{ak} - \alpha_a^Y - \alpha_k^Y - \alpha_0^Y. \quad (3.11)$$

Powyższe parametry α spełniają następujące warunki $\sum_a \alpha_a^Y = \sum_k \alpha_k^Y = \sum_a \alpha_{ak}^Y = \sum_k \alpha_{ak}^Y = 0$. Logarytmy liczebności Y_{ak} wyrażone wzorem (3.7) według propozycji Purcella i Kisha (1980) można przedstawić jako sumę składników odpowiadających dwóm strukturom:

- asocjacji – odpowiada ona zestawowi interakcji α_{ak} ($a = 1, \dots, A; k = 1, \dots, K$), które określają związek pomiędzy wierszami i kolumnami; z teoretycznego punktu widzenia, gdy dwa wymiary są niezależne od siebie, to $\alpha_{ak} = 0$ dla wszystkich par (a, k) ;
- alokacji – jest ona określona przez parametry $\alpha_0, \alpha_a, \alpha_k$ ($a = 1, \dots, A; k = 1, \dots, K$); struktura ta zawiera informacje o przeciętnej wartości, a także wadze poszczególnych jej wierszy i kolumn.

Zastosowanie techniki IPF prowadzi do przeniesienia struktury asocjacji tabeli X na oszacowania tabeli Y . Równość interakcji w szacowanej tabeli X oraz zastępczej tabeli Y , tj.:

$$\alpha_{ak}^Y = \alpha_{ak}^X, \quad a = 1, \dots, A; \quad k = 1, \dots, K. \quad (3.12)$$

stanowi założenie estymatora SPREE.

Na koniec powyższych rozważań warto również nadmienić, iż w artykule Szymkowiaka i Józefowskiego (2014) wykazano, że w celu zastosowania algorytmu IPF do wypełnienia środka tabeli kontyngencji nie są konieczne dane wejściowe ze spisu czy rejestru administracyjnego. W takim przypadku „punktem startowym” algorytmu IPF może być macierz złożona z samych jedynek. Jednak wyniki uzyskane w oparciu o takie postępowanie mogą istotnie odbiegać od rezultatów uzyskanych w sytuacji, gdy dane w tabeli X pochodzą od zmiennych merytorycznie powiązanych (zbliżone definicje, inny okres referencyjny przy tych samych definicjach) ze zmienną objaśnianą. Dlatego też uzyskane w ten sposób wyniki należy traktować ze szczególną ostrożnością, ponieważ struktura wejściowej tabeli kontyngencji ulega znacznym zmianom, pomimo iż sam algorytm IPF osiąga zbieżność. Stąd też wykorzystanie informacji pochodzących z dodatkowych źródeł (spisy, rejestry administracyjne) w celu wypełnienia środka tabeli kontyngencji mają zdecydowanie większe uzasadnienie merytoryczne.

3.4. GSPREE dla tabeli dwuwymiarowej

Kolejnym krokiem w rozwoju metodologii estymacji SPREE było rozwiązanie dla małych obszarów oparte na uogólnionych modelach liniowych zaproponowane przez Markera (1999). Choć opisana pierwotnie procedura nie zawierała estymatora SPREE, wskazywano na to, że estymator SPREE da się wyrazić jako model log-liniowy. Ta koncepcja została rozwinięta w pracy Nobla (2002), w której autorzy zastosowali podejście zaproponowane przez Nelderera (1972), w celu zaprezentowania estymacji SPREE w ramach uogólnionych modeli liniowych (ang. *Generalized Linear Model* – GLM). Wśród przykładów tej klasy modeli można wymienić regresję Poissona, regresję logistyczną oraz modele log-liniowe. Podejścia te są jednak poza zakresem rozważań w niniejszej rozprawie.

Ważnym etapem w rozwoju estymacji typu SPREE, które zostało w szerszym zakresie zaprezentowane i zaimplementowane w niniejszej rozprawie doktorskiej, jest podejście

przedstawione przez Zhanga i Chambersa (2004). Autorzy w swoim artykule zaproponowali zastosowanie uogólnionego SPREE (ang. *Generalized Structure Preserving Estimation* – GSPREE), wykazując, że oszacowania bezpośrednie wartości w środku tabeli Y wraz z wartościami tabeli zastępczej X mogą być użyte do modelowania struktury asocjacji w szacowanej tabeli Y . W zaproponowanym przez autorów uogólnionym liniowym modelu strukturalnym (ang. *Generalized Linear Structural Model* – GLSM) zakłada się, że interakcje w tabeli Y zawierającej szukane wartości są proporcjonalne do odpowiednich interakcji w tabeli zastępczej X .

Jest to zatem złagodzenie założenia, które występuje w klasycznym SPREE. Ponadto, uwzględniając dodatkowo w modelu efekty losowe na poziomie obszarów, w literaturze przedmiotu zaproponowano rozszerzenie uogólnionego SPREE w formie uogólnionego mieszanego liniowego modelu strukturalnego (ang. *Generalized Linear Structural Mixed Model* – GLSMM) (Zhang i Chambers, 2004).

Model GLSM

Omawiając rozwój estymacji typu SPREE warto zwrócić uwagę na model GLSM przedstawiony przez Zhanga i Chambersa (2004). Zaproponowali oni podejście szacowania struktury asocjacji θ^Y obejmującej nieznanne frakcje wewnątrzobszarowe $\theta_{ak}^Y = Y_{ak}/Y_a$ ($a = 1, \dots, A$, $k = 1, \dots, K$) przy wykorzystaniu struktury asocjacji tabeli zastępczej θ^X oraz struktury asocjacji oszacowań bezpośrednich $\hat{\theta}^Y$.

Założyli, że relacje pomiędzy strukturą asocjacji szukanej tabeli θ^Y i zastępczej tabeli θ^X opisać można za pomocą modelu GLSM:

$$\mu_{ak}^Y = \lambda_k + \beta \mu_{ak}^X, \quad \sum_{k=1}^K \lambda_k = 0, \quad (3.13)$$

gdzie funkcja wiążąca zdefiniowana jest następująco:

$$\mu_{ak}^Y = \log(\theta_{ak}^Y) - \frac{1}{K} \sum_{l=1}^K \log(\theta_{al}^Y). \quad (3.14)$$

Wyrażenia μ_{ak}^X zdefiniowane są w analogiczny sposób.

Korzystając z log-liniowej reprezentacji (3.7) wyrażenie (3.14) można sprowadzić do postaci:

$$\mu_{ak}^Y = \alpha_k^Y + \alpha_{ak}^Y. \quad (3.15)$$

Po podstawieniu do modelu (3.13) wyrażenia (3.15) i sumowaniu stronami w ramach obszarów $a = 1, \dots, A$ otrzymuje się równoważną jego postać:

$$\alpha_{ak}^Y = \beta \alpha_{ak}^X, \quad (3.16)$$

$$\alpha_k^Y = \lambda_k + \beta \alpha_k^X. \quad (3.17)$$

Ponieważ zawsze można dobrać takie λ_k , aby przy danym równaniu (3.16) spełnione było równanie (3.17), stąd model (3.13) zapisany w postaci (3.16)-(3.17) sprowadza się do równania (3.16), które określić można jako strukturalne założenie estymatora GSPREE. Oznacza ono, że struktura asocjacji tabeli Y jest proporcjonalna do struktury asocjacji w tabeli zastępczej X , gdzie β to współczynnik proporcjonalności. Model spełniający takie założenie określa się jako proporcjonalny model interakcji (ang. *proportional interactions model* – PIM). Należy również podkreślić, że klasyczny estymator SPREE jest specjalnym przypadkiem estymatora GSPREE, w którym $\beta = 1$.

Model GLSMM

Kolejne podejście w ramach estymacji typu SPREE reprezentuje model GLSMM. Zhang i Chambers (2004) zaproponowali tu złagodzenie założenia o proporcjonalności struktury asocjacji poprzez uwzględnienie efektów losowych w uogólnionym estymatorze SPREE. Do jego konstrukcji wykorzystali uogólniony mieszany liniowy model strukturalny postaci:

$$\mu_{ak}^Y = \lambda_k + \beta \mu_{ak}^X + v_{ak}, \quad (v_{a2}, \dots, v_{aK})^T \stackrel{IID}{\sim} N(\mathbf{0}, \Sigma), \quad v_{a1} = - \sum_{j=2}^K v_{aj}, \quad (3.18)$$

który oprócz efektów stałych λ_k uwzględnia również efekty losowe v_{ak} .

Postępując podobnie jak w przypadku modelu GLSM, tj. przechodząc na log-liniową reprezentację modelu (3.18), a następnie sumując go stronami, otrzymujemy równoważny układ równań:

$$\alpha_k^Y = \lambda_k + \beta \alpha_k^X + \bar{v}_{.k} = \lambda_k + \beta \alpha_k^X + O_p(A^{-1/2}), \quad (3.19)$$

$$\alpha_{ak}^Y = \beta \alpha_{ak}^X + v_{ak} - \bar{v}_{.k} = \beta \alpha_{ak}^X + v_{ak} + O_p(A^{-1/2}), \quad (3.20)$$

gdzie $\bar{v}_{.k} = A^{-1} \sum_a v_{ak}$. Wraz ze wzrostem liczby obszarów A , składnik $\bar{v}_{.k}$ dąży do zera, stąd jego wpływ na strukturę alokacji asymptotycznie zanika. Brak wpływu składników losowych na strukturę alokacji można byłoby osiągnąć narzucając ograniczenie $\bar{v}_{.k} = 0$,

jednak takie rozwiązanie mogłoby prowadzić do problemów z estymacją parametrów modelu. Zhang i Chambers (2004) sugerują alternatywne rozwiązanie, polegające na standaryzacji składników losowych ex-post tak, aby w ramach każdej z kategorii $k = 1, \dots, K$ sumowały się do 0, wykazując niewielki wpływ takiego zabiegu na ostateczne oceny estymatora GLSMM. Mając powyższe na uwadze, strukturalne założenie modelu GLSMM, podobnie jak w przypadku modelu GLSM, można sprowadzić do jednego równania:

$$\alpha_{ak}^Y = \beta \alpha_{ak}^X + v_{ak}. \quad (3.21)$$

Model spełniający powyższe założenie określa się jako proporcjonalny model interakcji z efektami mieszanymi (ang. *proportional interactions mixed model* – PIMM).

3.5. GSPREE dla trójwymiarowej tabeli kontyngencji

Rozszerzeniem podejścia omawianego w części 3.4 jest metoda pozwalająca na estymację liczebności w tabeli trójwymiarowej, w której oprócz podziału według obszarów $a = 1, \dots, A$ i kategorii $k = 1, \dots, K$ zmiennej grupującej, uwzględnia się dodatkowe przekroje według kategorii $j = 1, \dots, J$ zmiennej, która określana będzie jako warstwująca. Niech Y_{ajk} oznaczają szukane wartości, gdzie a oznacza obszar, k – kategorię zmiennej grupującej, j – kategorię zmiennej warstwującej, zaś X_{ajk} to odpowiadające im wartości w tabeli zastępczej.

Podobnie jak w przypadku tabel dwuwymiarowych, również i tu Zhang i Chambers (2004) rozważają log-liniową reparametryzację tabeli Y :

$$\log(Y_{ajk}) = \alpha_0^Y + \alpha_a^Y + \alpha_j^Y + \alpha_k^Y + \alpha_{aj}^Y + \alpha_{ak}^Y + \alpha_{jk}^Y + \alpha_{ajk}^Y. \quad (3.22)$$

gdzie:

$$\alpha_0^Y = \frac{1}{AJK} \sum_a \sum_j \sum_k \log Y_{ajk}, \quad (3.23)$$

$$\alpha_j^Y = \frac{1}{AK} \sum_a \sum_k \log Y_{ajk} - \alpha_0^Y, \quad (3.24)$$

$$\alpha_a^Y = \frac{1}{JK} \sum_j \sum_k \log Y_{ajk} - \alpha_0^Y, \quad (3.25)$$

$$\alpha_k^Y = \frac{1}{AJ} \sum_a \sum_j \log Y_{ajk} - \alpha_0^Y, \quad (3.26)$$

$$\alpha_{aj}^Y = \frac{1}{K} \sum_k \log Y_{ajk} - \alpha_a^Y - \alpha_j^Y - \alpha_0^Y, \quad (3.27)$$

$$\alpha_{ak}^Y = \frac{1}{J} \sum_j \log Y_{ajk} - \alpha_a^Y - \alpha_k^Y - \alpha_0^Y, \quad (3.28)$$

$$\alpha_{jk}^Y = \frac{1}{A} \sum_a \log Y_{ajk} - \alpha_j^Y - \alpha_k^Y - \alpha_0^Y, \quad (3.29)$$

$$\alpha_{ajk}^Y = \log Y_{ajk} - \alpha_{aj}^Y - \alpha_{ak}^Y - \alpha_{jk}^Y - \alpha_a^Y - \alpha_j^Y - \alpha_k^Y - \alpha_0^Y. \quad (3.30)$$

Należy zaznaczyć, że przypadek trójwymiarowej tabeli kontyngencji można sprowadzić do postaci dwuwymiarowej jeśli wstawimy $i = a \times j$, a następnie do Y_{ik} oraz X_{ik} zastosujemy opisane wcześniej modele dla tabel dwuwymiarowych. Warto dodać, że uwzględnienie w tabeli dodatkowego wymiaru j zapewnia większą elastyczność w modelowaniu, a co za tym idzie, również w estymacji.

Modele GLSM

Jednym z podejść zaproponowanych w odniesieniu do trójwymiarowej tabeli kontyngencji jest model GLSM, w którym estymacji podlega struktura asocjacji tabeli θ^Y obejmującej frakcje w ramach przekroju obszaru i warstwy, tj. $\theta_{ajk}^Y = Y_{aj}^{-1} Y_{ajk}$ (Zhang i Chambers, 2004). Propozycja ta bazuje na modelu GLSM postaci:

$$\mu_{ajk}^Y = \lambda_{jk} + \beta \mu_{ajk}^X, \quad \sum_{k=1}^K \lambda_{jk} = 0. \quad (3.31)$$

gdzie

$$\mu_{ajk}^Y = \log(\theta_{ajk}^Y) - K^{-1} \sum_l \log(\theta_{ajl}^Y) = \alpha_k^Y + \alpha_{jk}^Y + \alpha_{ak}^Y + \alpha_{ajk}^Y. \quad (3.32)$$

Składnik μ_{ajk}^X zdefiniowany jest analogicznie. Sumując stronami po a oraz j równanie (3.31), uzyskuje się następujące równania:

$$\alpha_k^Y = \bar{\lambda}_{.k} + \beta \alpha_k^X, \quad (3.33)$$

$$\alpha_{jk}^Y = (\lambda_{jk} - \bar{\lambda}_{.k}) + \beta \alpha_{jk}^X, \quad (3.34)$$

$$\alpha_{ak}^Y = \beta \alpha_{ak}^X, \quad (3.35)$$

$$\alpha_{ajk}^Y = \beta \alpha_{ajk}^X. \quad (3.36)$$

Składniki λ_{jk} można zawsze dobrać tak, aby przy danych równaniach (3.35)-(3.36) spełnione były równania (3.33)-(3.34). Zatem strukturalne założenie estymatora GSPREE dla tabeli trójwymiarowej określone jest ostatecznie przez równania (3.35)-(3.36), a model (3.31) określić można jako model proporcjonalnych interakcji – PIM.

Rozważa się również podejście, w którym zamiast przyjęcia wspólnego współczynnika proporcjonalności dla wszystkich warstw, dopuszcza się, aby β różniła się w zależności od kategorii zmiennej grupującej j . Model GLSM przyjmuje wtedy postać:

$$\mu_{ajk}^Y = \lambda_{jk} + \beta_j \mu_{ajk}^X, \quad \sum_{k=1}^K \lambda_{jk} = 0. \quad (3.37)$$

W takiej sytuacji trójwymiarową tabelę $\{Y_{ajk}\}$ podzielić można na J dwuwymiarowych tabel $\{Y_{a1k}\}, \dots, \{Y_{aJk}\}$. Niech

$$\{\alpha_{0|j}^Y, \alpha_{a|j}^Y, \alpha_{k|j}^Y, \alpha_{ak|j}^Y\} \quad (3.38)$$

będą parametrami modelu log-liniowego, które określają strukturę asocjacji tabeli dwuwymiarowej z warstwy j . Łatwo można wykazać relacje pomiędzy parametrami (3.38) modelu log-liniowego dla tabeli dwuwymiarowej, a parametrami modelu (3.22) dla tabeli trójwymiarowej, m.in.:

$$\alpha_{k|j}^Y = \alpha_k^Y + \alpha_{jk}^Y, \quad (3.39)$$

$$\alpha_{ak|j}^Y = \alpha_{ak}^Y + \alpha_{ajk}^Y \quad (3.40)$$

Podobną parametryzację można przeprowadzić w przypadku tabeli zastępczej X . Model (3.37) można wtedy przekształcić do następującego układu równań:

$$\alpha_{k|j}^Y = \lambda_{jk} + \beta_h \alpha_{k|j}^X, \quad (3.41)$$

$$\alpha_{ak|j}^Y = \beta_j \alpha_{ak|j}^X. \quad (3.42)$$

Z dowolności λ_{jk} wynika, że powyższy układ można sprowadzić do równania (3.42). Zatem (3.37) jest modelem proporcjonalnych interakcji – PIM, a bazujący na nim estymator GLSM spełnia strukturalne założenie w postaci równania (3.42).

Model GLSMM

Wśród podejść znajdujących zastosowanie w trójwymiarowej tabeli kontyngencji można wskazać także prezentowany wcześniej uogólniony model liniowy z efektami losowymi

(GLSMM) (Zhang i Chambers, 2004). Model ten w takiej sytuacji przyjmuje postać:

$$\mu_{ajk}^Y = \lambda_{jk} + \beta_j \mu_{ajk}^X + v_{ajk}, \quad (v_{aj2}, \dots, v_{ajk})^T \stackrel{IDD}{\sim} N(0, \Sigma_j), \quad v_{aj1} = -\sum_{l=2}^K v_{ajl}, \quad (3.43)$$

gdzie v_{ajk} oznacza składnik losowy.

W powyższym modelu możemy przyjąć, że β_j i Σ_j są równe we wszystkich warstwach, bądź mogą być zróżnicowane pomiędzy warstwami, a więc model (3.43) można rozpatrywać w czterech różnych konfiguracjach: (β_j, Σ_j) , (β_j, Σ) , (β, Σ_j) i (β, Σ) . W przypadku, gdy parametry β_j i Σ_j różnią się w zależności od j , otrzymujemy *warstwowany* model GLSMM. Jeżeli natomiast wybierzemy wariant, w którym parametry β_j i Σ_j nie są zależne od j , tj. (β, Σ) , wtedy model (3.43) sprowadzić można do równań:

$$\alpha_k^Y = \bar{\lambda}_{.k} + \beta \alpha_k^X + \bar{v}_{.k}, \quad (3.44)$$

$$\alpha_{jk}^Y = (\lambda_{jk} - \bar{\lambda}_{.k}) + \beta \alpha_{jk}^X + \bar{v}_{.jk} - \bar{v}_{.k}, \quad (3.45)$$

$$\alpha_{ak}^Y = \beta \alpha_{ak}^X + \bar{v}_{.ak} - \bar{v}_{.k}, \quad (3.46)$$

$$\alpha_{ajk}^Y = \beta \alpha_{ajk}^X + v_{ajk} - \bar{v}_{.ak} - \bar{v}_{.jk} + \bar{v}_{.k}. \quad (3.47)$$

Pomijając $\bar{v}_{.jk}$ oraz $\bar{v}_{.k}$, widzimy, że składnik losowy v_{ajk} można rozłożyć na dwa składniki: $w_{ak} = \bar{v}_{.ak}$ i $u_{ajk} = v_{ajk} - \bar{v}_{.ak}$, które wpływają odpowiednio na α_{ak}^Y i α_{ajk}^Y . Jednak w procesie estymacji $\{Y_{dhi}\}$ potrzebne są wartości v_{ajk} , więc efektywniej jest zastosować jeden składnik v_{ajk} zamiast dwóch niezależnych: w_{ak} i u_{ajk} .

3.6. Estymacja parametrów modeli GSPREE

3.6.1. Estymacja parametrów modelu GLSM dla dwuwymiarowej tabeli kontyngencji

Przyjmijmy następujące oznaczenia:

$$\theta_a := (\theta_{a1}^Y, \dots, \theta_{aK}^Y)^T, \quad (3.48)$$

$$\eta_a := (\mu_{a1}^Y, \dots, \mu_{aK}^Y)^T, \quad (3.49)$$

$$\mu_a := (\mu_{a1}^X, \dots, \mu_{aK}^X)^T, \quad (3.50)$$

$$\xi := (\lambda_2, \dots, \lambda_A, \beta)^T. \quad (3.51)$$

Model GLSM (3.13) zaproponowany przez Zhanga i Chambersa (2004) można wyrazić w zapisie macierzowym jako wielowymiarowy uogólniony model liniowy:

$$\eta_a = \mathbf{Z}_a \boldsymbol{\xi} =: \mathbf{g}(\theta_a), \quad (3.52)$$

gdzie konstrukcja macierzy \mathbf{Z}_a jest następująca:

$$\mathbf{Z}_a = \left(\mathbf{D} \mid \mu_a \right), \quad (3.53)$$

$$\mathbf{D} = \begin{pmatrix} -\mathbf{1}_{K-1}^T \\ \mathbf{I}_{(K-1) \times (K-1)} \end{pmatrix}. \quad (3.54)$$

Odwrotność funkcji wiążącej $\mathbf{g}(\theta_a)$ we wzorze (3.52) przyjmuje wtedy postać:

$$\theta_a = h(\eta_a) = (h_1(\eta_a), \dots, h_K(\eta_a)), \quad (3.55)$$

gdzie:

$$h_k(\eta_a) = \exp(\eta_{ak}) / \sum_{k'=1}^K \exp(\eta_{ak'}). \quad (3.56)$$

Niech $\tilde{\theta}_a$ będzie nieobciążonym bezpośrednim estymatorem szukanych frakcji θ_a . Załóżmy, że estymatory $\tilde{\theta}_a$ są niezależne pomiędzy obszarami i oznaczmy macierz ich kowariancji przez $G_a := \text{cov}(\tilde{\theta}_a, \tilde{\theta}_a)$.

Rozważmy linearyzację:

$$\mathbf{g}(\hat{\theta}_a) \approx \mathbf{g}(\theta_a) + \mathbf{g}'(\theta_a) (\hat{\theta}_a - \theta_a) =: \mathbf{u}_a, \quad a = 1, \dots, A, \quad (3.57)$$

gdzie $\mathbf{g}'(\theta_a)$ to macierz pochodnych cząstkowych funkcji $\mathbf{g}(\cdot)$ względem θ_a . Wektor \mathbf{u}_a zdefiniowany we wzorze (3.57) można wyrazić jako model liniowy postaci:

$$\mathbf{u}_a = \mathbf{Z}_a \boldsymbol{\xi} + \mathbf{e}_a, \quad (3.58)$$

$$\mathbf{e}_a = \mathbf{g}'(\theta_a) (\hat{\theta}_a - \theta_a), \quad (3.59)$$

gdzie

$$E(\mathbf{u}_a) = \mathbf{Z}_a \boldsymbol{\xi} = \mathbf{g}(\theta_a), \quad (3.60)$$

$$\text{Cov}(\mathbf{u}_a, \mathbf{u}_a) = \mathbf{g}'(\theta_a) \mathbf{G}_a [\mathbf{g}'(\theta_a)]^T =: \mathbf{V}_a. \quad (3.61)$$

Parametry modelu (3.58)-(3.59) można wtedy dopasować za pomocą iteracyjnej ważonej metody najmniejszych kwadratów (ang. *iterative weighted least squares* – IWLS).

Aktualizację wektora parametrów ξ w k -tym kroku procedury IWLS przeprowadza się następująco:

$$\xi^{(i)} = \left[\sum_{a=1}^A \mathbf{Z}_a^T \mathbf{V}_a^{(i-1)} \mathbf{Z}_a \right]^{-1} \sum_{a=1}^A \mathbf{Z}_a^T \mathbf{V}_a^{(i-1)} \mathbf{u}_a^{(i-1)}, \quad (3.62)$$

gdzie

$$\mathbf{V}_a^{(k-1)} = \mathbf{g}'(\theta_a^{(i-1)}) \mathbf{G}_a [\mathbf{g}'(\theta_a^{(i-1)})]^T, \quad (3.63)$$

$$\mathbf{u}_a^{(i-1)} = \mathbf{Z}_a \xi^{(i-1)} + \mathbf{g}'(\theta_a^{(i-1)}) (\tilde{\theta}_a - \theta_a^{(i-1)}), \quad (3.64)$$

przy czym wektor frakcji $\theta_a^{(i-1)}$ wyznacza się za pomocą wektora parametrów $\xi^{(i-1)}$ otrzymanego w poprzedniej iteracji algorytmu, tj:

$$\theta_a^{(i-1)} = \mathbf{h}(\mathbf{Z}_a \xi^{(i-1)}). \quad (3.65)$$

Po osiągnięciu zbieżności w ostatnim I -tym kroku algorytmu, otrzymuje się oszacowanie wektora parametrów $\hat{\xi} = \xi^{(I)}$ wraz z przybliżeniem macierzy kowariancji:

$$Cov(\hat{\xi}, \hat{\xi}) = \left(\sum_{a=1}^A \mathbf{Z}_a^T \mathbf{V}_a^{(I)} \mathbf{Z}_a \right)^{-1}. \quad (3.66)$$

3.6.2. Estymacja parametrów modelu GLSMM dla dwuwymiarowej tabeli kontyngencji

Model GLSMM dany wzorem (3.18) można zapisać w notacji macierzowej:

$$\mathbf{g}(\theta_a^Y) = \mathbf{Z}_a \xi + \mathbf{D} \mathbf{b}_a \quad (3.67)$$

gdzie $\mathbf{b}_a = (v_{a2}, \dots, v_{aK})^T$, natomiast \mathbf{Z}_a , ξ , \mathbf{D} zdefiniowane są wcześniej. Rozważmy mieszany model liniowy postaci:

$$\mathbf{u}_a = \mathbf{Z}_a \xi + \mathbf{D} \mathbf{b}_a + \mathbf{e}_a, \quad (3.68)$$

$$\mathbf{e}_a = \mathbf{g}'(\theta_a) (\tilde{\theta}_a - \theta_a), \quad (3.69)$$

gdzie \mathbf{b}_a i \mathbf{e}_a są niezależne.

Do szacowania parametrów modelu GLSMM, podobnie jak w przypadku GLSM, można zastosować metodę IWLS. W i -tym kroku tej metody aktualizowane są wektor parametrów, składniki losowe i macierz ich kowariancji:

$$\xi^{(i)} = \left[\sum_{a=1}^A \left(\mathbf{Z}_a^T \left(\mathbf{Q}_a^{(i-1)} \right)^{-1} \mathbf{Z}_a \right) \right]^{-1} \sum_{a=1}^A \left(\mathbf{Z}_a^T \left(\mathbf{Q}_a^{(i-1)} \right)^{-1} \mathbf{u}_a^{(i-1)} \right), \quad (3.70)$$

$$\mathbf{b}_a^{(i)} = \Sigma^{(i-1)} \mathbf{D}^T \left(\mathbf{Q}_a^{(i-1)} \right)^{-1} \left(\mathbf{u}_a^{(i-1)} - \mathbf{Z}_a \xi^{(i-1)} \right), \quad (3.71)$$

$$\Sigma^{(i)} = \mathbf{D}^{-1} \sum_a \left\{ \left(\mathbf{b}_a^{(i-1)} \right) \left(\mathbf{b}_a^{(i-1)} \right)^T + \left[\mathbf{D}^T \left(\mathbf{Q}_a^{(i-1)} \right)^{-1} \mathbf{D} + \left(\Sigma^{(i-1)} \right)^{-1} \right]^{-1} \right\}, \quad (3.72)$$

gdzie

$$\mathbf{Q}_a^{(i-1)} = \mathbf{D} \Sigma^{(i-1)} \mathbf{D}^T + \mathbf{g}' \left(\theta_a^{(i-1)} \right) \mathbf{G}_a \left[\mathbf{g}' \left(\theta_a^{(i-1)} \right) \right]^T, \quad (3.73)$$

$$\mathbf{u}_a^{(i-1)} = \mathbf{Z}_a \xi^{(i-1)} + \mathbf{D} \mathbf{b}_a^{(i-1)} + \mathbf{g}' \left(\theta_a^{(i-1)} \right) \left(\tilde{\theta}_a - \theta_a^{(i-1)} \right), \quad (3.74)$$

przy czym wektor frakcji $\theta_a^{(i-1)}$ wyznacza się za pomocą wektora parametrów $\xi^{(i-1)}$ i składników losowych $\mathbf{b}_a^{(i-1)}$ otrzymanych w poprzedniej iteracji algorytmu, tj:

$$\theta_a^{(i-1)} = \mathbf{h} \left(\mathbf{Z}_a \xi^{(i-1)} + \mathbf{D} \mathbf{b}_a^{(i-1)} \right). \quad (3.75)$$

W wyniku osiągnięcia zbieżności w ostatnim I -tym kroku algorytmu otrzymuje się oszacowanie frakcji $\hat{\theta}_a = \theta_a^{(I)}$, oszacowania wektora parametrów $\hat{\xi} = \xi^{(I)}$ wraz z macierzą kowariancji:

$$\text{Cov}(\hat{\xi}, \hat{\xi}) = \left(\sum_{a=1}^A \mathbf{Z}_a \left(\hat{\mathbf{Q}}_a \right)^{-1} \mathbf{Z}_a \right)^{-1}, \quad (3.76)$$

gdzie

$$\hat{\mathbf{Q}}_a = \mathbf{D} \hat{\Sigma} \mathbf{D}^T + \mathbf{g}' \left(\hat{\theta}_a \right) \mathbf{G}_a \left[\mathbf{g}' \left(\hat{\theta}_a \right) \right]^T, \quad (3.77)$$

a także oszacowania realizacji składnika losowego $\hat{\mathbf{b}}_a = \mathbf{b}_a^{(I)}$ wraz z macierzą kowariancji $\hat{\Sigma} = \Sigma^{(I)}$.

3.6.3. Estymacja parametrów modelu GLSM oraz GLSMM dla trójwymiarowej tabeli kontyngencji

Przywołując przykład przedstawiony przez Zhanga i Chambersa (2004) w celu oszacowania parametrów modeli GLSM (3.31) i GLSMM (3.43) w sytuacji, gdy parametry (β_j, Σ_j) są zależne od warstwy, stosuje się procedurę IWLS niezależnie dla każdej j -tej warstwy. Sprowadza się to do obliczeń dla tabeli dwuwymiarowej. W przypadku pozostałych modeli GLSM dla tabel trójwymiarowych, należy odpowiednio skonstruować komponenty θ_a

i Z_a , aby uzyskać postać (3.52), zaś w przypadku pozostałych modeli GLSMM dla tabel trójwymiarowych odpowiedniej modyfikacji podlegają dodatkowo komponenty \mathbf{d}_a i \mathbf{D} w celu sprowadzenia modelu do postaci (3.69). Proces estymacji parametrów przebiega według procedury IWLS opisanej dla przypadku tabel dwuwymiarowych.

Technika IPF w modelach GSPREE

Końcowym etapem estymacji w przypadku estymatora GSPREE jest uzyskanie oszacowań zgodnych ze znaną strukturą alokacji tabeli Y . W tym celu podobnie jak w przypadku klasycznego estymatora SPREE wykorzystuje się technikę IPF. Po estymacji parametrów modelu GLSM\GLSMM wyznacza się oszacowania frakcji wewnątrzobszarowych przekształcając funkcję wiążącą (3.14) i podstawiając dopasowane wartości $\hat{\mu}_{ak}^Y$ uzyskując w ten sposób następującą postać:

$$\hat{\theta}_{ak}^Y = \frac{\exp \hat{\mu}_{ak}^Y}{\sum_{l=1}^K \exp \hat{\mu}_{al}^Y}. \quad (3.78)$$

Za pomocą metody IPF przechodzi się następnie z oszacowań frakcji wewnątrzobszarowych $\hat{\theta}_{ak}^Y$ na oszacowania wartości globalnych \hat{Y}_{ak} , zgodnych ze znanymi wartościami brzegowymi wierszy $Y_{a.}$ i kolumn $Y_{.k}$ przy zachowaniu struktury asocjacji, która została dopasowana w oparciu o model GLSM/GLSMM.

W przypadku tabel trójwymiarowych po wyznaczeniu oszacowań frakcji $\hat{\theta}_{ajk}$ procedura IPF wykonywana jest niezależnie w ramach każdej z warstw $j = 1, \dots, J$, przy czym zapewniana jest zgodność ze znanymi wartościami brzegowymi $Y_{aj.}$ i $Y_{.jk}$.

3.7. Inne rozszerzenia estymatora typu SPREE

W literaturze przedmiotu można zauważyć ciągły rozwój metodyki związanej z estymacją typu SPREE. W odpowiedzi na pojawiające się problemy badawcze, głównie związane z jakością danych, proponowane są kolejne modyfikacje estymatora SPREE zarówno w ujęciu klasycznym, jak i modelowym. Przykładowo Zhang (2009) zaproponował modyfikację modeli GLSM oraz GLSMM, biorąc pod uwagę sytuację kiedy w badaniu reprezentacyjnym stanowiącym podstawę konstrukcji struktury alokacji w tym estymatorze występują braki danych. Zaproponowane podejście jest szczególnie cenione w obszarze badań

prowadzonych przez statystykę publiczną. Zjawisko braków odpowiedzi w badaniach reprezentacyjnych realizowanych przez krajowe urzędy statystyczne stanowi bowiem jedno z głównych źródeł błędów nielosowych mających wpływ na jakość estymacji.

Inną, przedstawioną w literaturze przedmiotu modyfikacją klasycznego estymatora SPREE, jest tzw. rozszerzony SPREE (ESPREE). Został on zaproponowany w pracach Isidro (2010) oraz Isidro i współautorów (2016). W estymacji ESPREE zakłada się, że szacowana zmienna, odnosząca się do ubóstwa, nie jest obserwowana w ramach badania spisowego, stąd metoda ta wychodząc od klasycznego SPREE przechodzi do wersji rozszerzonej, w której wykorzystuje się do uzyskania oszacowań, tzw. danych pseudospisowych (ang. *pseudo-census data*), pozyskiwanych w oparciu o metodę EBP (ang. *Empirical Best Prediction*) lub ELL⁹. Obie te metody prowadzą do integracji danych spisowych i danych pochodzących z badania reprezentacyjnego w celu uzyskania oszacowań parametrów dla małych obszarów. W podejściu tym przyjmuje się założenie, że dane do dwóch wykorzystywanych zbiorów – spisowego i pochodzącego z badania reprezentacyjnego – są gromadzone w tym samym czasie. Przejęcie takiego założenia jest szczególnie ważne z punktu widzenia obszarów zastosowań estymatora odnoszących się do statystyki społecznej. Zazwyczaj bowiem dane dotyczące dochodów, spożycia lub ubóstwa nie są obserwowane w spisach powszechnych. Specyfika generowania oszacowań w obydwu metodach EBP i ELL jest taka, że mamy do dyspozycji wiele replikacji, stąd uchyla się założenie o stałej populacji spisowej na rzecz przyjęcia założenia o tzw. superpopulacji. To z kolei umożliwia traktowanie standardowych błędów szacunków jako mierników obejmujących zmienność danych ze spisu i badania reprezentacyjnego.

Kolejną modyfikację estymacji SPREE zaproponowali Berg i Fuller (2014). W ich podejściu estymator SPREE oparty jest na nieliniowych modelach mieszanych. W szczególności brany pod uwagę jest tu model regresji logistycznej.

Dokonując przeglądu zaproponowanych w literaturze przedmiotu podejść związanych z estymacją typu SPREE warto również przytoczyć pracę doktorską Luny (2016). Zawarto w niej propozycję rozszerzenia estymatora GSPREE na wielowymiarowy SPREE (ang. *Multivariate Structure Preserving Estimator – MSPREE*), w którym współczynnik β może się różnić pomiędzy kategoriami zmiennych w ramach zdefiniowanych kolumn tabeli kon-

⁹Skrót pochodzi od pierwszych liter nazwisk trzech twórców metody – Chris Elbers, Jean Olson Lanjouw oraz Peter Lanjouw.

tyngencji. Przykładowo w przypadku rynku pracy współczynnik ten jest różny dla kategorii pracujący, bezrobotni oraz bierni zawodowo. Warto dodać, że wartość współczynnika beta jest stała dla wszystkich obszarów zdefiniowanych w wierszach tabeli kontyngencji. Praca Luny (2016) zawiera również propozycję wielowymiarowego SPREE z efektami mieszanymi (ang. *Mixed MSPREE* – MMSPREE), który stanowi rozszerzenie MSPREE i bierze dodatkowo pod uwagę efekt losowy na poziomie każdej pojedynczej komórki stanowiącej środek tabeli kontyngencji.

Estymację typu SPREE poruszono także w pracy Dostała i współpracowników (2016). W celu uzyskania oszacowań w ramach podgrup populacji w niemieckiej edycji spisu 2011 rozbudowano podejście SPREE o alternatywną funkcję odległości χ^2 . W artykule wykazano, że jej zastosowanie poprawia jakość szacunków w ramach estymacji dla małych domen. Jest to modyfikacja klasycznego podejścia, w którym jak podkreślili Ireland oraz Kullback (1968), minimalizuje się tzw. funkcję dywergencji informacyjnej w ramach działania algorytmu IPF¹⁰.

Na koniec rozważań przedstawionych w tym podrozdziale warto wspomnieć o jednej z najnowszych propozycji modyfikacji estymacji typu SPREE dotyczącej wykorzystania w szacunku alternatywnych źródeł zmiennych pomocniczych. Podejście to zostało opisane w artykule (Koebe i in., 2021). Zaproponowano w nim, aby w latach międzypisowych włączyć do estymacji parametrów kluczowych statystyk dla małych obszarów informacje pomocnicze pochodzące ze zdjęć satelitarnych. Główna idea podejścia sprowadza się do aktualizacji rozkładów brzegowych dla populacji. Rozważane podejście zostało zilustrowane na przykładzie szacunku wielowymiarowego wskaźnika ubóstwa (ang. *multidimensional poverty index* – MPI). Chodziło szczególnie o aktualizację rozkładów brzegowych, szczególnie w regionach charakteryzujących się silnym wzrostem ludności poprzez aktualizację rozkładów brzegowych dla populacji.

¹⁰Więcej szczegółów można znaleźć w artykule Wu i Lu (2016).

3.8. Przegląd praktycznych zastosowań estymatorów typu SPREE

Niniejszy podrozdział zawiera przykłady praktycznych zastosowań estymatorów typu SPREE w odniesieniu do obszaru rynku pracy.

3.8.1. Australia

Metody estymacji typu SPREE w ujęciu klasycznym są stosowane w praktyce badań statystycznych między innymi przez australijską agencję rządową National Skills Commission¹¹, która dostarcza informacji o sytuacji na australijskim rynku pracy. Dzięki wykorzystaniu klasycznego estymatora typu SPREE publikowane są oszacowania o liczbie osób bezrobotnych, stopie bezrobocia oraz osobach aktywnych ekonomicznie na poziomie małych obszarów, którymi są tzw. statystyczne obszary poziomu 2 (ang. *Statistical Areas Level 2 – SA2*) grupujące społeczności powiązane ze sobą gospodarczo¹². W celu uzyskania szacunków wykorzystywane są rejestry administracyjne: Aktualni odbiorcy zasiłku dla młodych (ang. *Youth Allowance*) oraz Aktualni odbiorcy zasiłku „Na dobry początek” \Dla osób poszukujących pracy (ang. *Newstart Allowance* \ *JobSeeker Payment*). Dodatkowo uwzględnia się także informacje pochodzące z badania Labour Force Survey (LFS) oraz dane spisowe na temat odpowiednich kategorii rynku pracy na poziomie obszarów statystycznych SA2. Celem implementacji estymatora typu SPREE jest uzyskanie szacunków charakterystyk rynku pracy dla małych obszarów, zgodnych z aktualnymi szacunkami z badania LFS tworzącego strukturę alokacji konstruowanych tabel. Informacje pomocnicze z rejestrów oraz spisu powszechnego wykorzystywane są do ustanowienia struktury asocjacji dla konstruowanych tabel. Szczegóły w zakresie wyznaczania tabel dla poszczególnych kategorii rynku pracy oraz wskaźników zamieszczono w publikacji pt. *Small Area Labour Markets (2021)*.

¹¹<https://www.nationalskillscommission.gov.au/>

¹²Podział Australii na wspomniane obszary SA2 można znaleźć pod linkiem <https://dbr.abs.gov.au/absmaps/index.html>.

3.8.2. Wielka Brytania

Urząd Statystyczny Wielkiej Brytanii przygotowując się do przeprowadzenia spisów powszechnych w roku 2021 w ramach prac przygotowawczych podjął działania w celu przetestowania estymatora GSPREE, będącego rozszerzeniem estymatora typu SPREE. Umożliwia on prowadzenie szacunków, w sytuacji gdy niespełnione jest założenie o aktualności rozkładów spisowych oraz w procesie estymacji można wykorzystać dodatkowe zmienne pomocnicze. Pierwsze prace w tym zakresie podjęto w ramach współpracy Urzędu Statystycznego Wielkiej Brytanii oraz Uniwersytetu Southampton, por. (Luna i in., 2014). Szacowaną zmienną była liczba ludności według przynależności etnicznej w podziale na grupy wieku, w ramach jednostek przestrzennych zdefiniowanych jako tzw. *local authority districts* – LAD, które odpowiadają poziomowi NUTS 3. Uzyskanie takich bieżących informacji w Wielkiej Brytanii nie jest możliwe na podstawie żadnego z dostępnych źródeł danych z osobna. Dlatego też podjęto prace testowe w zakresie zastosowania estymatora typu GSPREE. Celem było uzyskanie aktualnych danych w tabeli kontyngencji dla 2014 roku na podstawie informacji z czterech źródeł danych. Jednym ze źródeł było doroczne badanie ludności (ang. *Annual Population Survey* – APS). Jest ono tak zaprojektowane, aby dostarczyć informacji o zmiennych demograficznych i społeczno-ekonomicznych na poziomie lokalnym. Prowadzone jest na podstawie próby liczącej około 250 tysięcy osób. Składa się na nie badanie LFS oraz dodatkowa próba (ang. *boost sample* – ELLFS), która ma na celu zapewnienie minimalnej wielkości próby ludności czynnej ekonomicznie na poziomie samorządu lokalnego odpowiedzialnego za edukację (ang. *local education authority*). Liczebność próby jest jednak często zbyt mała (czasami nawet zerowa), aby dostarczyć wiarygodnych szacunków na poziomie lokalnym, w przekroju wszystkich grup etnicznych. Kolejne źródło danych stanowiły bezpośrednie szacunki śródroczne (ang. *MidYear Estimates* – MYE) liczby ludności w poszczególnych grupach wieku na poziomie LAD dla 2014 roku. Otrzymywane są one za pomocą metody składników kohortowych (ang. *cohort component method*), w której aktualizacja szacunków spisowych odbywa się poprzez uwzględnienie składników zmian demograficznych (liczba urodzeń, zgonów oraz migracje). Następnym źródłem informacji był coroczny spis szkolny (ang. *English School*

Census – ESC), obejmujący populację osób w wieku 5 do 15 lat¹³. Uwzględniając opisane powyżej dwa źródła danych dla osób w wieku szkolnym rozważono trzy modele z efektami mieszanymi o rosnącym stopniu złożoności. W pierwszym modelu wzięto pod uwagę jedynie dane pochodzące ze spisu powszechnego przeprowadzonego w 2011 r., rozważanego jako jedyne źródło danych (szacunki w przekroju LAD x przynależność etniczna). W drugim modelu uwzględniono dezagregację na grupy wieku w przekroju 0—4, 5—15, 16 i więcej lat. Natomiast trzeci model był zbliżony do modelu drugiego i łączył dwa źródła danych tj. spis powszechny z 2011 roku oraz coroczny spis szkolny dla 2014 roku. Otrzymane na podstawie rozpatrywanych modeli oszacowania charakteryzowały się podobnymi niskimi wartościami obciążenia i błędu średniokwadratowego dla większości obszarów LAD. Jednak uzyskane wyniki walidacji modelu wskazywały, że zastosowanie najbardziej złożonego – trzeciego modelu charakteryzuje się najlepszym dopasowaniem, choć szacowane proporcje wewnątrz małych obszarów są podobne zarówno w drugim, jak i trzecim modelu (Correa-Onel i in., 2016).

3.8.3. Nowa Zelandia

Praktyczne zastosowanie estymatora typu SPREE oraz prace testowe w zakresie wykorzystania modelu log-linowego w Nowej Zelandii opisano w artykule Qiao (2005). Urząd Statystyczny Nowej Zelandii (ang. *Statistics New Zealand* – SNZ) jest zobligowany do monitorowania i dostarczania informacji o zjawiskach społeczno-ekonomicznych i monitorowania zmian w czasie na poziomie małych obszarów (rozumianych jako poziom terytorialny – ang. *Territorial Local Authorities* – TLAs) i domen w ramach przekroju terytorialnego TLA – płeć\grupa wieku\przynależność etniczna. Jedną z pożądaných informacji jest liczba osób bezrobotnych. SNZ wykorzystuje estymator typu SPREE, aby dopasować dane wewnątrz tabeli kontyngencji na temat bezrobocia rejestrowanego (dane Ministerstwa Rozwoju Społecznego), tak żeby były one zgodne z bardziej wiarygodnymi i aktualnymi wartościami brzegowymi tabeli pochodzącymi z badania *Household Labour Force Survey* (HLFS). Badanie HLFS swoim zakresem w przybliżeniu odpowiada prowadzonemu w Polsce Badaniu Aktywności Ekonomicznej Ludności. Pierwszym źródłem danych,

¹³Źródło to zawiera prawie całkowity stopień pokrycia populacji w wieku 5—15 lat w szkołach państwowych, z wyjątkiem szkół niepaństwowych gdzie odsetek pokrycia kształtuje się na poziomie około 7% populacji uczniów uczęszczających do szkół prywatnych.

stanowiącym strukturę asocjacyjną tabeli, są dane na temat liczby zarejestrowanych bezrobotnych osób, których dostarcza Ministerstwo Rozwoju Społecznego. Dane dostępne są dla wszystkich wymaganych przekrojów, natomiast ich definicja nie jest zgodna z przyjętą przez Międzynarodową Organizację Pracy. Dla różnych subdomen populacji wykorzystuje się korelacje pomiędzy obu źródłami. Jest to podstawa do połączenia obydwu źródeł. Drugim źródłem danych stanowiącym strukturę alokacyjną tabeli kontyngencji, jest badanie reprezentacyjne HLFS, które oparte jest na próbie około 30 tysięcy osób. Próba przyjęta w badaniu umożliwia dostarczanie wiarygodnych bezpośrednich szacunków w 10 z 16 regionów – *Regional Councils* – RC. Jest ona jednak niewystarczająca dla zdefiniowanych wcześniej małych obszarów, ze względu na zbyt duży błąd szacunku. Wykorzystując procedurę iteracyjnego proporcjonalnego dopasowania wartości wewnątrz tabeli kontyngencji dostosowywane są do zdefiniowanych wartości brzegowych poprzez kolejne iteracje, aż do uzyskania pełnej zgodności. Implementacja klasycznego estymatora typu SPREE nie daje jednak możliwości dołączenia dodatkowych zmiennych umożliwiających poprawę jakości modelu. Brak jest również diagnostyk oceniających jakość uzyskanego dopasowania struktury alokacyjnej do asocjacyjnej tabeli kontyngencji. Prowadzone są jednak dalsze prace w tym zakresie. Testowaniu poddaje się kolejne zmodyfikowane podejścia.

3.8.4. Kanada

Praktyczne zastosowanie estymatora typu SPREE zaproponowano w opracowaniu Hidioglou i Pataka (2009). Opisano w nim prace zrealizowane w 2008 roku w ramach studium wykonalności, podjęte na zlecenie Agencji Zasobów Ludzkich i Rozwoju Umiejętności (*Human Resources and Skills Development Canada*). W pracach tych chodziło o poprawę jakości szacunków w odniesieniu do danych o wykonywanych zawodach (na poziomie kodów 3–cyfrowych lub 4–cyfrowych) w przekroju prowincji. Badanie Labour Force Survey (LFS) prowadzone przez Kanadyjski Urząd Statystyczny dostarcza, między innymi, bezpośrednich szacunków o wykonywanych zawodach (na poziomie kodów 2–cyfrowych) w przekroju prowincji. Szacunki dla kodów 3–cyfrowych i 4–cyfrowych nie są publikowane ze względu na niską precyzję uzyskiwanych wyników estymacji. Informacje tego typu dostarczane są przez Kanadyjski Urząd Statystyczny z akceptowalną precyzją jedynie w ramach narodowych spisów powszechnych, które odbywają się co 5 lat. W ramach prowadzonych prac badawczych wykorzystano dane pochodzące ze spisów przeprowadzo-

nych w 2001 oraz 2006 roku, a także dane z badania LFS prowadzonego od maja 2001 do maja 2006 roku. Dokonano oszacowań mnożąc liczebności dla kodów 2-cyfrowych przez proporcję owych liczebności uzyskanych za pomocą opracowanej procedury. Procedura ta składała się z kilku kroków. Pierwszym z nich było oszacowanie liczebności wykonywanych zawodów dla kodów 3-cyfrowych za pomocą metody SPREE, polegającej na dostosowaniu struktury asocjacyjnej zawierającej informacje ze spisu ludności do aktualnych danych w strukturze alokacji (aktualnych oszacowań na podstawie danych z badania LFS), przy zachowaniu relacji pomiędzy zmiennymi w strukturze asocjacyjnej (ustalonych na podstawie danych spisowych). Następnie oszacowano liczebności kodów 3-cyfrowych w ramach poszczególnych kodów 2-cyfrowych na podstawie danych spisowych, aby dostosować je do liczebności dla kodów 2-cyfrowych w przekroju prowincji oraz kodów 3-cyfrowych na poziomie kraju oszacowanych na podstawie LFS. Kolejnym krokiem było połączenie proporcji uzyskanych za pomocą estymatora SPREE z proporcjami wynikającymi z badania LFS. Ostatnim krokiem było pomnożenie połączonych proporcji przez liczebności kodów 2-cyfrowych dla każdej prowincji. Na podstawie przeprowadzonych prac badawczych sformułowano kilka wniosków. Podkreślono między innymi, że rzetelność uzyskiwanych szacunków na podstawie badania LFS dla niektórych zawodów w ramach kodów 2-cyfrowych była względnie niska. To z kolei powodowało niską precyzję szacunków przy wykorzystaniu estymatora typu SPREE dla kodów 3-cyfrowych. Zaznaczono również, że uzyskane liczebności nie różnią się zbyt wiele od szacunków bezpośrednich, ale ich precyzja jest znacznie wyższa, szczególnie dla małych prowincji. Zauważono również, że zysk na precyzji dla liczebności jest znacznie niższy niż zysk dla proporcji, ponieważ ograniczają go wartości odpowiednich liczebności brzegowych.

3.8.5. Polska

Również w Polsce, w ramach projektu Eurostatu pt. Poprawienie jakości unijnych spisów (2021 i po roku 2021), realizowanego przez Główny Urząd Statystyczny, Urząd Statystyczny w Poznaniu oraz pracowników Katedry Statystyki Uniwersytetu Ekonomicznego w Poznaniu, została podjęta próba implementacji estymatorów typu SPREE w obszarze niepełnosprawności. Podczas prowadzonych prac badawczych¹⁴ wykorzystano dane pochodzące z Narodowego Spisu Powszechnego Ludności i Mieszkań 2011, które wykorzystano

¹⁴Autor niniejszej rozprawy doktorskiej był również członkiem zespołu projektowego.

do konstrukcji struktury asocjacji w rozważanych tabelach kontyngencji (Szymkowiak i in., 2017a). Z kolei informacje z badania BAEL dotyczące osób z niepełnosprawnościami na rynku pracy odpowiadały strukturze alokacji. Podczas przeprowadzonych prac zrealizowano między innymi następujące zadania:

- dokonano estymacji liczby osób z niepełnosprawnościami z uwzględnieniem statusu na rynku pracy i w przekroju domen o większej szczegółowości niż publikowane na podstawie bieżących badań reprezentacyjnych tj. według województw i miejsca zamieszkania oraz według województw i płci;
- zaimplementowano podstawowe konstrukcje estymatora typu SPREE w ujęciu modelowym dla tabel dwuwymiarowych (Zhang i Chambers, 2004);
- przeprowadzono badania symulacyjne celem oceny własności wybranych estymatorów typu SPREE bazujące na danych z NSP 2011 oraz z badania BAEL w zakresie informacji dotyczącej niepełnosprawności.

Uzyskane wyniki estymacji zapewniły zgodność struktury alokacji ze strukturą asocjacji w analizowanych tabelach. Wskazano jednocześnie, że możliwość takiej aktualizacji bieżących informacji z badania reprezentacyjnego jest coraz bardziej ograniczona ze względu na stopniową dezaktualizację informacji spisowych wraz z upływem czasu od momentu jego przeprowadzenia. Wpływa to niewątpliwie na wiarygodność uzyskanych oszacowań dla lat popispowych. W tym kontekście należałoby rozważyć wykorzystanie danych z bieżących rejestrów dotyczących osób z niepełnosprawnościami w miejsce struktury alokacji rozważanych tabel kontyngencji. W konkluzjach i rekomendacjach pracy badawczej podkreślono również, że zasadne byłoby przetestowanie większego zakresu modeli typu SPREE, w tym w ujęciu tabel trójwymiarowych.

3.9. Podsumowanie

Przedstawiona w rozdziale charakterystyka statystyki małych obszarów, ze szczególnym uwzględnieniem estymacji typu SPREE, oraz przegląd zastosowań różnych podejść estymatora SPREE wskazują, iż ten rodzaj estymacji może stać się przedmiotem rozważań w praktyce badań statystycznych. Estymacja typu SPREE charakteryzuje się bowiem własnościami, które są pożądane i umożliwiają estymację na niskim poziomie agregacji

przestrzennej i dla szczegółowo zdefiniowanych domen, między innymi w odniesieniu do rynku pracy.

Bazując na przeglądzie literatury z zakresu metod estymacji typu SPREE podjęto decyzję, by w empirycznej części rozprawy doktorskiej zaproponować własne rozwiązania metodyczne odwołujące się do podejścia modelowego z uwzględnieniem wariantów GLSM oraz GLSMM zaproponowanych przez Zhanga i Chambersa (2004). W kolejnym rozdziale zostanie zaprezentowana autorska procedura badawcza łącznie z egzemplifikacją wybranych podejść w procesie estymacji wybranych charakterystyk rynku pracy w kontekście miejskich obszarów funkcjonalnych.

Rozdział 4

Rynek pracy na obszarach funkcjonalnych miast wojewódzkich - procedura badawcza

W niniejszym rozdziale zaproponowano autorską procedurę badawczą. Jej zastosowanie pozwala na dokonanie szacunków dla unikalnego poziomu agregacji przestrzennej – miejskich obszarów funkcjonalnych – z wykorzystaniem estymacji typu SPREE. W opisie procedury skoncentrowano się na elementach, które wyróżnione zostały w ramach fazy przygotowawczej takich jak: zakres potencjalnych źródeł danych planowanych do wykorzystania w badaniu oraz zakres analizowanych zmiennych, rekomendowane warianty technik estymacji pośredniej wraz z proponowanymi metodami oceny jakości szacunku, rozważane przekroje i poziom agregacji jednostek, dla których ma być prowadzony szacunek. Zweryfikowano pierwszą hipotezę badawczą, że estymacja typu SPREE pozwala na oszacowanie podstawowych charakterystyk rynku pracy na poziomie agregacji przestrzennej, tj. dla MOF OW, dla których obecnie nie ma pokrycia informacyjnego. Szerszej analizie poddano ocenę jakości oszacowań, w tym szczególnie precyzję szacunków liczby osób pracujących, bezrobotnych oraz biernych zawodowo otrzymanych na podstawie rozważa-

nych wariantów estymatorów. Pozwoliło to na zweryfikowanie drugiej hipotezy badawczej, że estymacja typu SPREE cechuje się lepszą precyzją szacunków charakterystyk rynku pracy w porównaniu z estymacją kalibracyjną stosowaną w badaniu BAEL. Na podstawie przeprowadzonej oceny jakości oszacowań według przyjętych kryteriów dokonano wyboru estymatora, który posłużył w rozdziale piątym do wyznaczenia podstawowych wskaźników dotyczących rynku pracy (współczynnika aktywności zawodowej, wskaźnika zatrudnienia oraz stopy bezrobocia).

4.1. Założenia procedury badawczej

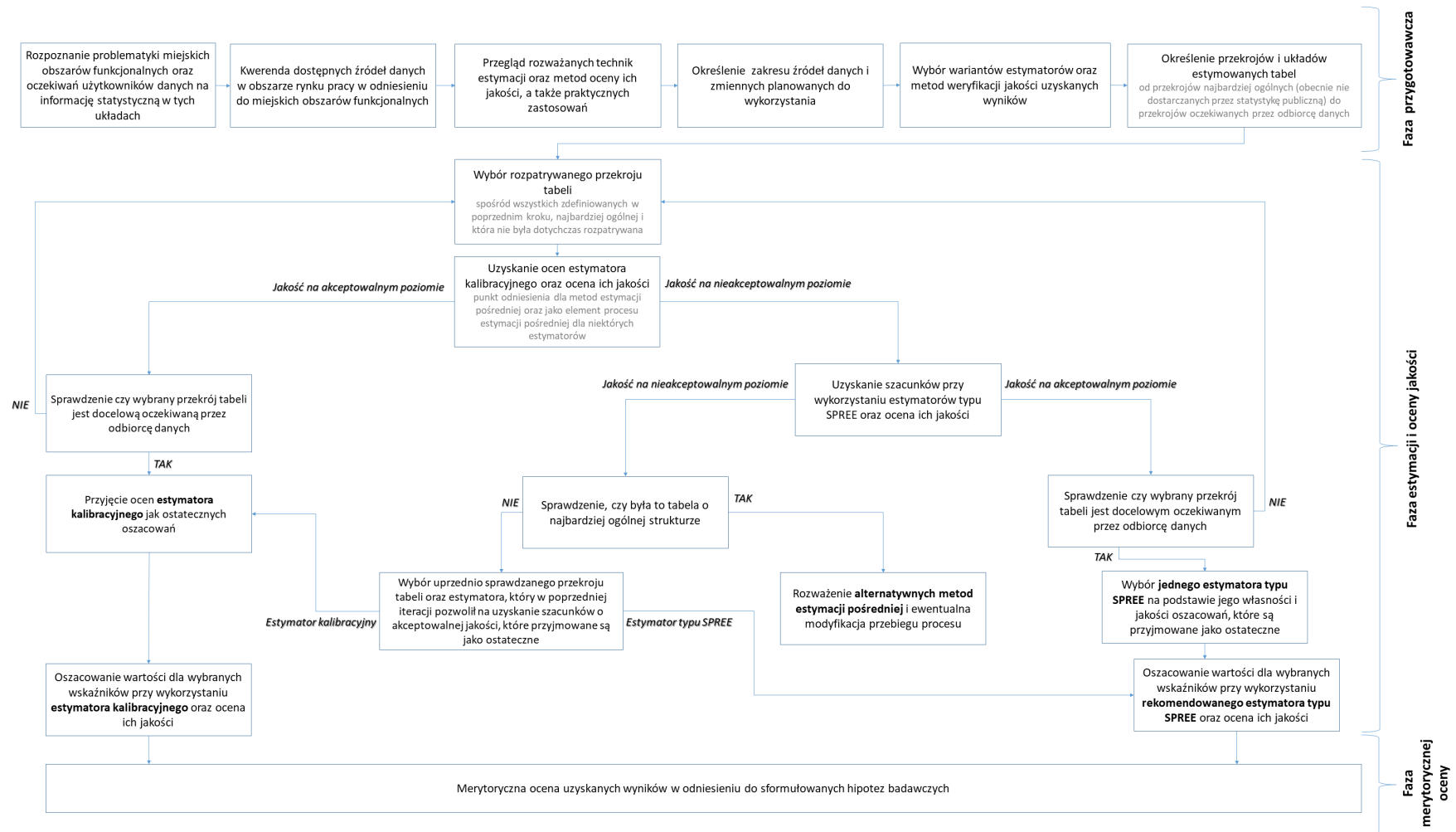
W ramach przeprowadzonego w pracy badania empirycznego, opierając się na wynikach analiz przedstawionych dotychczas w literaturze przedmiotu, zaproponowano autorską procedurę badawczą. Jej celem jest uzyskanie oszacowań współczynnika aktywności ekonomicznej ludności, wskaźnika zatrudnienia oraz stopy bezrobocia na poziomie miejskich obszarów funkcjonalnych wyznaczonych do realizacji instrumentu polityki spójności, jakim są Zintegrowane Inwestycje Terytorialne, w podziale na płeć i przyjęte grupy wieku. Schemat autorskiej procedury badawczej przedstawiony został na rysunku 4.1. Przy opracowywaniu procedury brano pod uwagę wskazówki i rekomendacje wypracowane w ramach projektu ESSnet on SAE (2012) oraz podejście zaproponowane w artykule Tzavidisa i współautorów (2018).

Powyższy schemat został podzielony na trzy fazy, które składają się z poszczególnych elementów określających wykonane prace teoretyczne i empiryczne. Pierwsze trzy elementy fazy przygotowawczej zaprezentowane na powyższym schemacie zostały opisane w trzech pierwszych rozdziałach niniejszej rozprawy doktorskiej.

Początkowym krokiem, opisanym szczegółowo w rozdziale pierwszym, było rozpoznanie tematyki miejskich obszarów funkcjonalnych, w tym między innymi, różnych definicji, sposobów ich wyznaczania oraz ich umiejscowienia w kontekście polityki krajowej i europejskiej. Przeanalizowano również kwestię zapotrzebowania użytkowników danych na informacje statystyczne w układzie wskazanych wyżej jednostek przestrzennych, w tym dotyczącego rynku pracy. W wyniku tych działań stwierdzono, że istnieje zapotrzebowanie na informacje na temat miejskich obszarów funkcjonalnych wyznaczonych do realizacji

Zintegrowanych Inwestycji Terytorialnych w latach 2014–2020. W oparciu o przeprowadzone analizy przyjęto, iż szczególnie poszukiwanymi informacjami o rynku pracy są dane na temat stopy bezrobocia, wskaźnika zatrudnienia oraz współczynnika aktywności ekonomicznej, w podziale na płeć oraz cztery grupy wieku (tj. 15–29, 30–39, 40–49, 50 i więcej).

Kolejnym krokiem w procedurze badawczej było sprawdzenie, czy w obecnie dostępnych źródłach danych znajdują się informacje niezbędne do wyznaczenia wyżej wymienionych wskaźników we wskazanych przekrojach. Chodziło między innymi o informacje o liczbie pracujących, bezrobotnych oraz biernych zawodowo zgodną z definicją przyjętą w BAEL.



Rysunek 4.1: Schemat procedury badawczej przeprowadzonej w ramach niniejszej rozprawy

Źródło: Opracowanie własne.

Ten etap opisany został w rozdziale drugim. Przeprowadzona kwerenda i analiza porównawcza źródeł danych wykazała, iż żadne z obecnie dostępnych źródeł nie umożliwia uzyskania informacji o rynku pracy w układzie miejskich obszarów funkcjonalnych. Analiza umożliwiła jednak wytypowanie w ramach ocenianych źródeł grupy zmiennych, które mogłyby być użyte w charakterze zmiennych pomocniczych, stosowanych w statystyce małych obszarów (SMO).

W związku z tym podjęto decyzję o włączeniu do badania metod SMO wykorzystujących estymację pośrednią. Estymacja pośrednia pozwala na połączenie informacji z kilku źródeł w ramach modelu statystycznego w celu oszacowania określonych parametrów. Przegląd rozważanych metod estymacji oraz ocen ich jakości, a także zaleceń i wskazówek co do przeprowadzenia procesu estymacji pośredniej został przedstawiony w rozdziale trzecim. Ze względu na opisane w rozdziale trzecim własności poszczególnych metod szacunku zdecydowano się na wykorzystanie w badaniu empirycznym estymatorów GLSM oraz GLSMM zaproponowanych przez Zhanga i Chambersa (2004).

Kolejny krok ujęty w schemacie obejmuje wskazanie planowanych do wykorzystania źródeł danych (podrozdział 4.2), określenie zakresu szacowanych parametrów oraz wybór miejskich obszarów funkcjonalnych, na poziomie których przeprowadzony zostanie proces estymacji (podrozdział 4.3). Etap ten został szczegółowo opisany w niniejszym rozdziale. Kolejny etap odnosi się do wyboru metody estymacji ze wskazaniem na konkretne warianty estymatorów oraz wybór sposobu oceny jakości uzyskanych oszacowań (podrozdział 4.4). W dalszej kolejności ustalone zostały przekroje tabel, dla których dokonano oszacowania absolutnych wartości parametrów takich jak: liczba osób pracujących, bezrobotnych oraz biernych zawodowo. Przekroje tabel określone zostały od najbardziej ogólnego poziomu agregacji jednostek do docelowego. Oszacowanie wartości absolutnych (w postaci wartości globalnych) jest konieczne do tego, aby w dalszej kolejności wyznaczyć wielkości względne - odpowiednie wskaźniki w zakładanych przekrojach. Możliwe scenariusze realizacji procesu estymacji tak pośredniej, jak i kalibracyjnej w obszarze rynku pracy zostały zaprezentowane na schemacie przedstawionym na rysunku 4.1 w ramach fazy estymacji i oceny jakości. Ostatnią fazą procedury postępowania w przypadku uzyskania akceptowalnej jakości oszacowań liczby osób pracujących, bezrobotnych oraz biernych zawodowo w określonym przekroju jest obliczenie wartości wytypowanych wskaźników dla tego przekroju wraz z oceną jakości uzyskanych oszacowań. Końcowym elementem całej procedury

badawczej była interpretacja i merytoryczna ocena uzyskanych oszacowanych wskaźników w odniesieniu do sformułowanych hipotez badawczych.

4.2. Opis zakresu wykorzystanych źródeł danych

Kwerenda źródeł danych przeprowadzona w rozdziale drugim wykazała, że Badanie Aktywności Ekonomicznej Ludności (BAEL) dostarcza bogatej informacji o wielu kategoriach wyróżnionych w ramach rynku pracy, pożądanych także z punktu prowadzenia niniejszego badania empirycznego. Chodzi tu nie tylko o rodzaj informacji, ale również zgodności definicyjne. Informacje te nie są jednak dostępne na poziomie miejskich obszarów funkcjonalnych. Jedynym źródłem zawierającym informacje o liczbie osób pracujących i bezrobotnych w analizowanych tu przekrojach, które stanowią składowe do wyliczanych wskaźników, jest źródło ZROPiB (Zintegrowany Rejestr Osób Pracujących i Bezrobotnych). Przyjęte w nim definicje zmiennych dotyczących osób pracujących i bezrobotnych są inne niż w badaniu BAEL. Z kolei wykazana korelacja pomiędzy zmiennymi w badaniu BAEL oraz w rejestrze ZROPiB (por. 2.6) daje podstawy, aby zmienne z tego rejestru wykorzystać w charakterze zmiennych pomocniczych w procesie estymacji typu SPREE. Dodatkowo mając informację o liczbie i strukturze ludności w wymaganych przekrojach na podstawie bieżących bilansów ludności oraz wykorzystując dane ze źródła ZROPiB, możliwe jest wyliczenie zmiennej zastępczej odpowiadającej liczbie osób biernych zawodowo.

Biorąc pod uwagę przyjęty w estymacji typu SPREE podział tabeli kontyngencji na strukturę asocjacji oraz strukturę alokacji (por. 3.3) poniżej opisano wykorzystane źródła danych w procesie estymacji.

4.2.1. Struktura asocjacji

Do ustalenia rozkładów warunkowych tabeli kontyngencji w ramach prowadzonych analiz korzystano z trzech źródeł danych (BAEL, ZROPiB oraz Bank Danych Lokalnych (BDL) na temat liczby i struktury ludności). Liczba i zakres wykorzystanych źródeł są uzależnione od przyjętego w przeprowadzonych pracach empirycznych wariantu estymatora.

Zmienne zastępcze

Zmienne zastępcze są wykorzystywane w procesie estymacji we wszystkich rozpatrywanych wariantach estymatora GLSM oraz GLSMM. Jak wspomniano w rozdziale trzecim, zmienna zastępcza jest w dużej mierze zgodna ze zmienną oryginalną (w tym przypadku z odpowiednimi kategoriami rynku pracy uzyskanymi z badania BAEL). Jednak nie pokrywa ona dokładnie tej samej zbiorowości co zmienna oryginalna. Powodem mogą być różnice w definicjach zmiennych lub różny okres referencyjny dla zmiennej oryginalnej i zastępczej. Podczas prowadzonych prac do środka analizowanych tabel kontyngencji wstawiono dane na temat liczby osób pracujących oraz bezrobotnych ze źródła ZROPiB (dane średnioroczne dla 2017 roku)¹. Z kolei zmienną zastępczą odpowiadającą liczbie osób biernych zawodowo wyznaczono na podstawie informacji o liczbie i strukturze ludności faktycznie zamieszkałej w Polsce, w odpowiednich przekrojach z BDL oraz informacji pochodzącej z ZROPiB. Wielkości te policzono jako średnią arytmetyczną z liczby ludności faktycznie zamieszkałej w wieku 15 lat i więcej według stanu na 30 VI oraz 31 XII 2017 roku. Przeliczenia takiego dokonano w celu przybliżenia poziomu zmiennej do przeciętnej wartości liczby ludności w 2017 roku. Następnie informacje te posłużyły do wyznaczenia liczby osób biernych zawodowo. Obliczono ją jako różnicę zdefiniowanej wyżej liczby ludności oraz liczby osób pracujących i bezrobotnych określonej w oparciu o ZROPiB. Otrzymane wielkości wstawiono do środka tabeli kontyngencji. Również w badaniu BAEL pozostała część ludności, po skorygowaniu o liczbę osób pracujących i bezrobotnych, zaliczana jest do osób biernych zawodowo. Warto dodać, że badana populacja w BAEL nie pokrywa się w pełni ze strukturami ludności zawartymi w BDL. Dane o stanie i strukturze ludności zamieszczone w BDL oraz te służące do wyznaczenia wag finalnych w BAEL opracowuje się przy zastosowaniu metody bilansowej². Z kolei dane ludnościowe z BDL, obejmują w odróżnieniu od tych wykorzystywanych BAEL³, również osoby w gospodarstwach zbiorowego zakwaterowania oraz osoby bezdomne, a także stałych mieszkańców Polski przebywających czasowo za granicą bez względu na okres ich nieobecności. Nie obejmują jednak osób przybyłych z zagranicy na pobyt czasowy bez względu na okres ich

¹Metodyka pozyskania tych danych opisana jest w rozdziale 2, w podpunkcie 2.1. Natomiast analiza porównawcza pomiędzy BAEL, a ZROPiB znajduje się w podpunkcie 2.4.

²Punktem wyjścia są w niej wyniki ostatniego spisu powszechnego, podlegające korekcie na podstawie bieżących informacji o ruchu naturalnym w zakresie urodzeń i zgonów, a także o migracji na pobyt stały i czasowy.

³Syntetyczny opis ludności podlegającej badaniu BAEL znajduje się w rozdziale 2 w podpunkcie 2.1., natomiast szerszy opis dotyczący tego zagadnienia znajduje się w publikacji GUS (2019c).

czasowego przebywania. Różnice te powodują, że zmienną „osoby biernie zawodowo” wyznaczoną na podstawie danych ludnościowych z BDL oraz informacji ze ZROPiB można traktować jako zmienną zastępczą. W niniejszej rozprawie jest ona oznaczona skrótem DLMR oznaczającym Dane Ludnościowe Minus Rejestrowe⁴.

Wszystkie powyżej opisane zmienne zastępcze wykorzystywane są w rozpatrywanych przekrojach i układach tabel. Na ich podstawie konstruowane są struktury asocjacji w tabelach zastępczych, które z kolei są wykorzystywane w procesie estymacji w ramach rozważanych modeli GSPREE.

Zmienne z badania reprezentacyjnego

Biorąc pod uwagę jedynie warianty estymatora GLSM zakładające, że parametr $\beta \neq 1$ w procesie estymacji wykorzystywane były szacunki kalibracyjne z badania BAEL dla 2017 roku⁵ w rozważanych w badaniu empirycznym układach i przekrojach. Uzyskane z tej części badania oszacowania kalibracyjne liczby osób pracujących, bezrobotnych oraz biernych zawodowo wykorzystywane są w procesie estymacji do korygowania interakcji występujących w tabelach zastępczych⁶.

4.2.2. Struktura alokacji

Wartości brzegowe tabel kontyngencji we wszystkich rozważanych estymatorach typu SPREE i jego wariantach pochodziły z badania BAEL (dane średnioroczne dla 2017 roku). Wartości brzegowe dla wierszy są tożsame ze stanem i strukturą ludności w gospodarstwach domowych wyznaczoną dla tego badania⁷. Do tych struktur kalibrowane są wagi w badaniu BAEL. Wyjątek stanowią tutaj dwa podregiony (Gorzowski oraz Zielonogórski), które nie są uwzględniane w procesie kalibracji wag. Wynika to z tego, że korekta wag dokonywana jest dla struktury ludności na poziomie województw w podziale na płeć i grupy wieku, bez uwzględnienia podziału na podregiony. Natomiast wartości brzegowe kolumn oraz – w przypadku podejścia trzykrokowego – wartości brzegowe dla poszczegół-

⁴Porównanie liczby osób biernych zawodowo w podziale na płeć na poziomie województw w 2017 roku pomiędzy DLMR a BAEL znajduje się w aneksie na rysunku A.1.

⁵Wybrane wyniki prac wykorzystujące w pewnym elemencie to źródło danych były prezentowane przez autora niniejszej rozprawy między innymi na konferencji pt. „230 Lat Statystyki Publicznej w Polsce” <https://konferencja230.stat.gov.pl>.

⁶Więcej informacji na temat metodologii znajduje się w rozdziale 3.

⁷Więcej informacji na temat procesu kalibracji znajduje się w publikacjach (GUS, 2018a,d).

nych wariantów zmiennej grupującej (wyznaczonych przez płeć i/lub grupy wieku) były sumą wag w odpowiednich przekrojach na podstawie badania BAEL.

4.3. Zakres szacowanych kategorii i podkategorii zmiennych

Tabela 4.1 zawiera wykaz analizowanych kategorii zmiennych oraz symboli, które są stosowane w dalszej części opisu i wizualizacji wyników.

Tabela 4.1: Wykaz zmiennych i ich kategorii

Zmienna	Symbol zmiennej	Kategoria	Symbol kategorii
Jednostka podziału terytorialnego	JT	Dolnośląskie	02
		Kujawsko-Pomorskie	04
		Lubelskie	06
		Podr. Gorzowski	08G
		Podr. Zielonogórski	08Z
		Łódzkie	10
		Małopolskie	12
		Mazowieckie	14
		Opolskie	16
		Podkarpackie	18
		Podlaskie	20
		Pomorskie	22
		Śląskie	24
		Świętokrzyskie	26
Warmińsko-Mazurskie	28		
Wielkopolskie	30		
Zachodniopomorskie	32		
Dodatkowy podział jednostek terytorialnych	DP	Miejskie obszary funkcjonalne ośrodków wojewódzkich	MOF OW
		Pozostały obszar jednostki terytorialnej	POJT
Status na rynku pracy	ST	Pracujący	PRAC
		Bezrobotni	BEZR
		Bierni zawodowo	BIER
Płeć	PL	Mężczyźni	M
		Kobiety	K
Grupy wieku	GW	15–29 lat	15–29
		30–39 lat	30–39
		40–49 lat	40–49
		50 i więcej	50+

Źródło: Opracowanie własne.

Z kolei w tabeli 4.2 zamieszczono wykaz analizowanych podkategorii zmiennych oraz symboli, które są stosowane w dalszej części opisu i wizualizacji wyników. Tabela 4.2 zawiera

szczegółowy wykaz podkategorii zmiennych odnoszących się do poszczególnych MOF OW zbliżonych do nazewnictwa przyjętego w oficjalnych dokumentach strategicznych inicjujących ustanowienie poszczególnych obszarów ZIT. Przyjęte w pracy doktorskiej symbole podkategorii zmiennych w zamyśle mają ułatwić ich umiejscowienie na mapie Polski. Należy dodać, że dla poszczególnych POJT nie ustalono symboli podkategorii, gdyż w rozprawie nie są one analizowane bez odniesień do powiązanych z nimi MOF OW.

Tabela 4.2: Podkategorie zmiennych

Symbol kategorii	Podkategoria	Symbol podkategorii
MOF OW	MOF Subregionu Centralnego	Katowic
	Warszawski OF	Warszawy
	MOF Gdańsk–Gdynia–Sopot	Trójmiasta
	Łódzki OM	Łodzi
	Krakowski OF	Krakowa
	Wrocławski OF	Wrocławia
	MOF Poznań	Poznań
	Bydgosko-Toruński OF	Bdg.-Torunia
	Szczeciński OM	Szczecina
	Lubelski OF	Lublina
	Białostocki OF	Białegostoku
	Rzeszowski OF	Rzeszowa
	Kielecki OF	Kielc
	MOF Aglomeracji Opolskiej	Opola
	MOF Olsztyna,	Olsztyna
MOF Zielonej Góry	Zielonej Góry	
MOF Gorzowa Wielkopolskiego	Gorzowa Wlkp.	
POJT	Pozostała część województwa lub podregionu po wyłączeniu obszaru danego MOF OW	Brak

Źródło: Opracowanie własne.

Zachowując kompromis pomiędzy zgłaszanymi potrzebami użytkowników danych oraz biorąc pod uwagę specyfikę estymatorów typu SPREE zdecydowano się na dostarczenie zakładanej informacji statystycznej dla 17 miejskich obszarów funkcjonalnych ośrodków wojewódzkich (MOF OW)⁸ wyznaczonych do realizacji instrumentu ZIT w ramach perspektywy finansowej Unii Europejskiej na lata 2014–2020. Dodatkowo ze względu na przyjęte i opisane w pracy konstrukcje tabel (podrozdział 4.5) w ramach wszystkich rozpa-

⁸Przyjęto w części empirycznej pracy, iż oprócz pojęcia **miejskie obszary funkcjonalne ośrodków wojewódzkich (MOF OW)** zamiennie stosowane będą określenia: obszary funkcjonalne miast wojewódzkich, miejskie obszary funkcjonalne.

trywanych przekrojów i układów oszacowania uzyskano dla pozostałych części jednostek terytorialnych (POJT) okalających poszczególne MOF OW⁹.

4.4. Rozważane estymatory i ocena ich jakości

Tabela 4.3 zawiera wykaz wszystkich estymatorów zastosowanych w niniejszej rozprawie doktorskiej, która bazuje na przeglądzie estymatorów typu SPREE szczegółowo opisanym w rozdziale trzecim¹⁰.

Tabela 4.3: Opis zaimplementowanych wariantów estymatorów

Lp.	Opis estymatora	Przyjęty skrót
1	Estymator kalibracyjny	<i>CAL</i>
2	GLSM, gdzie $\beta = 1$ odpowiada klasycznemu SPREE dla dwuwymiarowej tabeli kontyngencji - zmienna warstwująca ujęta w układzie kolumn	<i>SPREE_{2,W}</i>
3	GLSM dla dwuwymiarowej tabeli kontyngencji - zmienna warstwująca ujęta w układzie kolumn	<i>GLSM_{2,W}</i>
4	GLSMM dla dwuwymiarowej tabeli kontyngencji - zmienna warstwująca ujęta w układzie kolumn	<i>GLSMM_{2,W}</i>
5	GLSM, gdzie $\beta = 1$ odpowiada klasycznemu SPREE - zmienna warstwująca ujęta w układzie wierszy	<i>SPREE_{2,L}</i>
6	GLSM dla dwuwymiarowej tabeli kontyngencji - zmienna warstwująca ujęta w układzie wierszy	<i>GLSM_{2,L}</i>
7	GLSMM dla dwuwymiarowej tabeli kontyngencji - zmienna warstwująca ujęta w układzie wierszy	<i>GLSMM_{2,L}</i>
8	GLSM, gdzie $\beta = 1$ odpowiada klasycznemu SPREE dla trójwymiarowej tabeli kontyngencji - zmienna warstwująca ujęta w układzie wierszy	<i>SPREE_{3,L}</i>
9	GLSM, gdzie parametr β jest zależny od warstwy w trójwymiarowej tabeli kontyngencji - zmienna warstwująca ujęta w układzie wierszy	<i>GLSM_{3S,L}</i>
10	GLSM, gdzie parametr β jest taki sam dla wszystkich warstw w trójwymiarowej tabeli kontyngencji - zmienna warstwująca ujęta w układzie wierszy	<i>GLSM_{3P,L}</i>
11	GLSMM, gdzie parametr β jest zależny od warstwy w trójwymiarowej tabeli kontyngencji - zmienna warstwująca ujęta w układzie wierszy	<i>GLSMM_{3S,L}</i>
12	GLSMM, gdzie parametr β jest taki sam dla wszystkich warstw w trójwymiarowej tabeli kontyngencji - zmienna warstwująca ujęta w układzie wierszy	<i>GLSMM_{3P,L}</i>

Źródło: Opracowanie własne.

⁹Poza dwoma wyjątkami (podregion gorzowski, podregion zielonogórski), były to pozostałe części województw.

¹⁰Przyjęte skróty do poszczególnych typów estymatorów będą stosowane podczas prezentowania uzyskanych wyników przedstawianych na rysunkach oraz tabelach w dalszej części pracy.

Zgodnie ze schematem postępowania (por. rysunek 4.1) w pierwszej kolejności przeprowadzono estymację kalibracyjną stanowiącą punkt odniesienia dla oceny wyników szacunków uzyskanych na podstawie wszystkich rozpatrywanych wariantów estymatora typu SPREE. Jest on wymieniony na pierwszej pozycji (Lp. 1) w tabeli 4.3 i oznaczony skrótem *CAL*. W schemacie procedury badawczej przyjęto założenie, że (por. rysunek 4.1), w przypadku otrzymania zadowalających wyników estymacji z wykorzystaniem jednego z rozpatrywanych estymatorów, biorąc pod uwagę jego efektywność, proces estymacji można by uznać za zakończony. Mając jednak świadomość, że przyjęte przekroje tabel zawierają domeny nieplanowane na etapie projektowania badania, należałoby wówczas dodatkowo dokładnie przeanalizować rozkład próby w danej domenie i sprawdzić, czy jest ona w nim reprezentowana przez wylosowane jednostki. Estymator kalibracyjny jest również używany jako jeden z komponentów podczas procesu estymacji z wykorzystaniem wszystkich wariantów estymatora GLSM oraz GLSMM, za wyjątkiem GLSM, gdzie $\beta = 1$. Następnie przeprowadzano proces estymacji pośredniej z wykorzystaniem wymienionych w tabeli 4.3 estymatorów, w zależności od rozpatrywanego przekroju i układu tabeli. Rozpoczynając od estymatora oznaczanego w badaniu empirycznym jako SPREE (szczególny przypadek estymatora GLSM z $\beta = 1$) poprzez GLSM oraz GLSMM dla tabel dwuwymiarowych, a następnie trójwymiarowych.

Każdy z tych wariantów mógł być dodatkowo rozpatrywany w ujęciu szerokim¹¹ (oznaczone subskrypcją *W*) – tzn. w kolumnach tabeli kontyngencji ujmowane są dodatkowe kategorie zmiennej warstwującej (płeć i/lub grupy wieku) – (Lp. 2, 3, 4) – oraz w ujęciu wąskim (oznaczone subskrypcją *L*) – tzn. dodatkowe kategorie zmiennej warstwującej (płeć i/lub grupy wieku) są ujmowane w wierszach tabeli kontyngencji – (Lp. 5, 6, 7). Następną grupę stanowiły estymatory dla trójwymiarowej tabeli kontyngencji (Lp. 8, 9, 10, 11) w ujęciu wąskim i dodatkowo w dwóch wariantach dla estymatorów GLSM oraz GLSMM w zależności od rodzaju przyjętego współczynnika β jednego dla całej tabeli (oznaczenie *P* w dolnym indeksie) lub zależnego od warstwy wyznaczonej poprzez przyjętą zmienną grupującą (oznaczenie *S* w dolnym indeksie).

W celu weryfikacji jakości powyższych wariantów estymatorów zdecydowano się na wyznaczenie współczynnika *CV*, określonego wzorem 3.3, zgodnie z analizą przeprowadzoną

¹¹Objaśnienia pojęć: ujęcie szerokie oraz wąskie znajduje się w Kursie „Pogromcy danych” http://pogromcydanych.icm.edu.pl/materials/1_przetwarzanie/24_tidyr.

w podrozdziale 3.1.2. Wariancję poszczególnych estymatorów, która jest składową wzoru na współczynnik CV , uzyskano poprzez zastosowanie metody bootstrap uwzględniającej schemat losowania próby w badaniu BAEL (*design based*).

Metoda bootstrap polega na wielokrotnym próbkowaniu z oryginalnej próby zgodnie ze schematem losowania stosowanym w badaniu reprezentacyjnym, z którego pochodzi oryginalna próba. W badaniu BAEL wykorzystuje się ją do wyznaczania wariancji estymatora kalibracyjnego. Procedura została szczegółowo opisana w publikacji GUS (2018a). Składa się ona z kilku etapów.

W pierwszym etapie odbywa się szacowanie wariancji, które jest realizowane na podstawie danych z prób kwartalnych. Są one losowane z uwzględnieniem podziału na warstwy na poziomie jednostek losowania pierwszego stopnia. Zaimplementowany wariant metody bootstrap realizowany jest oddzielnie w każdej warstwie, w celu otrzymania oszacowań odpowiedniego składnika wariancji wybranego estymatora. Następnie w każdej z warstw zawierających n_h wylosowanych jednostek pierwszego stopnia losowana jest podpróbka. Jest ona otrzymywana za pomocą losowania prostego ze zwracaniem o liczebności $n_h - 1$. W kolejnym kroku po otrzymaniu wylosowanych jednostek pierwszego stopnia do podpróbki bootstrapowej przenosi się wylosowane z nich do BAEL mieszkania łącznie z zamieszkującymi je respondentami. Taki sposób losowania powtarzany jest iteracyjnie B razy. Każdorazowo przeprowadzana jest modyfikacja wag, która obejmuje między innymi ich kalibrację. Poprzez takie działanie zapewniamy zgodność z danymi demograficznymi dla właściwego kwartału. Następnie dla każdej b -tej próby w oparciu o zmodyfikowane wagi wyznacza się ocenę estymatora kalibracyjnego interesującego nas parametru t_b^* . Natomiast w przypadku rozważanej w tej rozprawie estymacji pośredniej na tym etapie w oparciu o zmodyfikowane wagi, a także o dane z tabeli zastępczej wyznaczane są oceny estymatorów typu SPREE. Po wykonaniu B iteracji oszacowaniem wariancji estymatora badanego parametru \hat{t} w rozważanej warstwie, zgodnie z zasadą bootstrap jest wielkość:

$$\hat{V}(\hat{t}) = \frac{1}{B-1} \sum_{b=1}^B (\hat{t}_b^* - \hat{t})^2. \quad (4.1)$$

W celu otrzymania oszacowań wariancji dla rozważanej podpopulacji trzeba zsumować oszacowania wariancji otrzymane niezależnie we wszystkich jej warstwach. Przedstawioną powyżej procedurę daje się również zastosować z pewnymi modyfikacjami, do szacowania wariancji estymatorów pośrednich. Do wyznaczania wariancji dla danych publikacyjnych

w ramach badania BAEL przyjmuje się $B = 500$. Tę samą wartość przyjęto w niniejszej rozprawie dla wszystkich rozważanych wariantów estymatorów.

Po wyznaczeniu precyzji testowanych estymatorów oblicza się dla nich współczynniki CV , które należy oceniać według wspólnych kryteriów przyjętych dla wszystkich estymatorów. Ponieważ przyjęte w rozprawie przekroje, dla których dokonywany jest szacunek, nie odpowiadają przekrojom standardowo uwzględnianym w badaniu BAEL nie można ich oceniać na podstawie tych samych kryteriów, głównie ze względu na znacznie większą szczegółowość analizowanych szacunków. Przyjęto zatem zasady oceny zgodne z modelem obowiązującym w badaniu NSP 2011, opisanym w podrozdziale 2.5.1. Zakłada on, że $CV \leq 10\%$ uznaje się za akceptowalny poziom błędu szacunku. Z kolei szacunki, dla których $10\% < CV \leq 20\%$ powinny być traktowane z ostrożnością, jako mało precyzyjne. Przekroczenie progu $CV > 20\%$ traktowane jest jako wyraźny sygnał świadczący o braku wiarygodności uzyskanych oszacowań.

Należy mieć również świadomość tego, iż w badaniu BAEL występują liczne braki odpowiedzi (GUS, 2018d). Ponadto w niektórych przekrojach nieliczna jest reprezentacja osób bezrobotnych. Przeprowadzana jest również kalibracja wag do znanych struktur ludności. Wszystkie te elementy wpływają na obciążenie stosowanego estymatora kalibracyjnego. Pomiar obciążenia oraz jego analiza niestety nie jest publikowana w opracowaniach wyników badania. Dlatego w rozprawie także pominięto analizę obciążenia estymatorów pośrednich. Pomimo tych zastrzeżeń do estymatora kalibracyjnego, wydaje się, że można go potraktować jako punkt odniesienia do oszacowań pośrednich i sprawdzić, czy oszacowania są zbliżone czy też różnią się istotnie od siebie. Oceny takiej można dokonać na podstawie diagramów rozrzutu bądź też poprzez zmierzenie wielkości różnic wyrażonych stosunkiem jednych oszacowań względem drugich.

W pracy dodatkowym elementem oceny jakości estymatorów pośrednich, obok omawianej wcześniej miary precyzji CV , była dyskusja poświęcona wartościom współczynnika β i jego wpływu na oszacowania charakterystyk rynku pracy. W ocenie jakości uwzględniono również stopień zgodności uzyskanych oszacowań z danymi publikowanymi dla kategorii rynku pracy, których nie przewidziano w konstrukcji przyjętych tabel. Sprawdzone również, czy dany estymator zapewnia zgodność oszacowań w ramach zmiennych warstwujących, którymi były określone struktury ludności.

4.5. Rozważane układy tabel

W badaniu empirycznym weryfikowano możliwości przeprowadzenia estymacji kategorii rynku pracy w przekrojach tabel, które nie są obecnie dostarczane przez statystykę publiczną. Stanowią one odpowiedź na pojawiające się zapotrzebowanie zgłaszane ze strony użytkowników danych statystycznych. Kolejność konstruowania poszczególnych przekrojów tabel opiera się na schemacie przedstawionym na rysunku 4.1. Ich budowę rozpoczęto od przekroju ogółem bez podziału na zmienne warstwujące. W dalszej kolejności tworzono coraz bardziej szczegółowe tabele, aż do osiągnięcia docelowego podziału, oceniając na poszczególnych etapach jakość uzyskiwanych oszacowań.

Założony podział tabel – w zależności od szczegółowości dostarczanej informacji – obejmował cztery przekroje wyznaczone na podstawie skrzyżowania kategorii zmiennych opisanych w tabeli 4.1:

- 1) $JT \times DP \times ST$ – ogółem,
- 2) $JT \times DP \times ST \times PL$ – według płci,
- 3) $JT \times DP \times ST \times GW$ – według grup wieku,
- 4) $JT \times DP \times ST \times PL \times GW$ – według płci i grup wieku.

Układ tabeli przedstawiony na rysunku 4.2 był pierwszym, od którego rozpoczęto proces estymacji.

Została ona skonstruowana na najwyższym stopniu ogólności i zawiera informacje o liczbie osób pracujących, bezrobotnych oraz biernych zawodowo w podziale na miejskie obszary funkcjonalne oraz pozostały obszar jednostek terytorialnych. Wartości brzegowe wierszy są sumą wag ludności według definicji BAEL na poziomie poszczególnych jednostek podziału terytorialnego. Taki układ tabeli nie zapewnia jednak zgodności rozpatrywanych kategorii rynku pracy z danymi publikowanymi na poziomie poszczególnych województw. Środek tabeli stanowiący strukturę asocjacji (zaznaczony kolorem zielonym) zawiera 102 komórki, których wartości podlegają oszacowaniu. Źródłem zmiennych zastępczych w przypadku wszystkich rozpatrywanych estymatorów typu SPREE są informacje o osobach pracujących oraz bezrobotnych pochodzące ze ZROPiB oraz dane o osobach biernych zawodowo obliczone jako różnica pomiędzy danymi demograficznymi, a informacjami pochodzącymi

SPREE_{2,w} GLSM_{2,w} GLSMM_{2,w}

JPT	MOF OW			POJPT			OGÓŁEM
	PRAC	BEZR	BIER	PRAC	BEZR	BIER	
02	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
04	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
06	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
08G	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
08Z	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
10	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
12	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
14	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
16	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
18	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
20	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
22	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
24	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
26	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
28	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
30	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
32	ZROPiB	ZROPiB	DLMR	ZROPiB	ZROPiB	DLMR	BAEL
OGÓŁEM	BAEL	BAEL	BAEL	BAEL	BAEL	BAEL	BAEL

Rysunek 4.2: Przekroje tabel ogółem w podziale na kategorie rynku pracy

Objaśnienia: Wszystkie warianty estymatora GLSM oraz GLSMM wykorzystują dodatkowo oszacowania kalibracyjne do korekty interakcji występujących w tabeli zastępczej.

Estymator kalibracyjny był również używany jako punkt odniesienia do oszacowań pośrednich.

Źródło: Opracowanie własne.

ze źródła ZROPiB w odpowiednich przekrojach¹². Wartości brzegowe (zaznaczone kolorem niebieskim) pochodzą z kolei z BAEL-a i stanowią strukturę alokacji niniejszej tabeli kontyngencji. Dla wierszy są sumą skalibrowanych wag (za wyjątkiem województwa lubuskiego, które należało podzielić na dwa podregiony, aby uzyskać oszacowania dla MOF OW Zielonej Góry oraz Gorzowa Wielkopolskiego) na poziomie analizowanych jednostek terytorialnych. Sumy wag odpowiadają ogólnej liczbie ludności w danym przekroju zgodnie z definicją obowiązującą w BAEL-u.

W przypadku kolumn są to oszacowania na poziomie kraju dla poszczególnych kategorii rynku pracy w podziale na miejskie obszary funkcjonalne oraz pozostałą część jednostki terytorialnej. Wartości brzegowe kolumn są wartościami oszacowań dla poziomu całego kraju i pomimo nietypowego podziału (MOF OW, POJPT – który nie obowiązuje w badaniu BAEL) ze względu na dużą reprezentację w próbie na poziomie kraju można przyjąć, że są one wysoce wiarygodne. Tabela podlega procesowi estymacji w podejściu dwukrokowym dla trzech wariantów modeli: GLSM w którym $\beta = 1$, co sprawia, że jego działanie

¹²Więcej szczegółów o procesie wyliczania tego zasobu danych przedstawiono w punkcie 4.2.

jest tożsamy z klasycznym SPREE i tak też jest tutaj oznaczone; GLSM z szacowanym parametrem β , a także GLSMM w ujęciu wąskim, w którym żadne dodatkowe przekroje poza jednostkami podziału terytorialnego nie są wprowadzane na poziomie wierszy. Podczas procesu estymacji z wykorzystaniem wariantów estymatora GLSM z szacowanym parametrem β oraz GLSMM, do korekty interakcji zmiennych zastępczych w środku tabeli kontyngencji wykorzystywane są również oszacowane interakcje estymatora kalibracyjnego.

Kolejnym krokiem badawczym było wprowadzenie dodatkowego podziału tzn. uwzględnienie zmiennej „płeć”. Na rysunku 4.3 zamieszczone są trzy układy tabel, w ramach których szacowane były analizowane zmienne.

Rysunek 4.3: Przekroje tabel uwzględniające płeć w podziale na kategorie rynku pracy

a) $SPREE_{2,W}$ $GLSM_{2,W}$ $GLSMM_{2,W}$			b) $SPREE_{2,L}$ $GLSM_{2,L}$ $GLSMM_{2,L}$			c) $SPREE_{3L}$ $GLSM_{3P,L}$ $GLSM_{3S,L}$ $GLSMM_{3S,L}$ $GLSMM_{3P,L}$				
JT	DP	OGÓLEM	PL	JT	DP	OGÓLEM	PL	JT	DP	OGÓLEM
	PL				ST				ST	
	ST									
02	ZMIENNE ZASTĘPCZE	BAEL	M	02	ZMIENNE ZASTĘPCZE	BAEL	M	02	ZMIENNE ZASTĘPCZE	BAEL
04				04				04		
06				06				06		
...					
OGÓLEM	BAEL		OGÓLEM	BAEL		OGÓLEM	BAEL		BAEL	
			K	02			K	02		
				04				04		
				06				06		
					
			OGÓLEM	BAEL			OGÓLEM	BAEL		
							RAZEM	BAEL		
							K	02		
								04		
								06		
								...		
							RAZEM	BAEL		

Objaśnienia: Wszystkie warianty estymatora GLSM oraz GLSMM wykorzystują dodatkowo oszacowania kalibracyjne do korekty interakcji występujących w tabeli zastępczej.

Estymator kalibracyjny był również używany jako punkt odniesienia do oszacowań pośrednich.

Źródło: Opracowanie własne.

W ramach wszystkich powyższych układów tabel i wariantów estymatorów oszacowano wartości w 204 komórkach ($JT \times DP \times ST \times PL$ por. tabela 4.1) w środku każdej z tabel kontyngencji. Tabela a) na rysunku 4.3 uwzględnia dodatkowy podział ze względu na płeć w ramach kolumn i jest rozpatrywana dla trzech wariantów estymatorów SPREE, GLSM oraz GLSMM w podejściu dwukrokovym i w wąskim układzie tabel. Natomiast w Tabeli b) wprowadzono dodatkowy podział ze względu na płeć do wierszy tabeli, co zmienia ich układ na pionowy (ujęcie wąskie) z tym samym zestawem estymatorów co w Tabeli

a). Należy podkreślić, że ten układ pozwala na dopasowanie do wartości brzegowych w wierszach jednostek podziału terytorialnego dodatkowo w podziale na płeć. Natomiast w przypadku Tabeli a) występujący w kolumnach podział na płeć określa dopasowanie tej struktury do wartości z BAEL na poziomie całego kraju. Prowadzi to do niezgodności w wymiarze płci na poziomej JT, jeżeli przyjmiemy, że BAEL dostarcza wiarygodnych szacunków w tym przekroju.

Ostatni układ Tabel c) również jest tworzony na podstawie oszacowań trzech estymatorów SPREE, GLSM, GLSMM, z uwzględnieniem większej liczby ich wariantów. Dla tego układu tabel wszystkie estymatory rozpatrywane były w podejściu trzykrokovym, w ujęciu wąskim. Wówczas wartości ze środka tabeli kontyngencji w kolumnach nie są korygowane do ich wartości brzegowych na poziomie kraju, lecz ze względu na dodatkową zmienną określoną na poziomie wierszy, to znaczy według płci. W wyniku tego podejścia uzyskuje się dodatkowe agregaty w ramach tabeli kontyngencji, do której dopasowywane są wartości znajdujące się w ich środku. Dodatkowo, dla estymatora typu GLSM oraz GLSMM istnieje możliwość wyznaczenia jednego parametru β dla środka całej tabeli kontyngencji, bądź też osobno w ramach każdej zmiennej grupującej na poziomie wierszy, to znaczy według płci analizowanej grupy osób.

Na rysunku 4.4 zamieszczony jest trzeci zestaw tabel, wprowadzający dodatkowy podział na cztery grupy wieku bez uwzględnienia płci.

Rysunek 4.4: Przekroje tabel uwzględniające grupy wieku w podziale na kategorie rynku pracy

a) $SPREE_{2,W}$ $GLSM_{2,W}$
 $GLSMM_{2,W}$

JT	DP	OGÓŁEM
	GW	
	ST	
02	ZMIENNE ZASTĘPCZE	BAEL
04		
06		
...		
OGÓŁEM	BAEL	

b) $SPREE_{2,L}$ $GLSM_{2,L}$
 $GLSMM_{2,L}$

GW	JT	DP	OGÓŁEM
		ST	
		5-29	
04			
06			
...			
30-39	02	ZMIENNE ZASTĘPCZE	BAEL
	04		
	06		
	...		
OGÓŁEM		BAEL	

c) $SPREE_{3L}$ $GLSM_{3P,L}$ $GLSM_{3S,L}$
 $GLSMM_{3S,L}$ $GLSMM_{3P,L}$

GW	JT	DP	OGÓŁEM
		ST	
		15-29	
04			
06			
...			
RAZEM		BAEL	
30-39	02	ZMIENNE ZASTĘPCZE	BAEL
	04		
	06		
	...		
RAZEM		BAEL	
...	...	ZMIENNE ZASTĘPCZE	BAEL
....		BAEL	

Objaśnienia: Wszystkie warianty estymatora $GLSM$ oraz $GLSMM$ wykorzystują dodatkowo oszacowania kalibracyjne do korekty interakcji występujących w tabeli zastępczej.

Estymator kalibracyjny był również używany jako punkt odniesienia do oszacowań pośrednich.

Źródło: Opracowanie własne.

Takie podejście skutkuje tym, że wewnątrz tabeli jest bardziej złożone. Składa się ono z 408 komórek ($JT \times DP \times ST \times GW$ – por. tabela 4.1), dla których otrzymano oszacowania. Typy estymatorów w ujęciu dwukrokovym i trzykrokovym pozostają takie same, jak w przypadku wcześniej opisywanego podejścia uwzględniającego płeć.

Na rysunku 4.5 przedstawiono ostatni rozpatrywany układ tabel. Jest on kompilacją wcześniej rozpatrywanych, uwzględniających osobno płeć i grupy wieku. Jest to docelowy pożądaný poziom szczegółowości zmiennych i ich kategorii.

Rysunek 4.5: Przekroje tabel uwzględniające płeć i grupy wieku w podziale na kategorie rynku pracy

a) $SPREE_{2,W}$ $GLSM_{2,W}$
 $GLSMM_{2,W}$

JT	DP	OGÓŁEM
	PL	
	GW	
	ST	
02	ZMIENNE ZASTĘPCZE	BAEL
04		
06		
...		
OGÓŁEM	BAEL	

b) $SPREE_{2,L}$ $GLSM_{2,L}$
 $GLSMM_{2,L}$

PL	GW	JT	DP	OGÓŁEM
			ST	
M	15-29	02	ZMIENNE ZASTĘPCZE	BAEL
		04		
		06		
...				
M		
K	15-29	02		
		04		
		06		
K		
OGÓŁEM			BAEL	

c) $SPREE_{3,L}$ $GLSM_{3P,L}$ $GLSM_{3S,L}$
 $GLSMM_{3P,L}$ $GLSMM_{3S,L}$

PL	GW	JT	DP	OGÓŁEM
			ST	
M	15-29	02	ZMIENNE ZASTĘPCZE	BAEL
		04		
		06		
		...		
RAZEM			BAEL	
M	ZMIENNE ZASTĘPCZE	BAEL
		...		
		...		
...			BAEL	
K	15-29	02	ZMIENNE ZASTĘPCZE	BAEL
		04		
		06		
		...		
RAZEM			BAEL	
K	ZMIENNE ZASTĘPCZE	BAEL
		...		
		...		
...			BAEL	

Ob-

jaśnienia: Wszystkie warianty estymatora GLSM oraz GLSMM wykorzystują dodatkowo oszacowania kalibracyjne do korekty interakcji występujących w tabeli zastępczej.

Estymator kalibracyjny był również używany jako punkt odniesienia do oszacowań pośrednich.

Źródło: Opracowanie własne.

Wprowadzenie na raz dwóch dodatkowych zmiennych grupujących spowodowało kolejne zwiększenie złożoności tabel, w wyniku czego estymacji dokonano dla liczebności w 816 komórkach ($JT \times DP \times ST \times PL \times GW$ – por. tabela 4.1) wewnątrz każdej z rozpatrywanych tabel kontyngencji.

4.6. Analiza statystyczna uzyskanych szacunków dla kategorii rynku pracy

Przedstawione na rysunku 4.1 etapy estymacji i ocenę ich jakości rozpoczęto od estymatora kalibracyjnego dla przekroju tabeli ogółem. Ocena jakości bazująca na współczynniku CV wykazała, iż tylko jedno oszacowanie wartości globalnych wewnątrz tabeli kontyngencji charakteryzowało się nieakceptowalnym błędem szacunku według przyjętych w pracy progów (por. rysunek A.2 w aneksie). Dlatego też w kolejnym kroku przeprowadzono

estymację dla wszystkich wariantów estymatora typu SPREE, dla których uzyskano akceptowalne błędy szacunku (por. rysunek A.2). Uzyskane w procesie estymacji tabele wynikowe, pozwoliły jednocześnie pozytywnie zweryfikować pierwszą hipotezę badawczą, że estymacja typu SPREE umożliwia oszacowanie podstawowych charakterystyk rynku pracy na poziomie agregacji przestrzennej, tj. miejskich obszarów funkcjonalnych ośrodków wojewódzkich, dla których obecnie nie ma pokrycia informacyjnego.

Następnie proces estymacji prowadzono etapami, zgodnie z przyjętym schematem postępowania, przechodząc kolejno poprzez przekroje pośrednie (płeć – por. rysunek A.4, grupy wieku – por. rysunek A.6), aż do przekroju docelowego, tj. uwzględniającego jednocześnie podział na płeć i ustalone grupy wieku.

Na każdym etapie uzyskane oceny estymatora kalibracyjnego były nieakceptowalne. Stałoby one jednak punkt odniesienia dla zastosowanych metod pośrednich. Dodatkowo szacunki kalibracyjne w przypadku wariantów estymatorów GLSM oraz GLSMM, gdzie szacowany jest parametr $\beta \neq 1$ były wykorzystywane w ramach konstrukcji tych estymatorów. Rozpatrywane warianty estymatorów SPREE oraz GLSMM charakteryzowały się akceptowalnym błędem szacunku i dawały znaczną poprawę precyzji oszacowań.

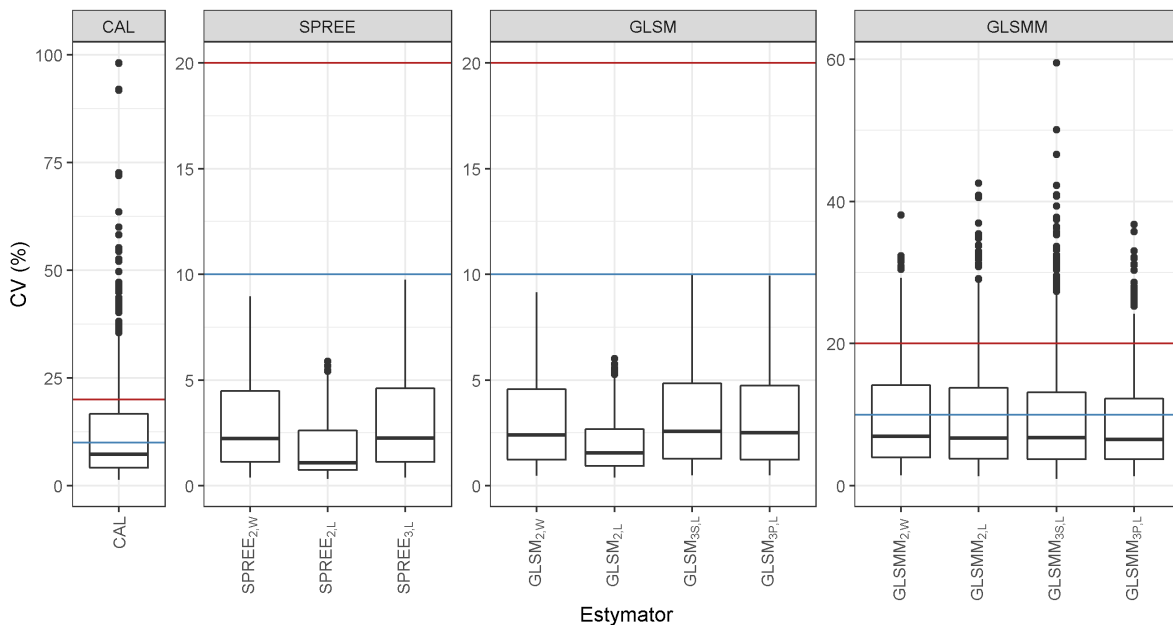
Ponieważ wartości bezwzględne uzyskane dla poszczególnych kategorii rynku pracy (liczba pracujących, bezrobotnych oraz biernych zawodowo w odpowiednich przekrojach i układach tabel) są elementem pośrednim w procesie pozyskania wartości uwzględnionych w badaniu wskaźników, pominięto ich szczegółową analizę. Uwagę skupiono na ocenie ich jakości, która determinuje ostateczny wybór estymatora służącego w dalszej części rozprawy do wyznaczenia wytypowanych wskaźników. Uzyskane wyniki branych pod uwagę ocen jakości oszacowań dla przekroju ogółem, w podziale na płeć, w podziale na grupy wieku znajdują się w aneksie w tabelach A.2–A.4 oraz na rysunkach A.2 – A.7. Przedstawiono tam również tabele zawierające oszacowania parametru β dla rozważanych wariantów estymatora typu SPREE w ramach poszczególnych przekrojów.

W niniejszym rozdziale przedstawiono wyniki stanowiące elementy oceny jakości otrzymanych szacunków dla przyjętego w pracy przekroju docelowego dotyczącego rynku pracy tj. w podziale na płeć i grupy wieku.

4.6.1. Analiza wartości współczynnika CV

Weryfikację jakości rozpoczęto od oceny precyzji oszacowań uzyskanych dla rozważanych wariantów estymatora typu SPREE (SPREE, GLSM i GLSMM) oraz estymatora kalibracyjnego (CAL).

Każdy wykres pudełkowy na rysunku 4.6 powstał na podstawie wszystkich uzyskanych ocen precyzji dla oszacowań liczebności w 802 komórkach wewnątrz poszczególnych układów tabel kontyngencji (por. 4.5).



Rysunek 4.6: Rozkład wartości współczynnika precyzji dla kategorii rynku pracy w układzie tabel uwzględniających płeć i grupy wieku

Uwaga: Przyjęte skale do oceny współczynnika CV są różne dla poszczególnych grup wariantów estymatorów.

Objaśnienia: Czerwoną linią zaznaczono wartość CV na poziomie 20%. Z kolei poziom $CV = 10\%$ wyznacza linia niebieska.

Źródło: Opracowanie własne.

Analizując rysunek 4.6 można zauważyć, że nieakceptowalny poziom precyzji ocen estymatora kalibracyjnego występuje dla prawie jednej czwartej szacowanych liczebności w komórkach wewnątrz przyjętej tabeli kontyngencji. Dla prawie połowy estymowanych liczebności w komórkach tabeli kontyngencji współczynnik precyzji przekracza 10%, a w skrajnym przypadku zbliża się do 100%. Z kolei w przypadku wariantów estymatorów SPREE i GLSM, uzyskane oceny precyzji wskazują na znaczną poprawę w stosunku do estymatora

kalibracyjnego, gdyż żadna z wartości CV nie przekracza ustalonego progu 10%, a wartości CV poniżej 5% uzyskano w przynajmniej 75% przypadków. Dla tych dwóch grup estymatorów najlepsze własności pod względem uzyskanych ocen precyzji wykazują estymatory $SPREE_{2,L}$ oraz $GLSM_{2,L}$, dla których wartość CV jedynie w kilku przypadkach nieznacznie przekracza 5%. Wszystkie warianty estymatora GLSMM posiadają wartości odstające przekraczające próg 20%, a w przypadku estymatorów $GLSMM_{2,L}$ oraz $GLSMM_{3S,L}$ występują obserwacje, dla których wartość CV przekracza 40%, dochodząc do prawie 60% w przypadku $GLSMM_{3S,L}$.

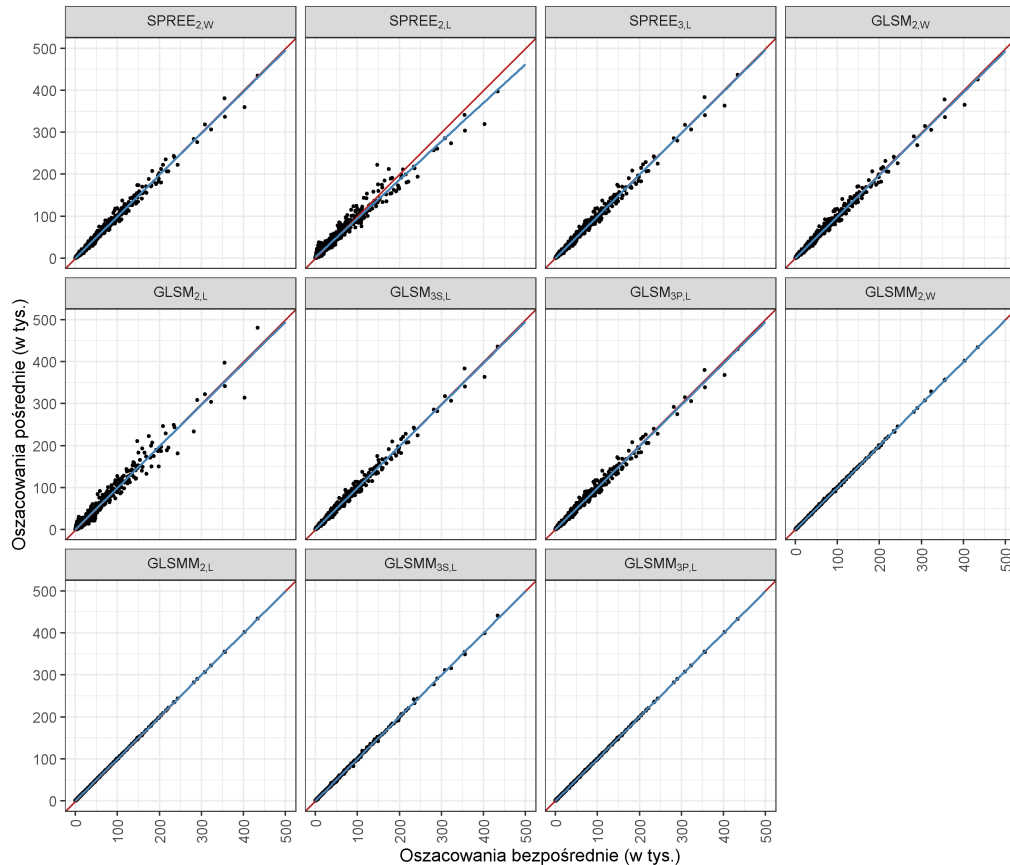
Przeprowadzone prace w zakresie oceny jakości precyzji szacunków jednoznacznie potwierdzają drugą hipotezę badawczą, że estymacja typu SPREE cechuje się lepszą precyzją szacunków charakterystyk rynku pracy w porównaniu z estymacją kalibracyjną stosowaną w badaniu BAEL.

4.6.2. Porównanie ocen estymatora kalibracyjnego z ocenami estymatorów pośrednich

Na rysunku 4.7 zamieszczono wykresy, które umożliwiają porównanie ocen estymatora kalibracyjnego z ocenami rozważanych estymatorów pośrednich. Z porównań wyłączone oceny estymatora kalibracyjnego, dla których współczynnik CV był powyżej 20%. Za przyjęciem takiego założenia przemawiała konieczność wykluczenia z porównań obserwacji, które posiadają nieakceptowalny poziom błędu.

Analizując poszczególne wykresy można zauważyć, że największe różnice w poziomie ocen estymatorów w porównaniu z ocenami estymatora kalibracyjnego widoczne są dla estymatora $SPREE_{2,L}$ oraz $GLSM_{2,L}$, chociaż w tym drugim przypadku są one już nieco mniejsze. Należy podkreślić, że dotyczy to estymatorów, które charakteryzowały się najwyższą precyzją spośród analizowanych estymatorów pośrednich zamieszczonych na rysunku 4.6. Pozostałe warianty estymatorów SPREE oraz GLSM charakteryzują się mniejszymi różnicami uzyskanych wartości oszacowań w porównaniu do estymatora kalibracyjnego. Z kolei wszystkie warianty estymatora GLSMM wykazują się wysoką zgodnością z oszacowaniami kalibracyjnymi. Podkreślić jednak należy, że jest to grupa estymatorów, którą cechowały się zdecydowanie wyższymi błędami szacunku wyrażonymi wartościami współczynnika

CV , w porównaniu do pozostałych rozważanych estymatorów pośrednich przedstawione na rysunku 4.6.



Rysunek 4.7: Porównanie szacunków pośrednich i kalibracyjnych w układzie tabel uwzględniających płeć i grupy wieku – $CV \leq 20\%$

Objaśnienia: Punkty układające się wzdłuż prostej oznaczonej kolorem czerwonym o równaniu $y = x$ świadczą o zgodności oszacowań wariantu estymatora pośredniego z kalibracyjnym. Natomiast niebieska linia jest prostą regresji dopasowaną do punktów na korelacyjnym diagramie rozrzutu.

Źródło: Opracowanie własne.

4.6.3. Analiza współczynników β

Kolejnym etapem weryfikacji wyników jest analiza parametrów β uzyskanych dla poszczególnych estymatorów.

Jednym z elementów, który należy wziąć przy ocenie pod uwagę jest poziom korekty struktury asocjacji tabeli zastępczej na podstawie danych z badania BAEL. Jest on wyrażony poprzez wartość parametru β , określanego również jako współczynnik proporcjonalności. Tabela 4.4 nie zawiera wariantów klasycznego estymatora SPREE, gdyż jak wynika z ich konstrukcji $\beta = 1$ nie jest modelowany z wykorzystaniem danych z BAEL.

Tabela 4.4: Oszacowania parametrów β rozpatrywanych wariantów estymatorów GLSM oraz GLSMM dla przekroju płci i grup wieku

Estymator	Warstwa	Parametr β
$GLSM_{2,W}$	–	0,948
$GLSM_{2,L}$	–	1,318
$GLSM_{3P,L}$	–	0,956
$GLSMM_{2,W}$	–	0,877
$GLSMM_{2,L}$	–	0,990
$GLSMM_{3P,L}$	–	0,915
$GLSM_{3S,L}$	Kobiety 15–29	0,856
	Kobiety 30–39	0,954
	Kobiety 40–49	0,940
	Kobiety 50+	0,998
	Mężczyźni 15–29	0,829
	Mężczyźni 30–39	0,991
	Mężczyźni 40–49	0,891
	Mężczyźni 50+	1,017
$GLSMM_{3S,L}$	Kobiety 15–29	0,615
	Kobiety 30–39	0,883
	Kobiety 40–49	0,927
	Kobiety 50+	0,957
	Mężczyźni 15–29	0,525
	Mężczyźni 30–39	0,929
	Mężczyźni 40–49	0,828
	Mężczyźni 50+	0,943

Źródło: Opracowanie własne.

Natomiast dla pozostałych wariantów estymatorów GLSM oraz GLSMM parametr ten jest estymowany, przy czym dla tabel dwuwymiarowych oraz trójwymiarowych w wariancie P otrzymujemy jedną wartość parametru dla całej tabeli. Natomiast w przypadku wariantu S jest ona zależna od warstwy wyznaczonej przez zmienną warstwującą. W przypadku niniejszego badania jest to płeć w podziale na grupy wieku. Wartość współczynnika proporcjonalności poniżej jedności powoduje osłabienie interakcji występujących w tabeli zastępczej, na podstawie informacji z BAEL-a. Wartość współczynnika powyżej jedności wskazuje na wzmocnienie interakcji w tabeli zastępczej na bazie informacji z badania BAEL. Analizując tabelę 4.4 można zauważyć wprowadzenie istotnych korekt poprzez parametr β w przypadku estymatora GLSMM szczególnie w warstwach wyznaczonych

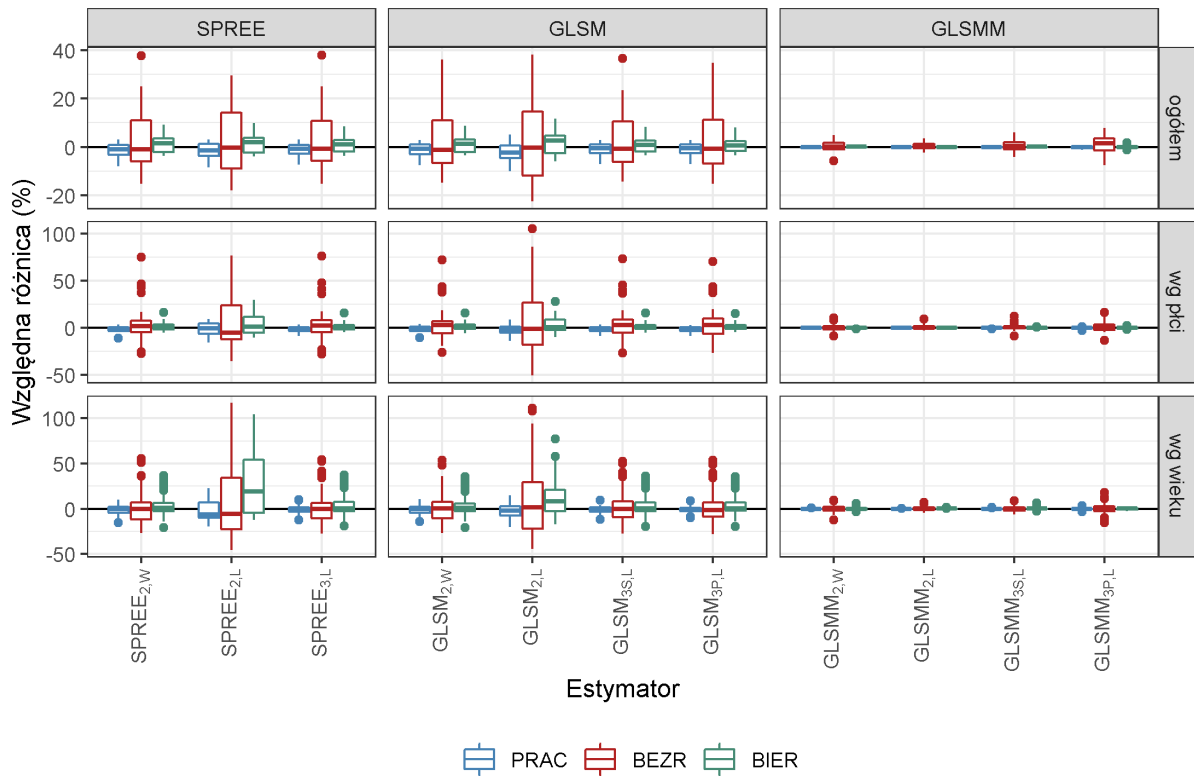
przez przekroje mężczyźni w wieku 15–29 lat oraz kobiety w wieku 15–29. Należy zauważyć, że grupa estymatorów GLSMM ze względu na swoją konstrukcję upodabnia się do estymatora kalibracyjnego. Im większa wariancja estymatora kalibracyjnego, tym większa wariancja składnika losowego stanowiącego komponent estymatorów GLSMM. Ma to często miejsce w przypadku przekrojów, które są słabo reprezentowane w próbie (co za tym idzie szacunki kalibracyjne nie są wiarygodne). Wydaje się jednak, że w takich przypadkach nie jest to pożądana własność estymatora, gdyż wymuszona na ich podstawie tak znaczna korekta interakcji tabeli zastępczej jest w takich przypadkach niepożądana.

4.6.4. Weryfikacja różnic pomiędzy oszacowaniami pośrednimi a publikowanymi informacjami

Kolejnym elementem, który wzięto pod uwagę weryfikując jakość uzyskiwanych oszacowań dla poszczególnych kategorii rynku pracy są różnice pomiędzy publikowanymi informacjami na poziomie województw, a otrzymanymi w wyniku zastosowania wariantów estymatorów typu SPREE. Chodzi tu przede wszystkim o takie zmienne jak liczba osób pracujących, bezrobotnych oraz biernych zawodowo – ogółem, w podziale na płeć i osobno grupy wieku. Porównań można dokonać tylko w takich przekrojach, gdyż oszacowania na niższym poziomie agregacji tj. dla województw w podziale na płeć i grupy wieku jednocześnie nie są publikowane.

Należy dodać, że w tabelach przedstawionych na rysunku 4.5 zsumowane do poziomu województw oszacowania dla poszczególnych kategorii rynku pracy z wykorzystaniem estymatorów typu SPREE, nie są zgodne z opublikowanymi wartościami dla tych kategorii otrzymanymi przy użyciu estymatora kalibracyjnego. Różnica ta wynika z tego, że wartości wewnątrz tabel kontyngencji były korygowane do wartości brzegowych w wierszach, a nie do wartości poszczególnych kategorii rynku pracy uzyskanych z wykorzystaniem estymatora kalibracyjnego. Analizując rysunek 4.8 zauważyć można dużą zgodność pomiędzy ocenami otrzymanymi dla wszystkich wariantów estymatora typu GLSMM a ocenami estymatora kalibracyjnego w przypadku wszystkich kategorii rynku pracy ogółem oraz w podziale na płeć i grupy wieku. Zdecydowanie większe różnice zauważalne są natomiast dla pozostałych typów estymatora SPREE czyli podejścia klasycznego oraz GLSM. Szczególnie jest to widoczne w odniesieniu do liczby osób bezrobotnych. Należy zauważyć,

że różnice te zwiększają się wraz ze wzrostem poziomu szczegółowości. Największe różnice uwzględniające wartości odstające występują w przypadku estymatora $SPREE_{2,L}$ oraz $GLSM_{2,L}$, gdzie oszacowania pośrednie liczby osób bezrobotnych są w niektórych przypadkach ponad 100% wyższe niż dla estymatora kalibracyjnego.



Rysunek 4.8: Względne różnice pomiędzy ocenami estymatora kalibracyjnego, a poszczególnymi wariantami estymatorów typu SPREE w przyjętych przekrojach

Źródło: Opracowanie własne.

Warto jednak dodać, że liczba osób bezrobotnych ogółem oraz w rozpatrywanych przekrojach jest zdecydowanie rzadziej reprezentowana w próbie niż liczba osób pracujących i biernych zawodowo ze względu na bardzo niski poziom tego zjawiska w 2017 roku. Oszacowania pozostałych kategorii rynku pracy dla grup estymatorów SPREE oraz GLSM w porównaniu do estymatora kalibracyjnego w ramach rozpatrywanych przekrojów charakteryzują się zdecydowanie mniejszym zróżnicowaniem i są zbliżone do siebie.

Na koniec warto również zaznaczyć, że szacunki otrzymane dla wszystkich wariantów estymatora typu SPREE dla tabel trójwymiarowych były zgodne z przyjmowaną na poziomie wierszy zmienną warstwową. Zapewniało to spójność ze strukturą ludności w podziale na płeć i grupy wieku. Jest to kolejna własność badanych estymatorów, która przemawia

za ich użyciem w badaniu empirycznym i pomaga w ostatecznym wyborze najlepszego wariantu estymatora.

4.7. Podsumowanie

W ramach zaprezentowanej w niniejszym rozdziale procedury badawczej przedstawionej na rysunku 4.1 szczegółowo opisano podjęte w ramach niniejszej dysertacji rozważania teoretyczne oraz działania empiryczne. Potwierdzono pierwszą hipotezę badawczą, iż estymacja typu SPREE pozwala na oszacowanie podstawowych charakterystyk rynku pracy na poziomie agregacji przestrzennej, tj. miejskich obszarów funkcjonalnych ośrodków wojewódzkich, dla których obecnie nie ma pokrycia informacyjnego. W szerokim zakresie przeprowadzono analizę jakości uzyskanych oszacowań charakterystyk rynku pracy dla przyjętego w analizie przekroju w podziale na płeć i grupy wieku. Wyniki przeprowadzonego badania pozwoliły pozytywnie zweryfikować postawioną na wstępie pracy drugą hipotezę zakładającą, że rozpatrywane warianty estymatora typu SPREE charakteryzują się lepszą precyzją szacunków w porównaniu do estymatora kalibracyjnego. Biorąc pod uwagę takie wyznaczniki jak: wartości współczynników precyzji dla wariantów estymatorów pośrednich, wyniki porównań ocen badanych estymatorów pośrednich i estymatora kalibracyjnego oraz wartości współczynników β , a także wyniki porównań wartości dla poszczególnych kategorii rynku pracy z wartościami publikowanymi, podjęto decyzję o wyborze wariantu estymatora typu SPREE postaci $GLSM_{3P,L}$. Zostanie on wykorzystany do wyznaczania wskaźników rynku pracy: współczynnika aktywności ekonomicznej, wskaźnika zatrudnienia oraz stopy bezrobocia, co zostanie zaprezentowane w piątym rozdziale pracy. Istotną własnością tego estymatora jest to, że zapewnia on zgodność ze strukturą ludności przyjmowaną w badaniu BAEL, która jednocześnie odpowiada rozpatrywanej zmiennej warstwującej (płeć w podziale na 4 grupy wieku).

Zgodnie z konstrukcją estymatora $GLSM_{3P,L}$, ostateczny proces estymacji wykorzystujący technikę IPF działa osobno dla każdej warstwy, natomiast współczynnik proporcjonalności β jest wyznaczany jeden dla wszystkich warstw. Takie podejście jest zbieżne z ideą „pożyczania mocy” w estymacji pośredniej. Ponadto ten wariant estymatora jest bardziej odporny na oceny estymatora kalibracyjnego wartości globalnej w przekrojach, które są nielicznie reprezentowane w próbie. Biorąc pod uwagę wszystkie wyżej wskazane prze-

słanki merytoryczno-metodyczne dokonano ostatecznego wyboru estymatora $GLSM_{3P,L}$ jako tego, który charakteryzował się najlepszymi własnościami. Wybór narzędzia estymacji umożliwił przejście do kolejnego etapu prac jakim było uzyskanie oszacowań dla wskazanych w celu badania wskaźników, ich statystycznej weryfikacji oraz oceny sytuacji na rynku pracy na podstawie otrzymanych wyników. Ten etap badania zostanie przedstawiony w kolejnym rozdziale.

Rozdział 5

Wskaźniki rynku pracy dla obszarów funkcjonalnych miast wojewódzkich

Niniejszy rozdział składa się z dwóch części. W pierwszej dokonano analizy precyzji szacunków wskaźników rynku pracy uzyskanych na poziomie miejskich obszarów funkcjonalnych, w przekroju płci oraz z uwzględnieniem czterech grup wieku. Oceny precyzji dokonano w odniesieniu do szacunków estymatora kalibracyjnego, który stanowił tzw. estymator referencyjny. Ocenie poddano również różnice pomiędzy otrzymanymi oszacowaniami wskaźników, a oszacowaniami kalibracyjnymi, stanowiącymi ich komponenty. Wyniki przedstawione w pierwszej części, stanowią uzupełnienie analizy prowadzonej w rozdziale czwartym¹ poświęconej weryfikacji hipotezy badawczej zakładającej, że estymacja typu SPREE cechuje się lepszą precyzją szacunków charakterystyk rynku pracy w porównaniu z estymacją kalibracyjną stosowaną w badaniu BAEL.

W części drugiej odwołano się do ostatniego etapu procedury badawczej zaprezentowanej w rozdziale czwartym, jakim jest ocena sytuacji na rynku pracy w zdefiniowanych przekrojach (por. rys. 4.1). Celem tej części była przede wszystkim próba weryfikacji dwóch

¹W rozdziale czwartym przedstawiono porównanie precyzji oszacowań w odniesieniu do wartości absolutnych, w przeciwieństwie do rozdziału piątego, w którym analiza precyzji oszacowań odnosi się do wskaźników rynku pracy.

hipotez badawczych zakładających, że sytuacja na rynku pracy w miejskich obszarach funkcjonalnych, z uwzględnieniem płci i grup wieku, jest lepsza niż w pozostałej części województwa oraz że rynek pracy w ujęciu miejskich obszarów funkcjonalnych cechuje się wyraźnym zróżnicowaniem przestrzennym.

5.1. Ocena jakości oszacowań analizowanych wskaźników rynku pracy

Głównym celem badania empirycznego przeprowadzonego w ramach niniejszej pracy było oszacowanie podstawowych wskaźników rynku pracy dla niepublikowanego do tej pory poziomu agregacji, jakim jest przekrój miejskich obszarów funkcjonalnych z uwzględnieniem grup wiekowych oraz zmiennej płeć. Wśród rozważanych wskaźników uwagę skupiono na współczynniku aktywności ekonomicznej, wskaźniku zatrudnienia oraz stopie bezrobocia. Szacunku dokonano na podstawie estymatora $GLSM_{3P,L}$. Szczegółowe zestawienie wszystkich otrzymanych wyników uzyskanych dla trzech wskaźników wraz z oceną precyzji szacunku przedstawiono w aneksie w tabeli A.5. W niniejszym rozdziale ograniczono się do przeanalizowania w szerszym zakresie wyników otrzymanych dla najbardziej szczegółowego z rozpatrywanych przekrojów, tj. w podziale na płeć oraz cztery grupy wieku.

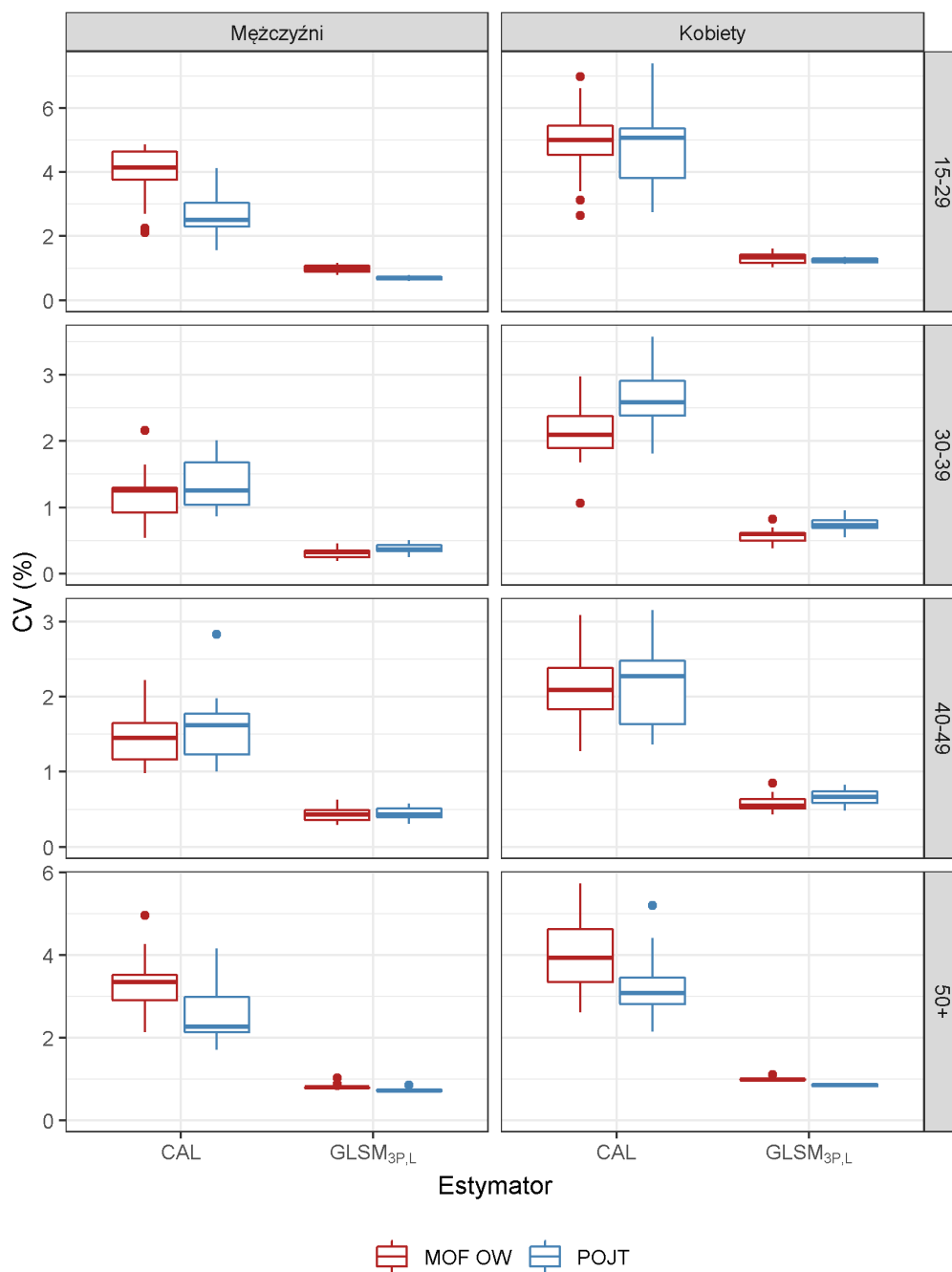
5.1.1. Analiza precyzji oszacowań wskaźników rynku pracy

W pierwszej kolejności ocenie poddana została precyzja szacunku wskaźników charakteryzujących rynek pracy wykorzystujących w swojej konstrukcji oceny estymatora $GLSM_{3P,L}$ w porównaniu ze wskaźnikami bazującymi na ocenach estymatora kalibracyjnego².

Jednym ze wskaźników charakteryzujących rynek pracy jest współczynnik aktywności ekonomicznej. Na rysunku 5.1 można zauważyć, że niezależnie od podziału na miejski obszar funkcjonalny (MOF OW) i pozostałą część jednostki terytorialnej (POJT), zarówno dla mężczyzn, jak i kobiet, w każdej z analizowanych grup wiekowych, uzyskane oceny precy-

²W przypadku rozpatrywanych wskaźników, które są ilorazami oszacowań dwóch wartości globalnych, należałoby de facto mówić o estymatorach ilorazowych. Ze względu na fakt, że w ich konstrukcji uwzględnia się oszacowania wartości globalnych z wykorzystaniem estymatora kalibracyjnego bądź $GLSM_{3P,L}$ w dalszej części wywodu, celem uproszczenia, estymatory te będą oznaczane odpowiednio przez $GLSM_{3P,L}$ oraz CAL .

zji CV estymatora kalibracyjnego oraz $GLSM_{3P,L}$ kształtowały się na poziomie niższym od przyjętego progu 10%.



Rysunek 5.1: Porównanie ocen precyzji współczynnika aktywności ekonomicznej

Uwaga: Poszczególne przekroje wieku przedstawione są na różnych skalach. Objaśnienie użytych skrótów umieszczone jest w tabelach 4.1, 4.2 oraz 4.3.

Źródło: Opracowanie własne.

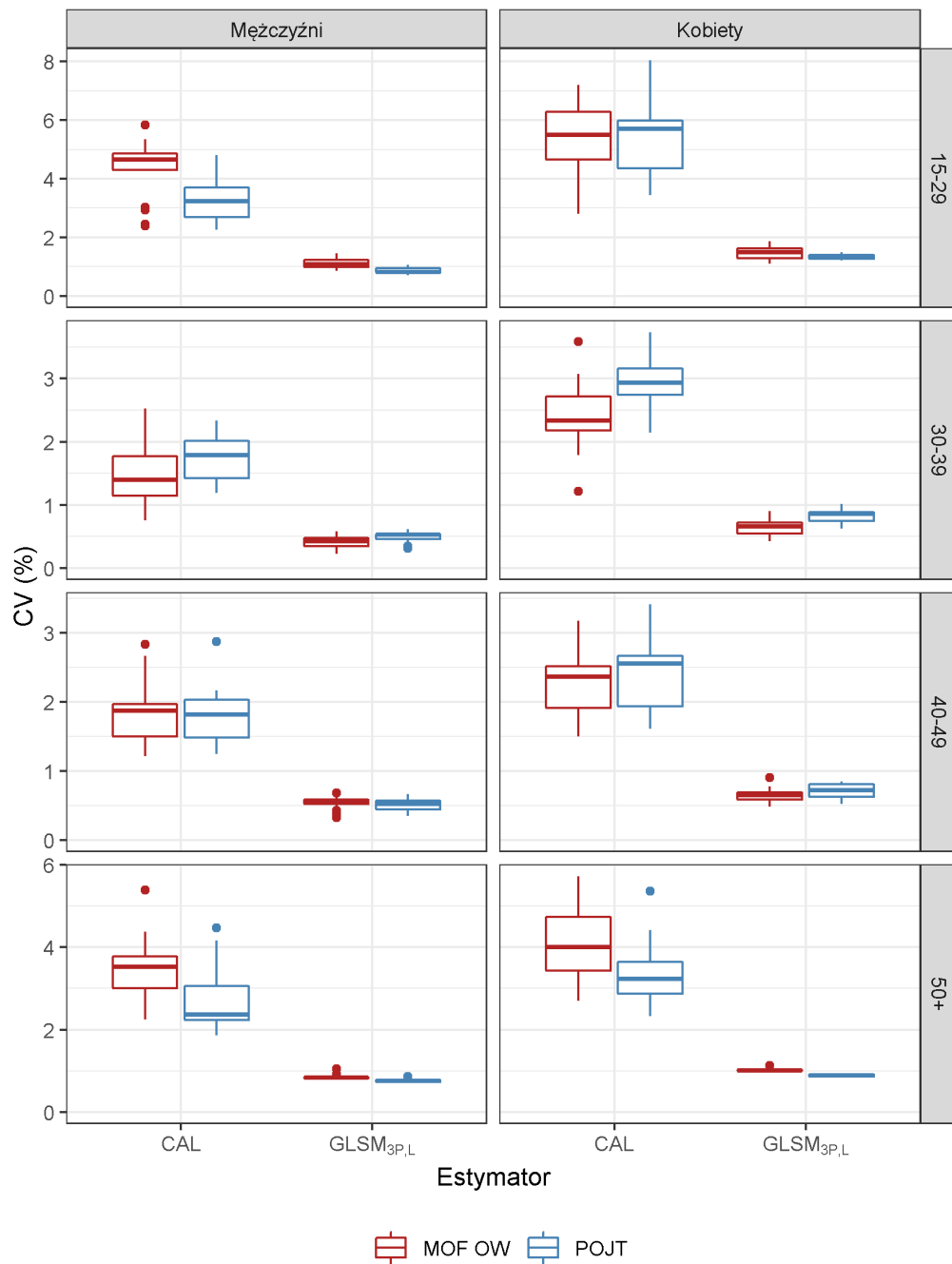
Widać jednak wyraźnie, że w każdym przypadku zastosowanie estymacji z wykorzystaniem $GLSM_{3P,L}$ dawało istotną poprawę precyzji. Nawet dla mężczyzn w grupie wieku 15–29

lat, gdzie uzyskano najwyższe wartości współczynnika CV , jego wartość nie przekroczyła 2%. Z kolei dla kobiet w tej samej grupie wiekowej skrajne oceny precyzji estymatora kalibracyjnego, zarówno dla MOF OW jak i POJT, oscylują koło 7%.

Zauważyć można również, że względny błąd oceny estymatora $GLSM_{3P,L}$ charakteryzuje się niewielkim zróżnicowaniem w porównaniu do estymatora kalibracyjnego. Widoczna jest również wyższa precyzja szacunków dla mężczyzn niezależnie od grupy wieku i zastosowanego estymatora.

Na rysunku 5.2 dokonano porównania ocen precyzji dla wskaźnika zatrudnienia. Porównując rysunki 5.1 oraz 5.2 można zauważyć, że relacje między ocenami precyzji współczynnika zatrudnienia dla obu estymatorów są podobne do tych, jakie uzyskano w przypadku współczynnika aktywności ekonomicznej. Jediną zauważalną różnicą jest to, że poziom współczynnika CV dla obu porównywanych estymatorów jest nieco wyższy w przypadku współczynnika aktywności zawodowej.

Kolejnym bardzo ważnym wskaźnikiem charakteryzującym rynek pracy, który został poddany analizie w kontekście oceny precyzji oszacowań, jest stopa bezrobocia (rysunek 5.3). Oceny błędów dla tego wskaźnika znacznie odbiegają od wartości uzyskanych dla współczynnika aktywności ekonomicznej oraz wskaźnika zatrudnienia. Dla wszystkich analizowanych przekrojów wartość współczynnika CV charakteryzująca precyzję estymatora kalibracyjnego (poza dwoma wartościami odstającymi dla mężczyzn i kobiet w wieku 15–29 lat) kształtuje się przynajmniej na poziomie 10%, dla którego, zgodnie z przyjętymi założeniami, należy zachować dużą ostrożność co do wiarygodności oszacowań. Jednak w większości przypadków współczynnik CV przekracza nieakceptowalny poziom 20%. Najwyższe wartości względnego błędu estymatora kalibracyjnego występują dla grup wiekowych 30–39 lat oraz 40–49 lat, w ramach których wyższe wartości otrzymano na obszarach MOF OW dla mężczyzn. W skrajnym przypadku sięgają one prawie 100%. Takie wysokie wartości współczynnika zmienności oszacowań dla estymatora kalibracyjnego wynikają z konstrukcji wskaźnika stopy bezrobocia. W grupach wiekowych 30–39 lat oraz 40–49 lat, szczególnie na obszarze MOF OW, zjawisko bezrobocia jest na tyle niskie w stosunku do liczby osób aktywnych zawodowo, że ma to wpływ na oceny precyzji.

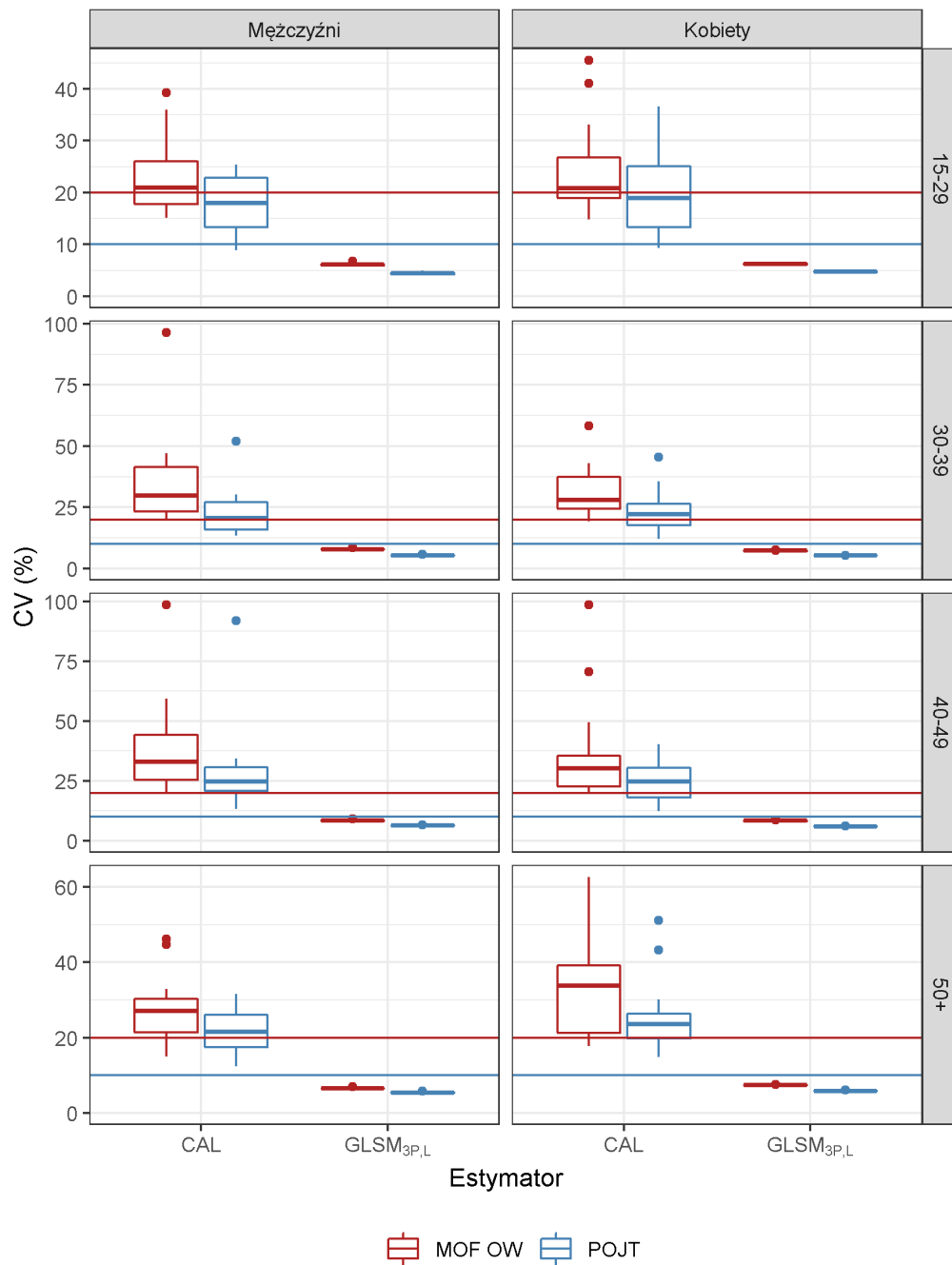


Rysunek 5.2: Porównanie ocen precyzji wskaźnika zatrudnienia

Uwaga: Poszczególne przekroje wieku przedstawione są na różnych skalach. Objasnienie użytych skrótów umieszczone jest w tabelach 4.1, 4.2 oraz 4.3.

Źródło: Opracowanie własne.

Z kolei oceny precyzji oszacowań tego wskaźnika uzyskanych przy pomocy estymatora $GLSM_{3P,L}$ są bardzo zbliżone do siebie i zaobserwować można wyraźny zysk na precyzji w stosunku do estymatora kalibracyjnego we wszystkich analizowanych przekrojach.

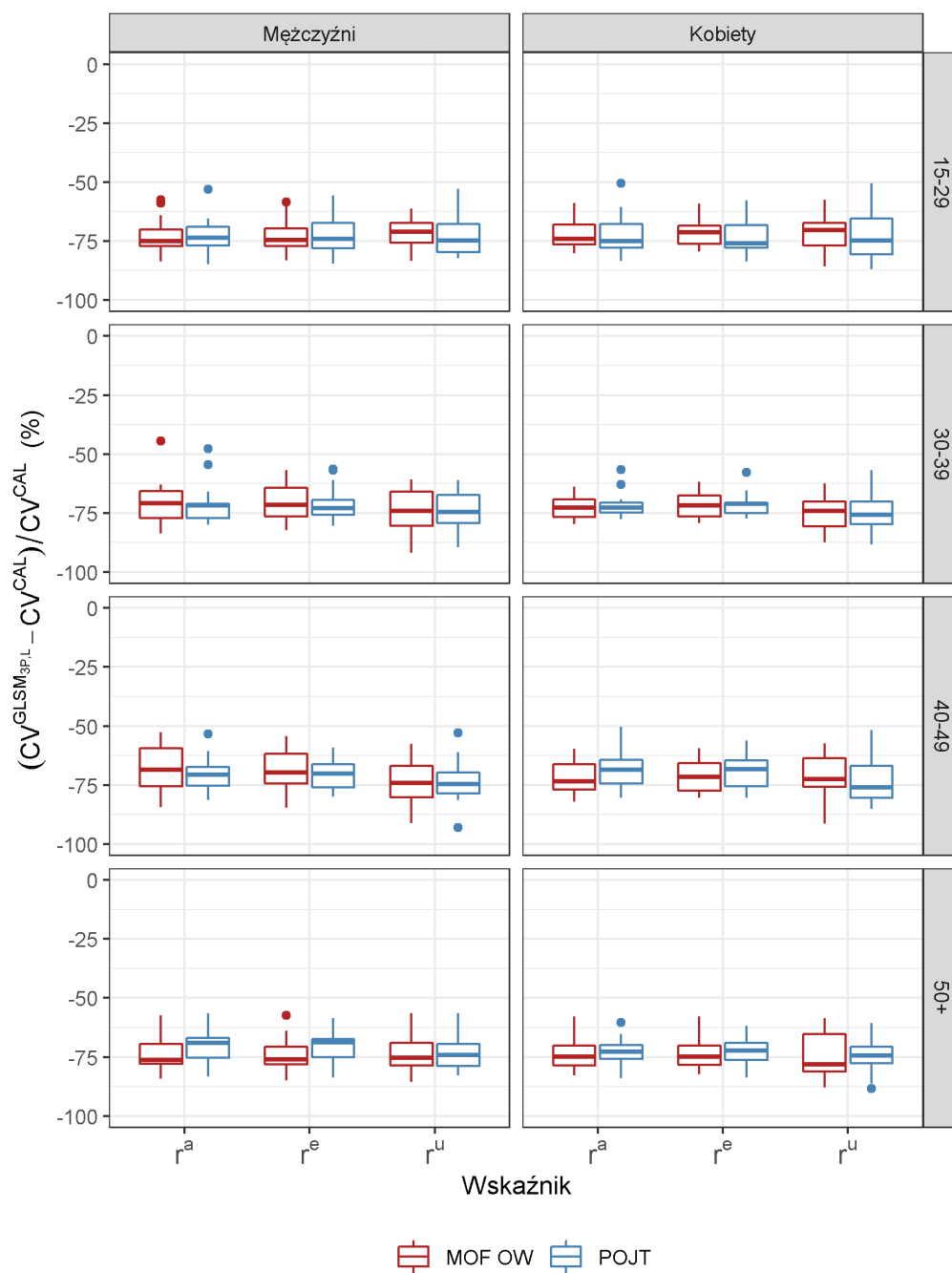


Rysunek 5.3: Porównanie ocen precyzji dla stopy bezrobocia

Uwaga: Poszczególne przekroje wieku przedstawione są na różnych skalach. Objaśnienie użytych skrótów umieszczone jest w tabelach 4.1, 4.2 oraz 4.3.

Źródło: Opracowanie własne.

Z zestawienia przedstawionego na rysunku 5.4 wynika, iż niezależnie od wybranego wskaźnika oraz jednostki przestrzennej (MOF OW, POJT) zastosowanie estymatora $GLSM_{3P,L}$ wpływa na poprawę precyzji szacunków.



Rysunek 5.4: Zestawienie zysku na precyzji dla poszczególnych wskaźników w analizowanych przekrojach wyznaczonych z wykorzystaniem estymatora $GLSM_{3P,L}$ w porównaniu z oszacowaniami za pomocą estymatora kalibracyjnego

Objaśnienia: r^a – współczynnik aktywności ekonomicznej, r^e – wskaźnik zatrudnienia, r^u – stopa bezrobocia

Źródło: Opracowanie własne.

Poprawa ta kształtuje się na co najmniej 50% poziomie, a w niektórych przypadkach przekracza nawet 90%. Wyjątek stanowią dwie obserwacje dla współczynnika aktywności

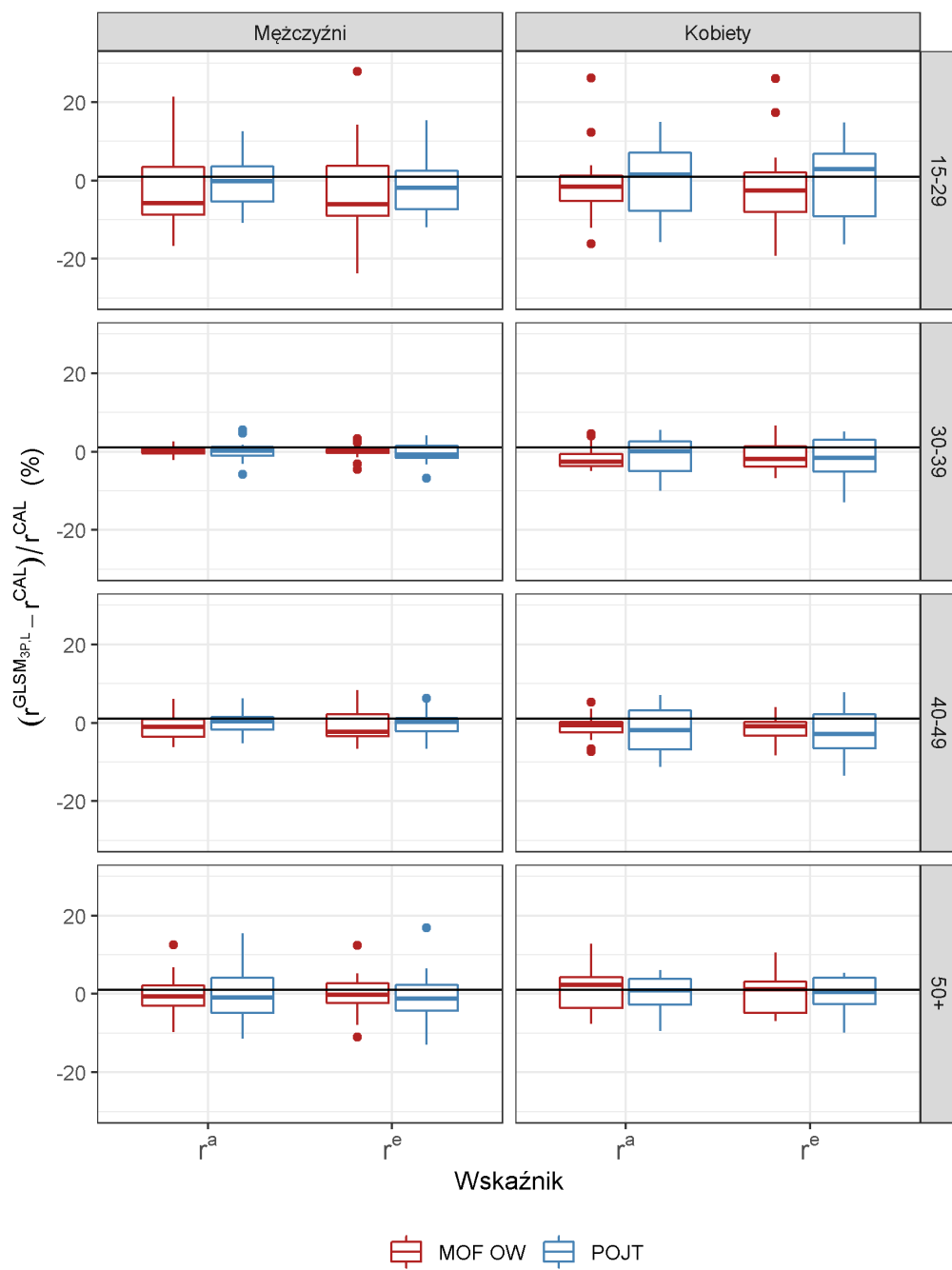
ekonomicznej ludności w grupie wieku 30–39 lat. Poziom poprawy precyzji jest na ogół zbliżony, niezależnie od estymowanego wskaźnika i rozpatrywanego przekroju.

Dokonując podsumowania oszacowań precyzji można jednoznacznie stwierdzić, że estymacja typu SPREE cechuje się lepszą precyzją szacunków charakterystyk rynku pracy (współczynnik aktywności ekonomicznej, wskaźnik zatrudnienia, stopa bezrobocia) w porównaniu z estymacją kalibracyjną stosowaną w badaniu BAEL. Pozytywnie zatem zwerdyktowano drugą hipotezę badawczą sformułowaną we wstępie pracy.

5.1.2. Analiza porównawcza wartości oszacowań wskaźników rynku pracy

W podrozdziale dokonano oceny zgodności oszacowań wszystkich trzech rozważanych wskaźników rynku pracy uzyskanych z wykorzystaniem estymatora $GLSM_{3P,L}$ oraz CAL w analizowanym przekroju uwzględniającym płeć, grupy wieku oraz wymiar przestrzenny (MOF OW oraz POJT). Wyniki zaprezentowane zostały na rysunkach 5.5 oraz 5.6. Celem przeprowadzonej analizy zgodności oszacowań było sprawdzenie na ile zbieżne (bądź nie) są oceny wskaźników uzyskanych dla dwóch rozważanych estymatorów.

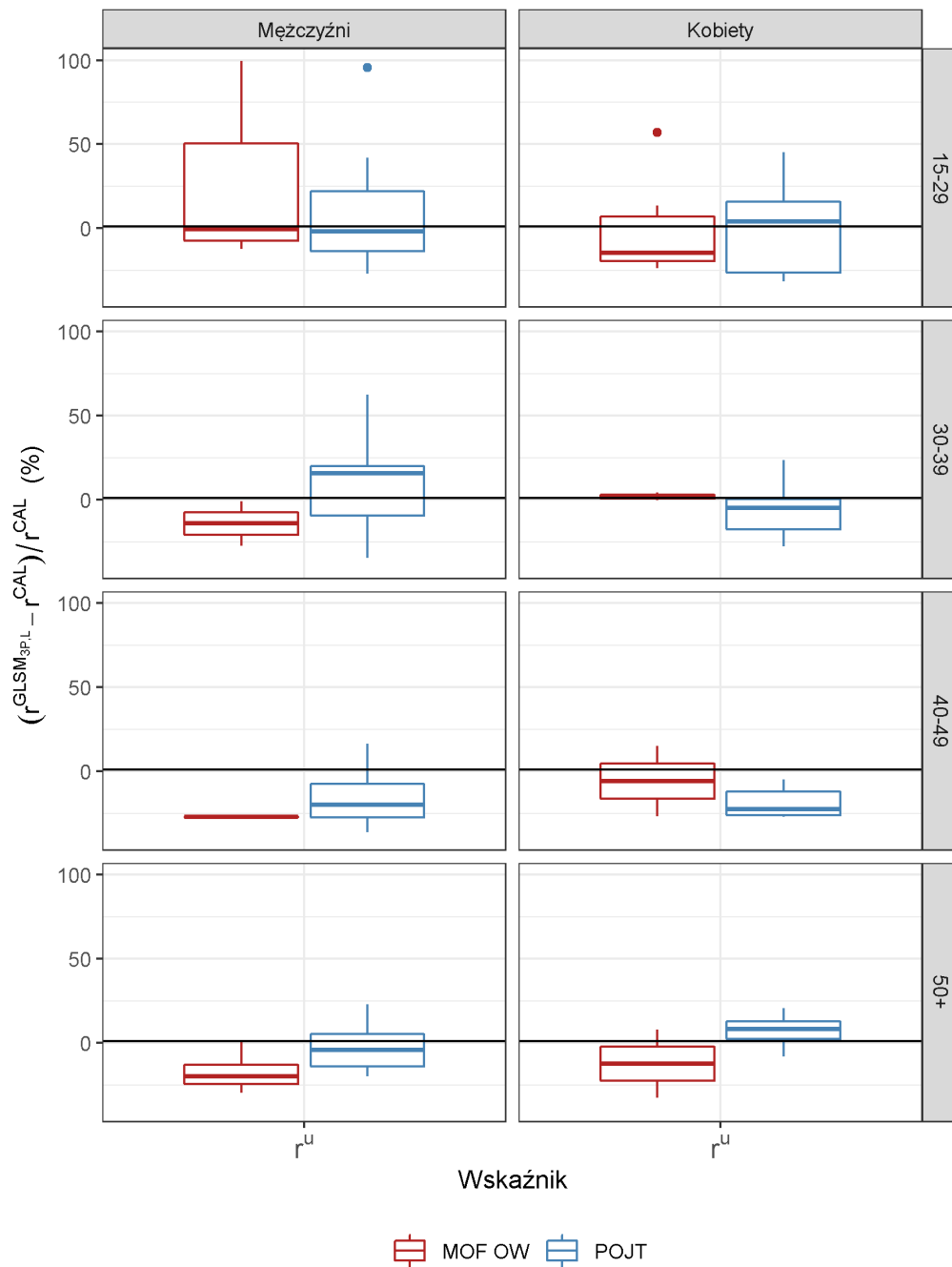
Na rysunku 5.5 przedstawiono względne różnice pomiędzy oszacowanymi wartościami współczynnika aktywności ekonomicznej oraz wskaźnika zatrudnienia uzyskanymi za pomocą estymatora $GLSM_{3P,L}$ i kalibracyjnego. Dodatnia różnica świadczy o wyższych oszacowaniach wskaźników z użyciem estymatora $GLSM_{3P,L}$ w porównaniu do użycia estymatora kalibracyjnego. Różnica ujemna oznacza sytuację odwrotną. Największe wartości, sięgające niemal 30% i świadczące o największych rozbieżnościach, odnotowano w grupie wiekowej 15–29 lat. Z kolei największa zgodność oszacowań pomiędzy porównywanymi estymatorami występuje dla grupy wieku 30–39 oraz 40–49 lat. Można również zauważyć, że płeć nie wpływa na zróżnicowanie oszacowań dla obu estymatorów.



Rysunek 5.5: Procentowe różnice pomiędzy oszacowaniami wskaźników uzyskanymi z wykorzystaniem estymatora $GLSM_{3PL}$ i kalibracyjnego

Objaśnienia: r^a – współczynnik aktywności ekonomicznej, r^e – wskaźnik zatrudnienia.

Źródło: Opracowanie własne.



Rysunek 5.6: Procentowe różnice pomiędzy oszacowaniami stopy bezrobocia uzyskanymi z wykorzystaniem estymatora $GLSM_{3P,L}$ i kalibracyjnego

Objaśnienia: r^u – stopa bezrobocia

Ze względu na nieakceptowalny poziom precyzji ($CV \geq 20\%$) usunięto 189 z 272 oszacowań.

Źródło: Opracowanie własne.

Różnice procentowe pomiędzy oszacowanymi wartościami stopy bezrobocia uzyskanymi za pomocą estymatora $GLSM_{3P,L}$ i kalibracyjnego przedstawiono na osobnym rysunku (por. 5.6). Wynikało to przede wszystkim z tego, że w przypadku tego wskaźnika odnotowano zdecydowanie wyższe wartości różnic. Także i dla stopy bezrobocia największe procentowe różnice występują w grupie osób 15–29 lat; w grupie mężczyzn dla dwóch obserwacji względne różnice sięgają nawet 100%. Dodatkowo w przypadku MOF OW dla przekroju mężczyzn w wieku 15–29 lat procentowa różnica dla kwartyła górnego wynosi 50%. Zauważyć można również, że w przypadku mężczyzn w MOF OW dla wszystkich grup wieku, z wyjątkiem grupy 15–29 lat, wartości stopy bezrobocia uzyskane za pomocą estymatora $GLSM_{3P,L}$ są niższe niż przy zastosowaniu estymatora kalibracyjnego.

W uzupełnieniu tej analizy w aneksie na rysunkach A.8, A.9 oraz A.10 przedstawiono diagramy rozrzutu, zawierające porównanie szacunków analizowanych wskaźników wyznaczonych z wykorzystaniem estymatora $GLSM_{3P,L}$ oraz kalibracyjnego.

5.2. Ocena sytuacji na rynku pracy – ujęcie przestrzenne

Druga część analizy obejmowała charakterystykę rozkładów oszacowań otrzymanych dla każdego z analizowanych w pracy wskaźników. W pierwszej kolejności pod uwagę wzięto współczynnik aktywności ekonomicznej. Rozkład ocen estymatorów dla tego współczynnika przedstawiono na wykresach pudełkowych (por. rys. 5.7). Na każdym z rysunków po lewej stronie umieszczono wykres pudełkowy przedstawiający wartości współczynnika dla miejskich obszarów funkcjonalnych ośrodków wojewódzkich (MOF OW). Po stronie prawej zaprezentowany jest wykres pudełkowy opisujący sytuację w pozostałej części jednostki terytorialnej (POJT). Poszczególne obszary MOF OW oraz okalające je pozostałe części (województwa lub podregionu) występujące na obu wykresach pudełkowymi są połączone niebieskimi liniami. Grubsza niebieska linia łącząca punkty pomiędzy wykresami pudełkowymi oznacza, że wskaźnik przyjął niższą wartość na obszarze danego MOF OW niż dla jego POJT. Natomiast cieńsza niebieska linia oznacza sytuację odwrotną. Dodat-

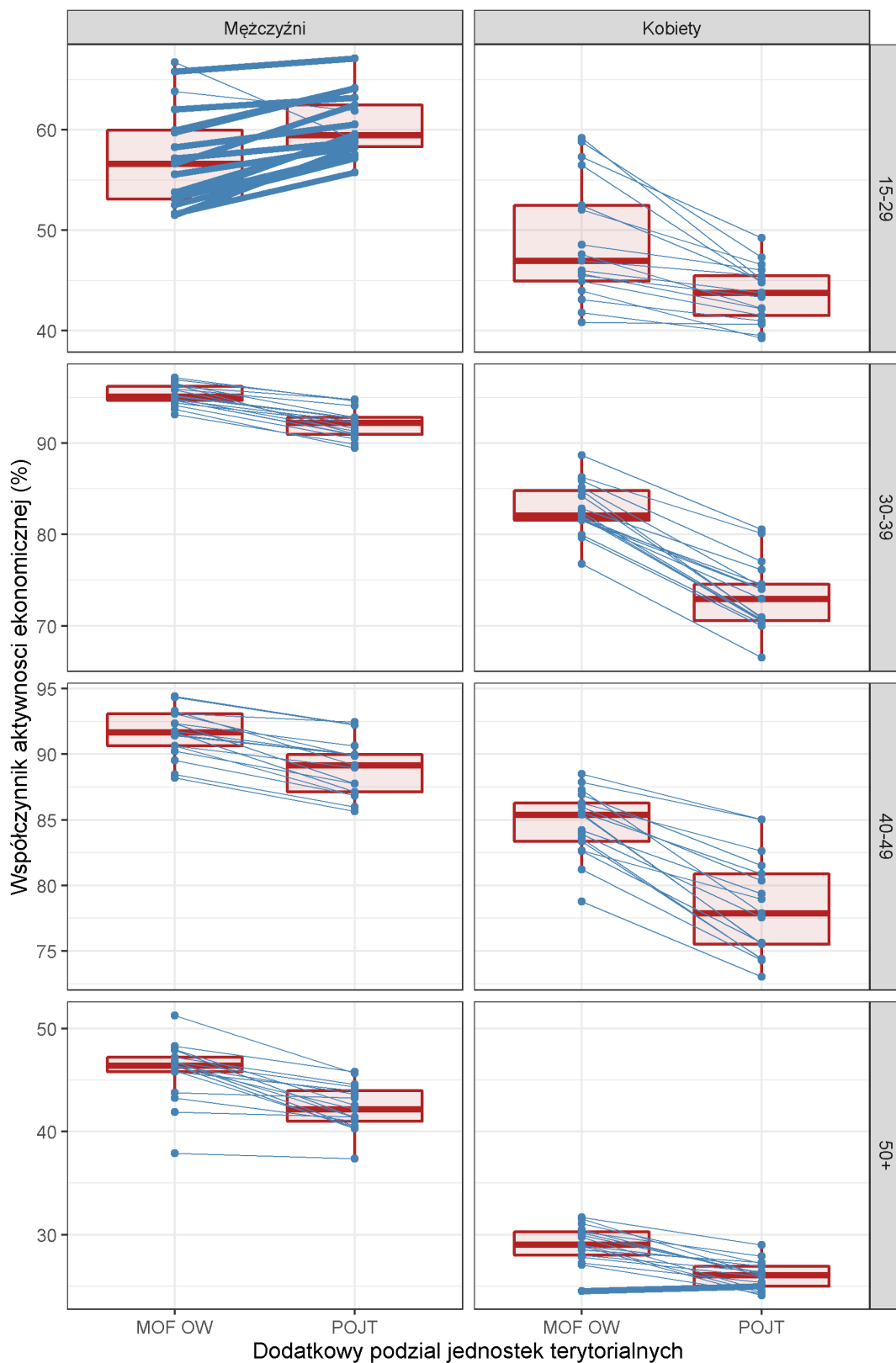
kowo w celu przestrzennej oceny analizowanego wskaźnika zaprezentowano kartogramy przedstawione na rysunku 5.8³.

Analizując wykresy pudełkowe na rysunku (5.7) oraz kartogramy na rysunku (5.8) można zauważyć, że najniższe wartości wskaźnika aktywności ekonomicznej ludności dotyczą grupy wiekowej 50 lat i więcej, niezależnie od płci. Warto podkreślić, że wartości tego wskaźnika są niższe dla kobiet we wszystkich grupach wiekowych, a szczególnie w grupie 50+. Wartości najwyższe występują w przedziale wieku 30–39 i 40–49 lat, niezależnie od płci.

Przechodząc do bardziej szczegółowych analiz z uwzględnieniem podziału badanej zbiorowości na MOF OW oraz POJT, na rysunku 5.7 możemy dostrzec, że w ramach trzech grup wiekowych (30–39, 40–49, 50 lat i więcej), biorąc pod uwagę również płeć, sytuacja w MOF OW, wyrażona współczynnikiem aktywności ekonomicznej, jest lepsza niż dla POJT. Jedyne wyjątkiem stanowi MOF OW Katowic, dla którego wartość wskaźnika dla kobiet w wieku 50 lat i więcej (24,5%) jest nieco niższa niż dla pozostałej części województwa (25,0%).

Z odwrotną relacją mamy do czynienia w przypadku oszacowań dla mężczyzn w wieku 15–29. Wartości współczynnika aktywności ekonomicznej są wyraźnie wyższe poza miejskimi obszarami funkcjonalnymi (wyjątek stanowią MOF OW Wrocławia oraz MOF OW Katowic). To przekłada się na wyższe wartości kwartyli tego wskaźnika dla obszarów POJT w porównaniu do rozkładu oszacowań dla MOF OW.

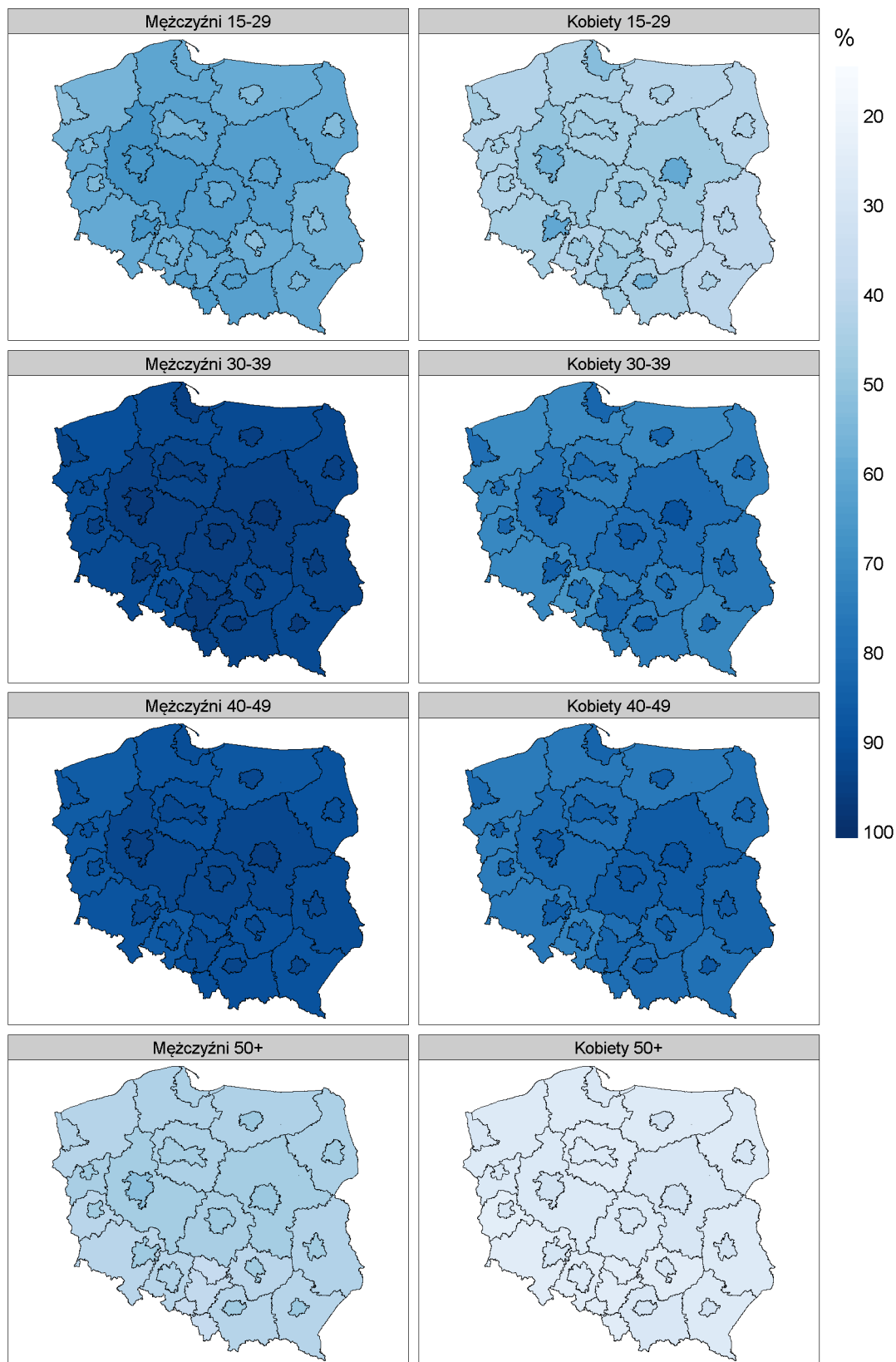
³Wartości wskaźników, które przywoływane są w prowadzonych analizach pochodzą z tabeli A.5 znajdującej się w aneksie, która zawiera wszystkie oszacowane wartości wskaźników.



Rysunek 5.7: Porównanie rozkładu współczynnika aktywności ekonomicznej pomiędzy MOF OW a POJT w ramach poszczególnych przekrojów

Uwaga: Oszacowania dla poszczególnych kategorii wieku przedstawione są na różnych skalach.

Źródło: Opracowanie własne.

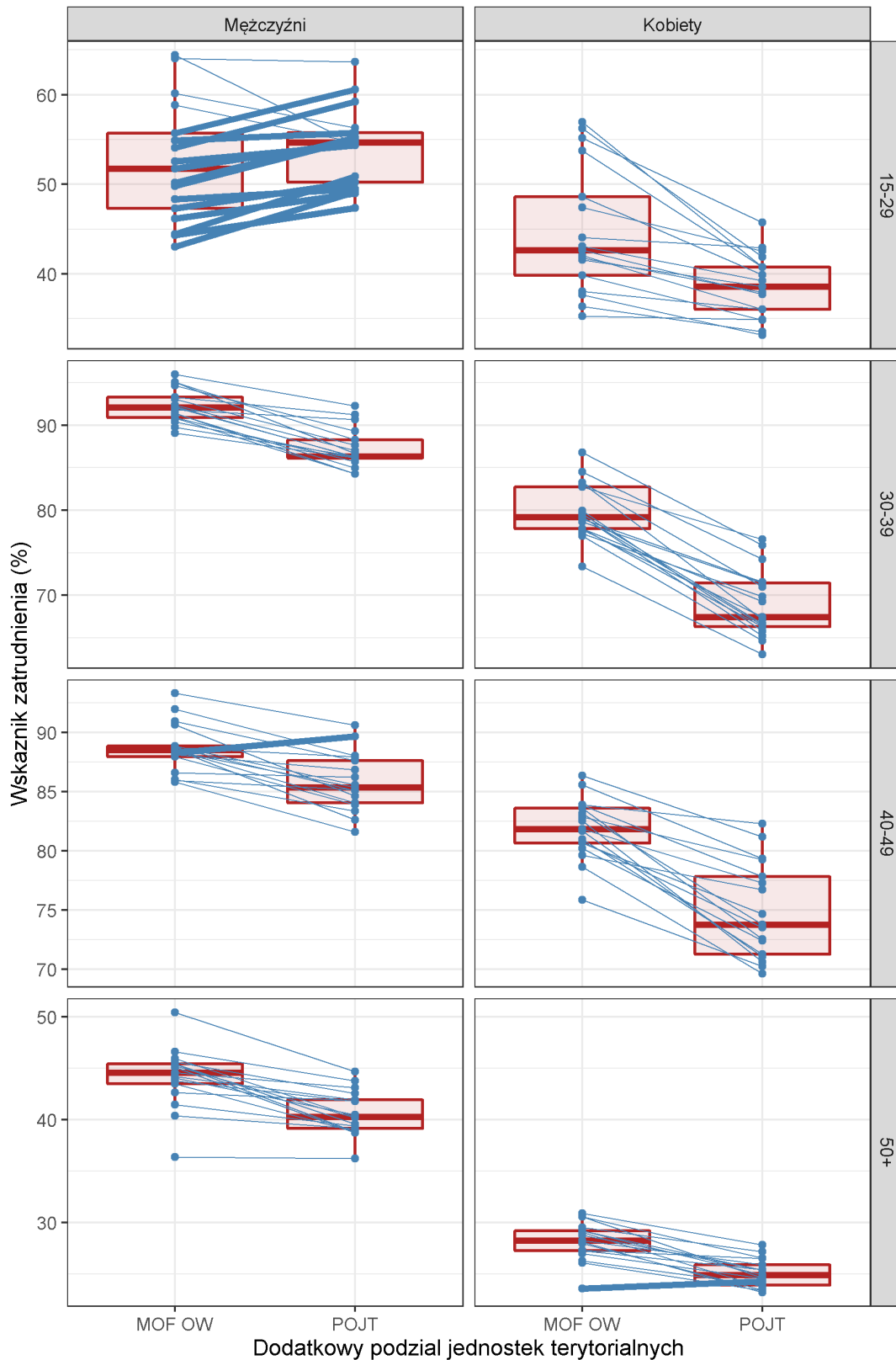


Rysunek 5.8: Przestrzenne zróżnicowanie współczynnika aktywności ekonomicznej w analizowanych przekrojach

Źródło: Opracowanie własne.

Analizując łącznie wykresy pudełkowe (por. rysunek 5.7) oraz kartogramy na rysunku (5.8), a także posiłkując się tabelą A.5 załączoną w aneksie można zauważyć, że największa różnica pomiędzy wartościami współczynnika aktywności ekonomicznej MOF OW i POJT, wskazująca na lepszą sytuację na rynku pracy w miejskich obszarach funkcjonalnych w porównaniu z pozostałą częścią województwa występuje w przypadku MOF OW Wrocławia, dla grupy kobiet w wieku 15–29 (MOF OW 59,2%, POJT 45,1%). Z kolei największe różnice na korzyść POJT w porównaniu do MOF OW odnotowano dla grupy mężczyzn w wieku 15–29 w przypadku MOF OW Kielc (odpowiednio 58,3% oraz 51,5%). Największa różnica pomiędzy miejskimi obszarami funkcjonalnymi ośrodków wojewódzkich pojawia się dla kobiet w wieku 15–29 lat w przypadku MOF OW Wrocławia (59,2%) i MOF OW Kielc (40,8%).

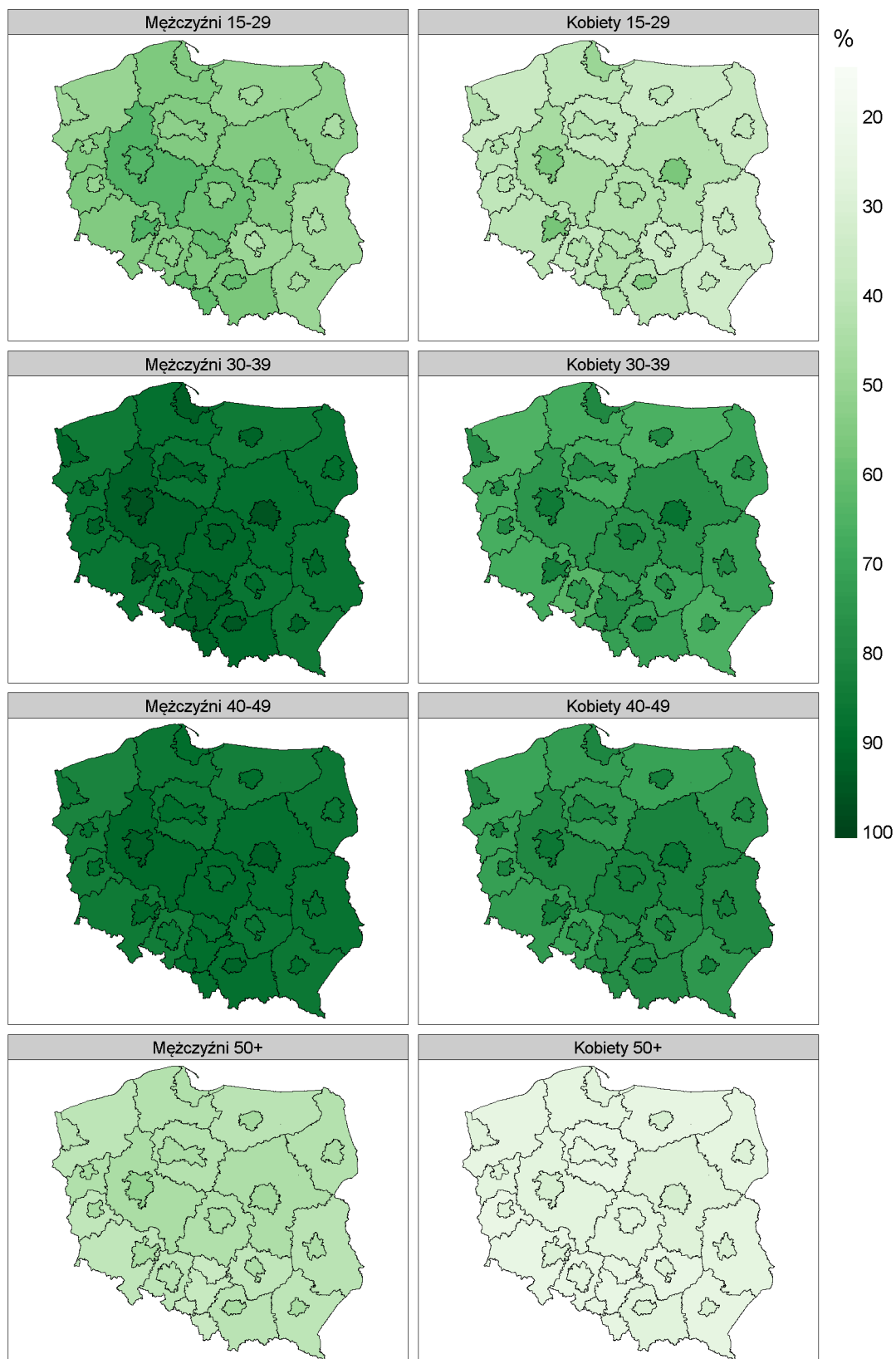
Drugim, z trzech uwzględnionych w pracy wskaźników, charakteryzujących sytuację na rynku pracy był wskaźnik zatrudnienia. Analizując wartości oszacowań tego wskaźnika przedstawione na rysunkach 5.9 oraz 5.10, można zauważyć, że zachowane zostały ogólne relacje pomiędzy domenami zaobserwowane w odniesieniu do opisanego wcześniej współczynnika aktywności ekonomicznej. Różnica dotyczyła jedynie poziomu wskaźnika zatrudnienia, który ze względu na jego konstrukcję, kształtuje się na niższym poziomie. Szczególnie niski poziom tego wskaźnika zanotowano dla grupy kobiet w wieku 50+ zarówno dla MOF OW jak i POJT. Podobnie jak przy współczynniku aktywności zawodowej największa różnica pomiędzy wartościami wskaźnika MOF OW i POJT, wskazująca na lepszą sytuację na rynku pracy w miejskich obszarach funkcjonalnych w porównaniu z pozostałą częścią województwa, występuje MOF OW Wrocławia, dla kobiet w wieku 15–29 (MOF OW 56,9%, POJT 40,8%).



Rysunek 5.9: Porównanie rozkładu wskaźnika zatrudnienia pomiędzy MOF OW a POJT w ramach poszczególnych przekrojów

Uwaga: Oszacowania dla poszczególnych kategorii wieku przedstawione są na różnych skalach.

Źródło: Opracowanie własne.

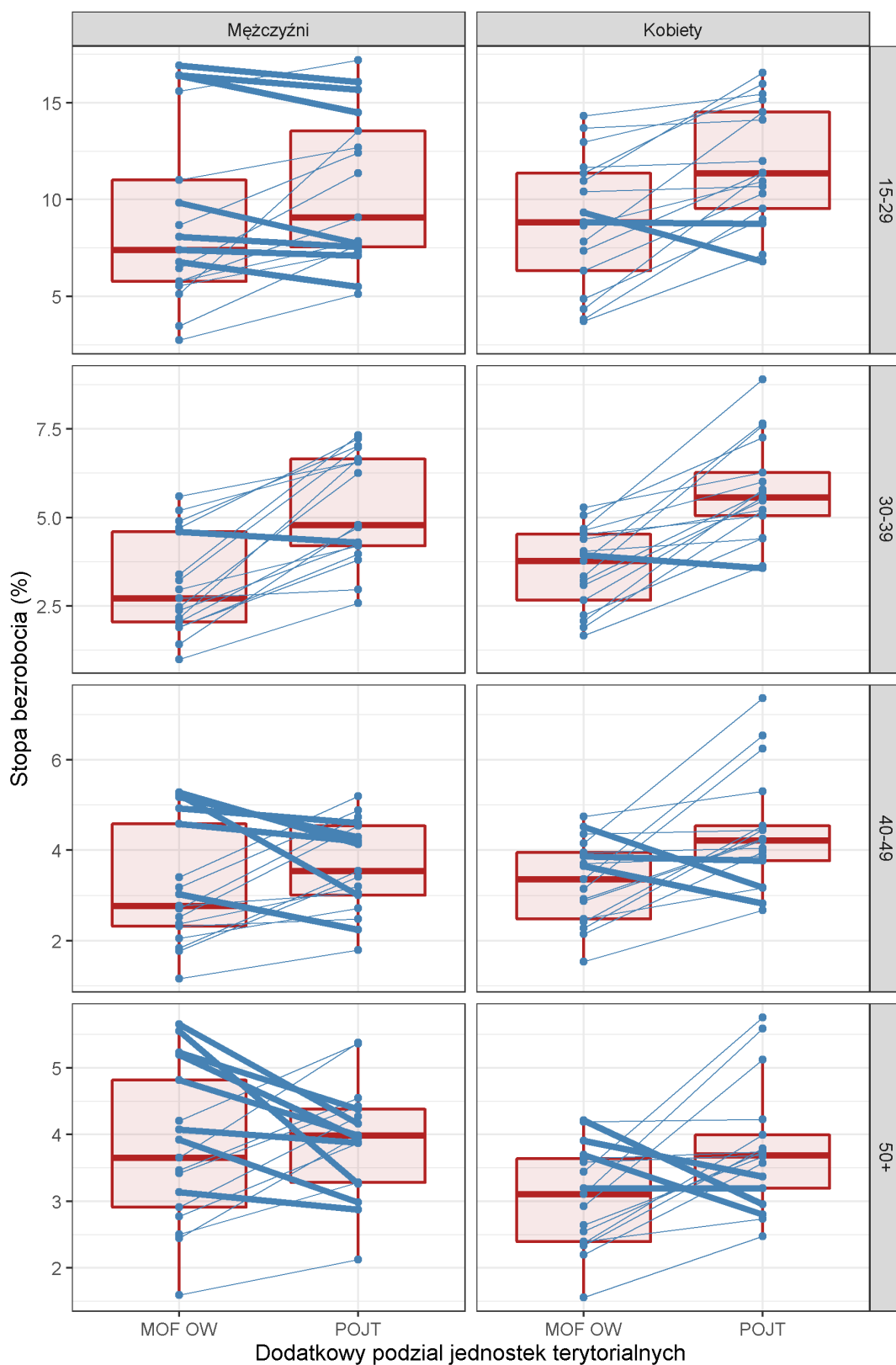


Rysunek 5.10: Przestrzenne zróżnicowanie wskaźnika zatrudnienia ludności w analizowanych przekrojach

Źródło: Opracowanie własne.

Natomiast największe różnica na korzyść POJT w porównaniu do MOF OW odnotowano dla MOF OW Białegostoku w grupie mężczyzn w wieku 15–29 (odpowiednio 44,4% oraz 50,9%). Największe różnice pomiędzy miejskimi obszarami funkcjonalnymi pojawiają się między MOF OW Wrocławia a MOF OW Kielc, dla kobiet w grupie wiekowej 15–29 lat, gdzie wartości wskaźnika zatrudnienia wyniosły odpowiednio 56,9% i 35,3%.

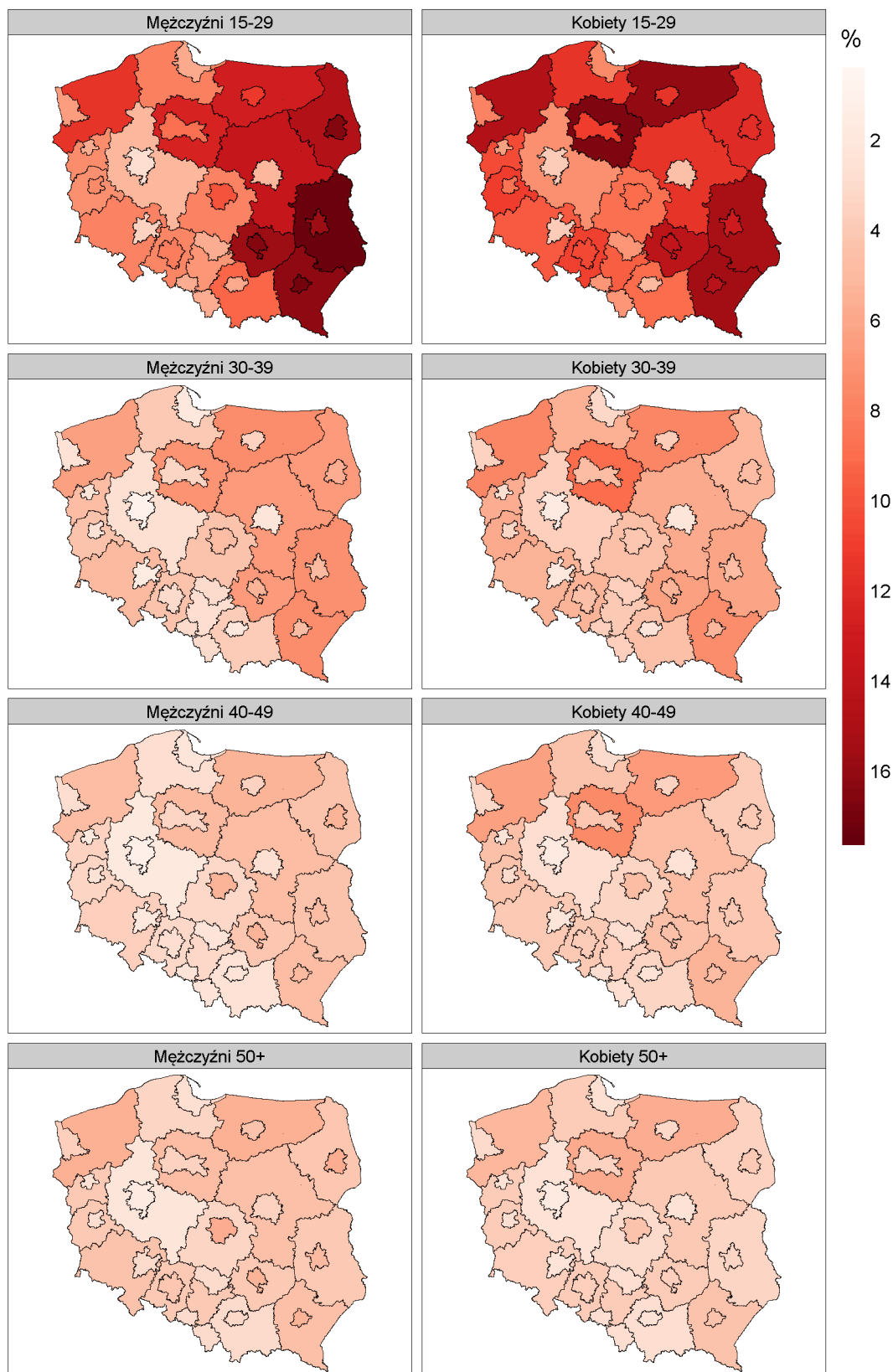
Ostatnim uwzględnionym w pracy wskaźnikiem charakteryzującym sytuację na rynku pracy była stopa bezrobocia. Rozkład ocen estymatorów dla tego wskaźnika przedstawiono w postaci wykresów pudełkowych na rysunku 5.11 oraz na kartogramach (por. rysunek 5.12). Obszar zmienności wartości stopy bezrobocia zdecydowanie odbiega od tego, co zostało zarejestrowane w przypadku dwóch pozostałych wskaźników. Zarówno wartości median, jak i dolnych i górnych kwartyli w ramach poszczególnych przekrojów wskazują na przeciętnie wyższy poziom stopy bezrobocia na obszarach POJT w porównaniu z MOF OW (wyjątek stanowi jedynie grupa mężczyzn w wieku 50 lat i więcej, dla których wartość kwartyli trzeciego dla POJT jest niższa niż jej odpowiednik dla MOF OW). Widać jednak, że w ramach każdego z przekrojów przynajmniej jeden MOF OW charakteryzuje się wyższą stopą bezrobocia aniżeli obszar go otaczający (POJT). Najwięcej takich przypadków odnotowano w grupie mężczyzn w wieku 15–29 lat oraz 50 lat i więcej. Nie można jednak jednoznacznie wskazać rodzaju domeny, w której sytuacja byłaby zdecydowanie lepsza na obszarach POJT niż MOF OW, jak to miało miejsce w odniesieniu do współczynnika aktywności ekonomicznej oraz wskaźnika zatrudnienia. Dla tych dwóch wskaźników taką domenę stanowili mężczyźni w wieku 15–29 lat. W przypadku stopy bezrobocia w domenie tej występują największe różnice w poziomie wskaźnika pomiędzy miejskimi obszarami funkcjonalnymi: od 2,7% dla MOF OW Poznania do 16,9% dla MOF OW Rzeszowa. Z kolei największa różnica na korzyść MOF OW w porównaniu do POJT została odnotowana dla MOF OW Warszawy w grupie mężczyzn w wieku 15–29 lat. Stopa bezrobocia wynosiła tam odpowiednio 5,1% i 13,5%. Z sytuacją odwrotną mamy do czynienia w przypadku MOF OW Katowic, gdzie stopa bezrobocia dla kobiet w grupie wieku 15–29 lat wynosi 9,3% a dla pozostałej części województwa przyjmuje wartość 6,8%.



Rysunek 5.11: Porównanie rozkładu stopy bezrobocia pomiędzy MOF OW a POJT w ramach poszczególnych przekrojów

Uwaga: Oszacowania dla poszczególnych kategorii wieku przedstawione są na różnych skalach.

Źródło: Opracowanie własne.



Rysunek 5.12: Przestrzenne zróżnicowanie stopy bezrobocia w analizowanych przekrojach

Źródło: Opracowanie własne.

Dokonując podsumowania oszacowań poszczególnych wskaźników (współczynnik aktywności ekonomicznej, wskaźnik zatrudnienia, stopa bezrobocia) i ich przestrzennego zróżnicowania można stwierdzić, że sytuacja na rynku pracy w miejskich obszarach funkcjonalnych z uwzględnieniem płci i grup wieku jest na ogół lepsza niż w pozostałej części województwa. Świadczą o tym wyższe wartości współczynnika aktywności ekonomicznej oraz wskaźnika zatrudnienia (z wyjątkiem mężczyzn w grupie wieku 15–29) w MOF OW w porównaniu z POJT. Z kolei w odniesieniu do stopy bezrobocia obserwuje się niższe wartości tego wskaźnika w MOF OW w porównaniu z POJT (mierzone medianą oszacowań), chociaż w każdym z analizowanych przekrojów można wskazać POJT, w którym wartość stopy bezrobocia była niższa niż dla odpowiadającego mu MOF OW. Sformułowana we wstępie trzecia hipoteza, że sytuacja na rynku pracy na obszarach funkcjonalnych miast wojewódzkich, mierzona oszacowanymi wskaźnikami, jest na ogół lepsza niż w pozostałej części okalającego go obszaru, znalazła w zdecydowanej większości przypadków swoje potwierdzenie. Podano również największe różnice w wartościach wskaźników pomiędzy analizowanymi MOF OW, co w pewnej mierze potwierdza czwartą hipotezę badawczą i świadczy o zróżnicowaniu przestrzennym rynku pracy pomiędzy tymi obszarami.

5.3. Taksonomiczna analiza zróżnicowania rynku pracy w miejskich obszarach funkcjonalnych ośrodków wojewódzkich

Na potrzeby oceny stopnia zróżnicowania sytuacji panującej na rynku pracy w miejskich obszarach funkcjonalnych ośrodków wojewódzkich wykorzystano wybrane metody analizy taksonomicznej: Topsis oraz hierarchiczną metodę grupowania Warda (Walesiak i Gatnar, 2009). Miało to na celu weryfikację postawionej hipotezy badawczej głoszącej, że rynek pracy w ujęciu miejskich obszarów funkcjonalnych ośrodków wojewódzkich cechuje się wyraźnym zróżnicowaniem przestrzennym. Podobnie, jak w przypadku estymacji typu SPREE, w analizie ograniczono się do miejskich obszarów funkcjonalnych ośrodków wojewódzkich (17 jednostek przestrzennych) wyznaczonych do realizacji instrumentu ZIT w perspektywie finansowej UE na lata 2014–2020.

W charakterze zmiennych diagnostycznych wykorzystano oszacowania trzech wskaźników: współczynnika aktywności ekonomicznej, wskaźnika zatrudnienia oraz stopy bezrobocia uzyskanych za pomocą estymatora $GLSM_{3P,L}$ w przekroju płci oraz czterech grup wieku. Wstępnie rozważano zatem 24 zmienne diagnostyczne (3 wskaźniki x 2 kategorie płci x 4 kategorie grup wieku) odnoszące się do 17 wyróżnionych w pracy miejskich obszarów funkcjonalnych ośrodków wojewódzkich. Zmienne dotyczące współczynnika aktywności ekonomicznej oraz wskaźnika zatrudnienia określone zostały jako stymulanty, a te związane ze stopą bezrobocia jako destymulanty.

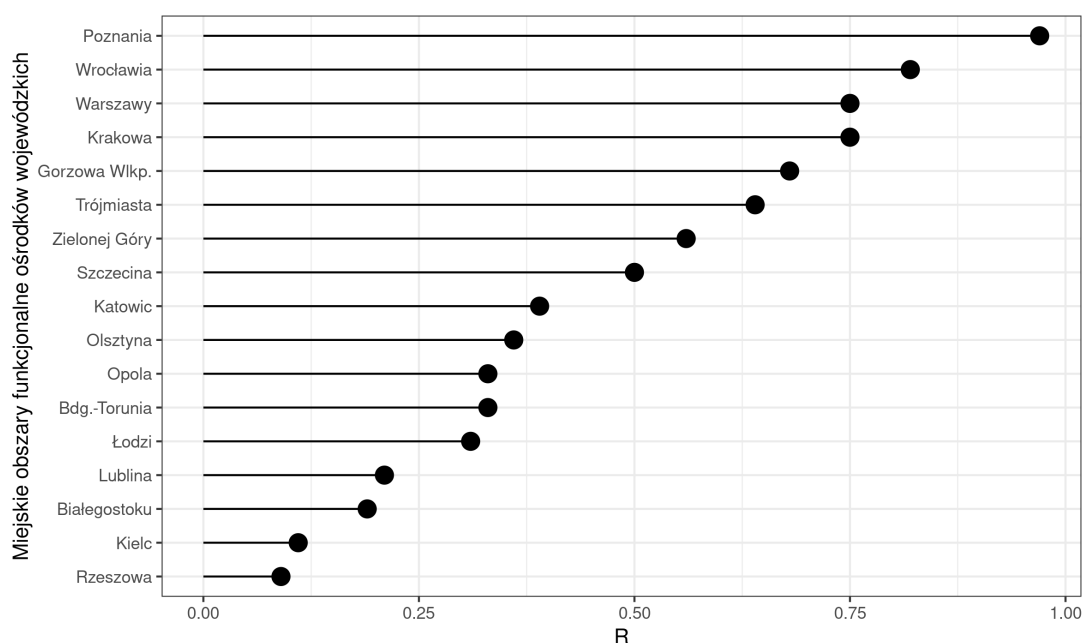
Na potrzeby utworzenia rankingu miejskich obszarów funkcjonalnych ośrodków wojewódzkich ze względu na sytuację panującą na rynku pracy skorzystano z metody Topsis (ang. *The Technique for Order of Preference by Similarity to Ideal Solution*) (Yoon i Hwang, 1981). Z kolei ich grupowania dokonano z zastosowaniem hierarchicznej metody Warda szeroko wykorzystywanej w tego typu analizach (Ward Jr, 1963).

W pierwszym kroku dokonano selekcji najważniejszych zmiennych diagnostycznych. Na potrzeby realizacji tego etapu wykorzystano metodę HINoV (Carmone Jr i in., 1999; Korzeniewski, 2012; Walesiak, 2008), której celem jest dobór mniejszej liczby zmiennych z pierwotnego zbioru przez eliminację tych, które zakłócają istniejącą strukturę klas w badanej zbiorowości obiektów. W wyniku zastosowania metody HINoV eliminacji zmiennych zakłócających w końcowym zbiorze zmiennych diagnostycznych pozostało 14⁴.

Na rysunku 5.13 przedstawiono ranking miejskich obszarów funkcjonalnych ośrodków wojewódzkich uzyskany z wykorzystaniem metody Topsis i ograniczonego zbioru zmiennych diagnostycznych. Można zauważyć, że zdecydowanie najlepszą sytuacją na rynku pracy charakteryzują się miejskie obszary funkcjonalne skupione wokół dużych miast wojewódzkich: Poznania, Wrocławia, Warszawy oraz Krakowa. Wyjątek stanowi tutaj miejski obszar funkcjonalny, którego miastem centralnym jest Łódź. Najkorzystniejsza sytuacja na rynku pracy w tych czterech miejskich obszarach funkcjonalnych jest związana bezpośrednio z dobrą sytuacją na rynku pracy w miastach centralnych stanowiących biegun rozwoju danego obszaru funkcjonalnego i jego pozytywnym wpływem na przemiany i rozwój gmin

⁴Były to następujące zmienne: ra_K_15_29, ra_K_40_49, ra_M_15_29, ra_M_30_39, re_K_15_29, re_K_30_39, re_M_15_29, re_M_30_39, re_M_40_49, ru_K_15_29, ru_K_30_39, ru_K_40_49, ru_K_50 oraz ru_M_50. Konstrukcja nazwy zmiennej składa się z trzech elementów: nazwy wskaźnika (ra - współczynnik aktywności ekonomicznej, re - wskaźnik zatrudnienia, ru - stopa bezrobocia), kategorii płci (K - kobiety, M - mężczyźni) oraz grupy wieku (15-29, 30-39, 40-49 i 50+).

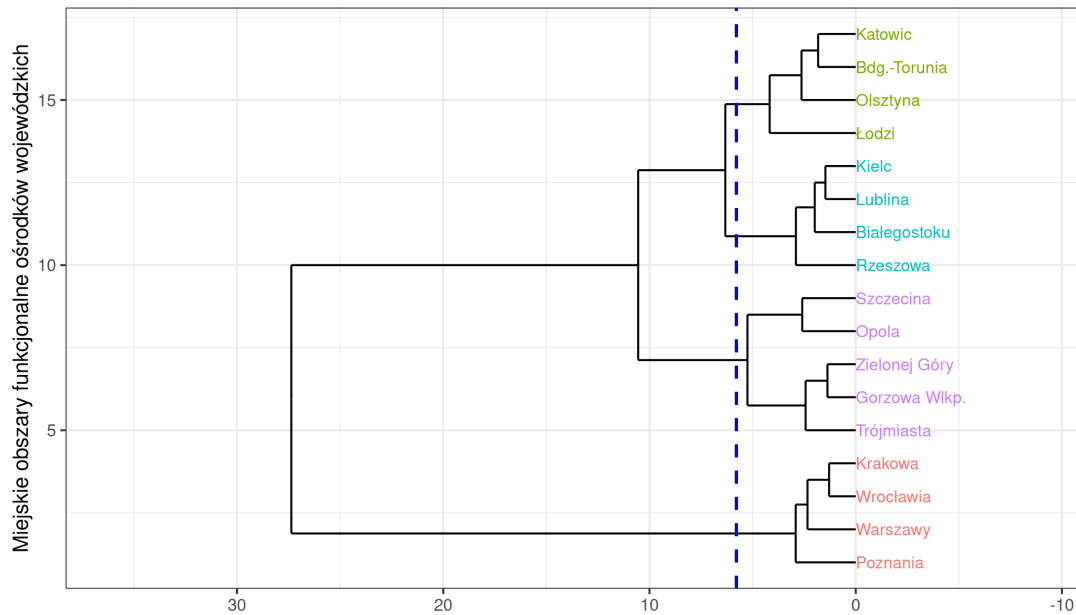
okalających ośrodek miejski. Niższa w rankingu pozycja miejskiego obszaru funkcjonalnego dla Łodzi może wynikać z ogólnie gorszej sytuacji panującej na rynku pracy w samej Łodzi w porównaniu ze wspomnianymi wcześniej czterema dużymi miastami. Łódź i jej obszar funkcjonalny charakteryzuje się bowiem wyższą stopą bezrobocia, niższym współczynnikiem aktywności ekonomicznej oraz wskaźnikiem zatrudnienia niż pozostałe największe miasta w Polsce (powyżej 500 tys. mieszkańców). Analizując ranking warto zwrócić uwagę na to, że najgorszą sytuację na rynku pracy zaobserwowano w miejskich obszarach funkcjonalnych, których rdzeń stanowią miasta zlokalizowane we wschodniej części Polski: Rzeszów, Kielce, Białystok oraz Lublin. Jest to pochodną na ogół gorszej sytuacji na rynku pracy w województwach Polski wschodniej: podlaskim, lubelskim, podkarpackim oraz świętokrzyskim.



Rysunek 5.13: Ranking miejskich obszarów funkcjonalnych ośrodków wojewódzkich – metoda TOPSIS

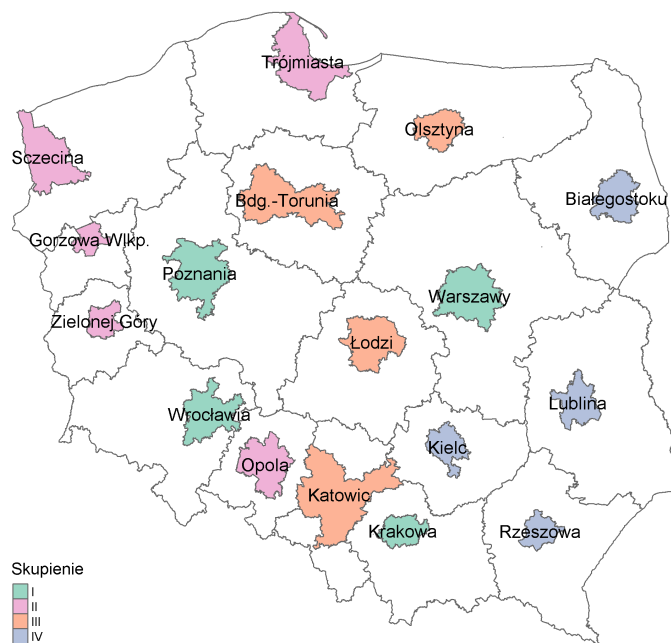
Źródło: Opracowanie własne.

Celem uzupełnienia opisu sytuacji panującej na rynku pracy w miejskich obszarach funkcjonalnych miast wojewódzkich pogrupowano je według ich podobieństwa. W badaniu użyto metodę Warda. Korzystając z kryterium Mojeny (1977) wyodrębniono cztery skupienia miejskich obszarów funkcjonalnych ośrodków wojewódzkich podobnych pod względem sytuacji panującej na rynku pracy (por. rysunek 5.14). Ich przestrzenne rozmieszczenie przedstawiono na rysunku 5.15.



Rysunek 5.14: Grupowanie miejskich obszarów funkcjonalnych ośrodków wojewódzkich metodą Warda ze względu na sytuację panującą na rynku pracy

Źródło: Opracowanie własne.



Rysunek 5.15: Grupowanie miejskich obszarów funkcjonalnych ośrodków wojewódzkich metodą Warda – ujęcie przestrzenne

Źródło: Opracowanie własne.

Można zauważyć, że wyniki grupowania uzyskane z wykorzystaniem metody Warda są w dużym stopniu zbieżne z rezultatami jakie otrzymano za pomocą metody Topsis. Pierwsze skupienie obejmuje miejskie obszary funkcjonalne ośrodków wojewódzkich Poznania, Warszawy, Wrocławia oraz Krakowa, w których sytuacja na rynku pracy była zdecydowanie najlepsza.

W skład drugiego skupienia wchodziły miejskie obszary funkcjonalne ośrodków wojewódzkich Szczecina, Opola, Zielonej Góry, Gorzowa Wielkopolskiego oraz Trójmiasta. W sporządzonym wcześniej rankingu dotyczącym sytuacji na rynku pracy uplasowały one się z reguły zaraz po miejskich obszarach funkcjonalnych skupionych wokół czterech dużych miast w Polsce. Oceniając przestrzenną lokalizację miejskich obszarów funkcjonalnych ośrodków wojewódzkich wchodzących w skład drugiego skupienia można zauważyć, że znajdują się one w północno-zachodniej części Polski. W skład trzeciego skupienia wchodziły MOF OW Katowic, Bydgoszczy i Torunia, Olsztyna oraz Łodzi. Również i te miejskie obszary funkcjonalne ośrodków wojewódzkich znajdowały się na zbliżonych pozycjach w przedstawionym powyżej rankingu. Położone są one w centralnej części Polski obejmującej pas od południa na północ. Ostatnia - czwarta grupa miejskich obszarów funkcjonalnych miast wojewódzkich obejmowała obszary skupione wokół Kielc, Lublina, Białegostoku i Rzeszowa, czyli te, które zlokalizowane są zlokalizowane we wschodniej części kraju. W rankingu charakteryzującym sytuację na rynku pracy znajdowały się one na ostatnich miejscach.

Podsumowując analizę zróżnicowania rynku pracy w miejskich obszarach funkcjonalnych ośrodków wojewódzkich z wykorzystaniem wybranych technik analizy skupień można zauważyć, że sytuacja na terenie miejskich obszarów funkcjonalnych miast powyżej 500 tys. mieszkańców z wyjątkiem Łodzi (skupienie I) jest lepsza niż w pozostałych tego rodzaju jednostkach administracyjnych. Sytuacja ta jest również korzystniejsza w MOF OW usytuowanych w północnej i zachodniej części Polski (skupienie II). Z kolei w miejskich obszarach funkcjonalnych zlokalizowanych wokół mniejszym miast wojewódzkich, przede wszystkim leżących w województwach wschodniej Polski, sytuacja na rynku pracy jest najmniej korzystna w porównaniu do pozostałych MOF OW (skupienie IV). Te obszary charakteryzują się na ogół wyższą stopą bezrobocia oraz niższymi wartościami wskaźnika zatrudnienia i współczynnika aktywności ekonomicznej. Taka sytuacja może być również związana z innymi czynnikami nieuwzględnionymi w powyższej analizie, np. niższymi

wynagrodzeniami w miastach wojewódzkich Polski wschodniej. W rankingu wszystkich miast wojewódzkich pod względem mediany rocznych dochodów mieszkańców MOF OW z IV skupienia plasują się w drugiej połowie jednostek (GUS, 2020a)⁵. Przeprowadzona powyżej analiza pozwoliła na pozytywne zweryfikowanie czwartej hipotezy badawczej, że rynek pracy w ujęciu miejskich obszarów funkcjonalnych ośrodków wojewódzkich cechuje się wyraźnym zróżnicowaniem przestrzennym.

5.4. Podsumowanie

Wyniki badania przedstawione w tym rozdziale umożliwiły weryfikację hipotez sformułowanych we wstępie niniejszej dysertacji. Biorąc pod uwagę estymator pośredni użyty do wyznaczenia wskaźników rynku pracy wykazano, że charakteryzuje się on lepszą precyzją niż estymator kalibracyjny. Własność ta została również potwierdzona w wyniku oceny precyzji szacunku przeprowadzonej w rozdziale czwartym. Pozytywnie zweryfikowano hipotezę badawczą zakładającą, że sytuacja na rynku pracy na obszarach funkcjonalnych miast wojewódzkich (MOF OW), mierzona współczynnikiem aktywności ekonomicznej ludności oraz wskaźnikiem zatrudnienia, jest na ogół lepsza niż w pozostałej części województwa czy podregionu (POJT). Jedynym wyjątkiem, w którym odnotowano odwrotną sytuację była grupa mężczyzn w wieku 15–29 lat. Wartości wskaźników dla kwartyli dla POJT były jednak tylko nieznacznie wyższe w porównaniu z wartościami dla MOF OW. Z kolei biorąc pod uwagę stopę bezrobocia, we wszystkich analizowanych przekrojach sytuacja na rynku pracy była lepsza w MOF OW niż POJT. Świadczą o tym wartości prezentowanych kwartyli. Należy jednak dodać, że w każdym przekroju istniały od jednego do ośmiu MOF OW charakteryzujących się wyższą stopą bezrobocia niż odpowiadający im obszar POJT. Warto jednak podkreślić, iż nie zawsze były to znaczne dysproporcje. W celu weryfikacji ostatniej hipotezy badawczej zakładającej, że rynek pracy w ujęciu miejskich obszarów funkcjonalnych ośrodków wojewódzkich cechuje się wyraźnym zróżnicowaniem przestrzennym, stworzono odpowiednie kartogramy umożliwiające analizę zróżnicowania przestrzennego współczynników charakteryzujących rynek pracy

⁵Wyniki zawarte w tej pracy eksperymentalnej zostały przygotowane przez Ośrodek Statystyki Miast Urzędu Statystycznego w Poznaniu i odnoszą się do 2018 r. Mogą być zatem dobrą zmienną „proxy” odzwierciedlającą sytuację na rynku pracy w stosunku do analiz prowadzonych w pracy doktorskiej, które dotyczą 2017 r.

w miejskich obszarach funkcjonalnych ośrodków wojewódzkich. Wykorzystano ponadto wybrane metody analizy taksonomicznej: Topsis na potrzeby utworzenia rankingu oraz Warda w celu wyodrębnienia jednorodnych grup MOF OW ze względu na sytuację na rynku pracy.

Zakończenie

Niniejsza praca jest próbą zaspokojenia popytu na dane o rynku pracy dla nietypowych przekrojów, jakimi niewątpliwie są miejskie obszary funkcjonalne. Zapotrzebowanie na tego rodzaju informacje zgłaszane jest przez różnych użytkowników danych statystycznych, szczególnie zaś przez instytucje szczebla rządowego i samorządowego zaangażowane w prowadzenie polityki społecznej. Dostęp do charakterystyk rynku pracy, zgodnych z definicjami przyjętymi przez Międzynarodową Organizację Pracy, na odpowiednim poziomie agregacji przestrzennej, stwarza również możliwość dokonywania analiz porównawczych zgodnych ze standardami międzynarodowymi. W związku z powyższym, po przeanalizowaniu dostępnych źródeł danych oraz technik estymacji, w ramach głównego celu rozprawy podjęto próbę pomiaru i oceny sytuacji na rynku pracy na obszarach funkcjonalnych miast wojewódzkich w 2017 roku. Analizę przeprowadzono na podstawie trzech wskaźników: współczynnika aktywności ekonomicznej ludności, wskaźnika zatrudnienia i stopy bezrobocia w przekrojach zdefiniowanych przez płeć i grupy wieku (15–29, 30–39, 40–49, 50 lat i więcej). Wskaźniki te zostały wyznaczone na podstawie wcześniej dokonanych oszacowań liczby osób pracujących, bezrobotnych oraz biernych zawodowo w przyjętych w pracy przekrojach. Przeprowadzone w ramach rozprawy badanie empiryczne pozwoliło na osiągnięcie zarówno celu głównego, jak i celu pobocznego. Wyniki analizy umożliwiły weryfikację założonych na wstępie hipotez badawczych.

Istotnym elementem pierwszego etapu prac była kwerenda dorobku naukowego w odniesieniu do zagadnienia miejskich obszarów funkcjonalnych, z uwzględnieniem sposobów ich

definiowania, zarządzania oraz delimitacji. Rozpoznano również ich umocowanie w ramach ustaw i strategii krajowych oraz ich znaczenie w kontekście wymiaru terytorialnego na poziomie Unii Europejskiej. W szerszym zakresie przedstawiono także Zintegrowane Inwestycje Terytorialne, będące jednym z instrumentów interwencji w ramach perspektywy finansowej Wspólnoty Europejskiej 2014–2020, który ma służyć przede wszystkim miastom i ich obszarom funkcjonalnym.

W pracy zidentyfikowano zakres dostępnej informacji w odniesieniu do obszarów funkcjonalnych miast oraz zgłaszane obecnie zapotrzebowanie informacyjne, szczególnie w odniesieniu do rynku pracy. Kierowano się głównie wskazaniem oraz rekomendacjami odbiorców danych, jakie znajdują się w dokumentach i raportach krajowych, takich jak: Założenia Krajowej Polityki Miejskiej do roku 2020 (MRR, 2012), Krajowa Polityka Miejska 2023 (MIiR, 2015a), Strategia na rzecz odpowiedzialnego rozwoju (Morawiecki, 2017), Krajowa Strategia Rozwoju Regionalnego 2030 (MIiR, 2019), Wskaźniki strategiczne na poziomie miejskich obszarów funkcjonalnych (Truskolaski i in., 2016) oraz w opracowaniach naukowych ((Batóg i in., 2015; Kaczmarek, 2015; Młodak, 2012; Rogalińska, 2017; Szafranek, 2017a; Szafranek i Kociuba, 2018). W analizie uwzględniono również mierniki dotyczące rynku pracy znajdujące się w opracowaniach poszczególnych strategii ZIT. Opierając się na przeprowadzonej kwerendzie potrzeb wybrano trzy pożądane wskaźniki z obszaru rynku pracy: współczynnik aktywności ekonomicznej, wskaźnik zatrudnienia oraz stopę bezrobocia, dla których dokonano kompleksowej oceny sytuacji na rynku pracy na poziomie miejskich obszarów funkcjonalnych miast wojewódzkich.

Następnym etapem realizacji pracy było sprawdzenie, czy w dostępnych źródłach danych (uzyskanych w badaniach pełnych i reprezentacyjnych prowadzonych przez statystykę publiczną oraz pozyskanych ze źródeł administracyjnych) znajdują się informacje o liczbie pracujących, bezrobotnych oraz biernych zawodowo (zgodnych z definicją przyjętą w BAEL), które są niezbędne do obliczenia wytypowanych wskaźników. Kwerenda źródeł danych wykazała, iż żadne z obecnie dostępnych źródeł nie zawiera pożądanej informacji dla miejskich obszarów funkcjonalnych niezależnie od przyjętej definicji. Przeprowadzona analiza porównawcza tych źródeł pozwoliła jednak na wskazanie potencjalnych źródeł zmiennych pomocniczych, które mogą być przydatne do estymacji pośredniej.

Mając to na uwadze, w kolejnym kroku podjęto decyzję o włączeniu do analizy jednej z metod statystyki małych obszarów odwołującej się do estymacji pośredniej, która

umożliwia połączenie informacji z kilku źródeł w ramach podejścia modelowego. Takie podejście pozwoliło na oszacowanie charakterystyk dotyczących rynku pracy wskazanych w ramach celu badawczego rozprawy. Na tym etapie szczególną uwagę zwrócono na to, aby wybrana technika estymacji zapewniała spójność z oszacowaniami bezpośrednimi na wyższym poziomie agregacji przestrzennej. Przeanalizowano możliwe podejścia pod kątem jakości otrzymanych szacunków pośrednich, a także zalecenia i wskazówki co do sposobu przeprowadzenia procesu estymacji pośredniej. W szerszym zakresie przeanalizowano estymatory typu SPREE, które charakteryzują się wspomnianą pożądaną własnością spójności szacunków (tzw. benchmarking). W analizie uwzględniono również przegląd literatury przedmiotu pod kątem implementacji tej grupy estymatorów. Ostatecznie zdecydowano się wykorzystać w części empirycznej dysertacji estymatory typu SPREE wykorzystujące modele GLSM oraz GLSMM, a opisane w pracy Zhanga i Chambersa (2004).

Kluczowym elementem w odniesieniu do całości wykonanych prac było zaproponowanie autorskiej procedury badawczej⁶, która dzieli na poszczególne etapy przebieg prac teoretycznych i empirycznych przeprowadzonych w ramach niniejszej rozprawy. Podziału dokonano na podstawie wskazówek i rekomendacji wypracowanych w ramach międzynarodowego projektu ESSnet on SAE (2012) oraz zaproponowanych w artykule Tzavidisa i współautorów (2018) etapów postępowania w procesie estymacji, a także z uwzględnieniem przedstawionych w tych opracowaniach schematów postępowania w przypadku użycia metod estymacji pośredniej. W toku prac wypracowano autorską propozycję procedury prowadzenia badania, która prowadzi do otrzymania oszacowań wskaźników w obszarze rynku pracy na niepublikowanym poziomie. Warto podkreślić, że schemat ten można stosunkowo łatwo przekształcić i dostosować do innego obszaru badawczego, w ramach którego wskazane jest użycie metod estymacji pośredniej (ubóstwo, niepełnosprawność czy statystyka przedsiębiorstw). Zgodnie z przyjętą procedurą badawczą, określono szczegółowy zakres źródeł danych i zmiennych, jakie miały być wykorzystane w części empirycznej pracy oraz dokonano wyboru estymatorów oraz metod oceny jakości uzyskanych szacunków. Określono przekroje i układy tabel. To pozwoliło doprecyzować zakres rozważanych miejskich obszarów funkcjonalnych. Na potrzeby pracy analizę zawężono do miejskich obszarów funkcjonalnych ośrodków wojewódzkich (MOF OW) wyznaczonych do realizacji instrumentu Zintegrowanych Inwestycji Terytorialnych w latach 2014–2020. Przyjęty

⁶Procedura ta została przedstawiona na rysunku 4.1.

układ tabel umożliwił uzyskanie oszacowań dla pozostałej części obszarów otaczających poszczególne MOF OW. Należy podkreślić, iż do tej pory informacje w takich przekrojach nie były dostępne, a zaproponowane w rozprawie przekroje tabel nigdy nie były publikowane przez statystykę publiczną. Zaproponowana procedura badawcza jest nowatorskim podejściem w estymacji typu SPREE, w którym aktualizacji poddaje się z reguły tabele otrzymane na podstawie wyników Narodowych Spisów Powszechnych. (por. National Skills Commission (2021); Luna i in. (2014); Correa-Onel i in. (2016); Qiao (2005); Hidiroglou i Patak (2009)).

Przeprowadzony zgodnie z zaproponowaną procedurą badawczą proces estymacji pozwolił na podstawie wszystkich rozpatrywanych estymatorów uzyskać oszacowania charakterystyk rynku pracy w zakładanych przekrojach. Analiza ocen precyzji uzyskanych dla poszczególnych wariantów estymatorów potwierdziła hipotezę badawczą zakładającą, że estymatory typu SPREE charakteryzują się lepszą precyzją szacunków, niezależnie od analizowanego przekroju w porównaniu z wykorzystywanym w badaniu BAEL estymatorem kalibracyjnym. Na podstawie wszystkich rozważanych metod oceny jakości estymatorów, z uwzględnieniem przesłanek statystyczno-merytorycznych, dokonano wyboru jednego estymatora typu SPREE - estymatora GLSM w wariancie $GLSM_{3P,L}$. Został on wykorzystany w empirycznej części pracy.

Otrzymane oszacowania mierników dotyczące aktywności zawodowej, zatrudnienia jak i bezrobocia w rozważanych układach tabel, umożliwiły kompleksową ocenę sytuacji na rynku pracy w badanych przekrojach, w szczególności na obszarach funkcjonalnych miast wojewódzkich w odniesieniu do sformułowanych hipotez badawczych. Dzięki uzyskanej informacji wskazano domeny wyróżnione w przekroju rzeczowo-przestrzennym, dla których sytuacja na rynku pracy wymaga szczególnego monitorowania. I tak na przykład w odniesieniu do stopy bezrobocia domenę taką stanowiły osoby w wieku 15–29 lat. Z kolei biorąc pod uwagę współczynnik aktywności ekonomicznej oraz wskaźnik zatrudnienia sytuacja mężczyzn w wieku poniżej 30 lat jest gorsza na obszarach MOF OW aniżeli w POJT. Niepokojąco niskie poziomy tych dwóch wskaźników zanotowano w grupie kobiet w wieku 50+, chociaż ich sytuacja jest lepsza na obszarach MOF OW w porównaniu do POJT.

Uzyskane z wykorzystaniem wybranych metod taksonomicznych wyniki pozwoliły ponadto utworzyć ranking poszczególnych MOF OW oraz pogrupować je ze względu na

sytuację na rynku pracy. Umożliwiło to przestrzenną analizę zróżnicowania tego zjawiska w układzie terytorialnym, dla którego obecnie nie są publikowane żadne informacje.

Należy podkreślić, że zaproponowane w rozprawie podejście metodyczne umożliwia opracowanie rozważanych tabel, z częstością większą niż roczna, np. w cyklu kwartalnym, który jest zgodny z dostępnością danych z badania BAEL. Jest ono także odporne na korekty granic obszarów funkcjonalnych miast wojewódzkich polegające na włączaniu lub wyłączeniu poszczególnych gmin.

Przedstawione wyniki prac nie wyczerpują wyzwań badawczych, które warto byłoby podjąć w kolejnych etapach. Wśród kluczowych zagadnień wartych zainteresowania, można wskazać na kilka zasadniczych wątków, które prawdopodobnie wyznaczać będą kierunki dalszych badań w omawianym obszarze:

- bazując na wypracowanej metodologii, należałoby cyklicznie prowadzić monitoring sytuacji na rynku pracy w przekroju rozważanych w pracy domen. Umożliwiłoby to diagnozę zachodzących zmian i przygotowanie działań prewencyjnych w przypadku pojawienia się niepożądanych zjawisk na rynku pracy;
- należałoby poddać głębszej analizie merytorycznej uzyskane oszacowania wskaźników rynku pracy. Szczególną uwagę warto by zwrócić na osoby w grupie wieku 15–29 lat oraz w wieku 50 lat i więcej, czy też bardziej ogólnie na współczynnik aktywności ekonomicznej kobiet w wybranych grupach wieku oraz jego powiązanie z rządowym programem 500 plus;
- warto również rozważyć, wykorzystując zaprezentowane źródła danych, przygotowanie i oszacowanie innych układów tabel dla wybranych kategorii rynku pracy, które uwzględniają dodatkowe zmienne warstwujące takie jak: grupy sekcji PKD, poziom wykształcenia, czy czas pozostawania bez pracy;
- wykorzystując estymację typu SPREE można przeanalizować osobno miasta wojewódzkie, gminy tworzące ich obszar funkcjonalny, pozostałą część województwa lub podregionu;
- odnosząc się do wyzwań metodologicznych, przeprowadzenie procesu estymacji pośredniej dla kolejnych lat umożliwiłoby zweryfikowanie stabilności oszacowań estymatorów w czasie;
- należałoby dokonać kalibracji wag dla podregionów zielonogórskiego oraz gorzowskiego, aby dostosować odpowiadające im struktury do danych demograficznych;

- należałoby zweryfikować jakość estymacji pośredniej poprzez przeprowadzenie analizy obciążenia uwzględniając podejścia oparte zarówno na schemacie losowania próby (*design-based*), jak i w ujęciu modelowym (*model-based*);
- z praktycznego punktu widzenia, najlepszą oceną zaproponowanej metodologii byłoby porównanie wartości kategorii rynku pracy i wskaźników obliczonych na podstawie danych NSP 2021 z wynikami uzyskanymi za pomocą podejścia przedstawionego w pracy;
- w kontekście prac rozważanych w dysertacji można rozważyć inne klasy estymatorów typu SPREE zaproponowane w literaturze przedmiotu (Luna Hernandez, 2016).

Podsumowując przeprowadzone w pracy rozważania, należy przypuszczać, że estymatory typu SPREE będą odgrywać w polskiej praktyce badań statystycznych coraz większą rolę. Wynika to z kilku zasadniczych powodów. Po pierwsze w systemie statystyki publicznej w Polsce coraz większą wagę zaczyna się przywiązywać do pozyskiwania informacji z pozastatystycznych źródeł danych (głównie rejestrów administracyjnych). Po drugie, z uwagi na światową pandemię COVID-19 wywołaną wirusem SARS-CoV-2, potrzebne będą kompleksowe opracowania sytuacji panującej na rynku pracy, który jest wyznacznikiem zmian zachodzących zarówno w gospodarce, jak i w życiu społecznym. Wreszcie, rosnący popyt na informacje na niskich poziomach agregacji przestrzennej oraz dla bardziej szczegółowo zdefiniowanych domen będzie wymuszać na instytucjach przeprowadzających badanie stosowanie bardziej wyrafinowanych metod. Estymacja typu SPREE w ujęciu zaprezentowanym w pracy ma szansę sprostać temu zapotrzebowaniu.

Bibliografia

- Australian Bureau of Statistics (2006). A Guide to Small Area Estimation. Version 1.1. Internal ABS document. Australian Bureau of Statistics.
- Barca, F. (2009). An Agenda for a Reformed Cohesion Policy. A Place-Based Approach to Meeting European Union Challenges and Expectations. European Communities, Brussels.
- Batóg, J., Mojsiewicz, M., Batóg, B. (2015). Monitorowanie strategii rozwoju w świetle delimitacji miejskich obszarów funkcjonalnych. *Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania Uniwersytetu Szczecińskiego*, 42(1):11–30.
- Beręsewicz, M. Szymkowiak, M. (2015). Big data w statystyce publicznej – nadzieje, osiągnięcia, wyzwania i zagrożenia. *Ekonometria*, 2:9–22.
- Berg, E. J. Fuller, W. A. (2014). Small area prediction of proportions with applications to the Canadian Labour Force Survey. *Journal of Survey Statistics and Methodology*, 2(3):227–256.
- Berry, B. J. L. (1967). Functional Economic Areas and Consolidated Urban Regions of the US. Social Sciences Research Council, New York.
- Bieszk-Stolorz, B. (2017). The impact of gender on routes for registered unemployment exit in Poland. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 12(4):733–749.

- Bieszk-Stolorz, B. Dmytrów, K. (2019). Spatial diversity of effectiveness of forms of professional activation in Poland in years 2008 - 2014 by poviats. *Oeconomia Copernicana*, 10(1):113–130.
- Binek, J., Opravil, Z., Chmelař, R., Svobodova, H. (2016). Cooperation and Mutual Relationships of Cities and their Hinterlands with Regard to the Operation of EU Integrated Development Instruments. *Quaestiones Geographicae*, 35(2):59–69.
- Birch, M. W. (1963). Maximum likelihood in three-way contingency tables. *Journal of the Royal Statistical Society: Series B*, 25(1):220–233.
- Broniszewski, M. (2017). Znaczenie zintegrowanego podejścia do obszarów funkcjonalnych w ujęciu regionalnym — studium przypadku. *Wrocławskie Studia Politologiczne*, 23:161–172.
- Brown, G., Chambers, R., Heady, P., Heasman, D. (2001). Evaluation of small area estimation methods—an application to unemployment estimates from the UK LFS. w: *Proceedings of Statistics Canada Symposium*.
- Bul, R. Walaszek, M. (2015). Dojazdy do pracy i szkół jako podstawowe kryterium delimitacji obszarów funkcjonalnych miast. *Rozwój Regionalny i Polityka Regionalna*, 29:119–138.
- Carmone Jr, F. J., Kara, A., Maxwell, S. (1999). Hinov: A new model to improve market segment definition by identifying noisy variables. *Journal of Marketing Research*, 36(4):501–509.
- Chłoń-Domińczak, A. Ptak-Chmielewska, A. (2020). Wykorzystanie danych rejestrowych do oceny sytuacji ekonomicznej mikro, małych i średnich przedsiębiorstw na przykładzie województwa kujawsko-pomorskiego. *Wiadomości Statystyczne. The Polish Statistician*, 65(6):52–68.
- Churski, P. (2018a). Podejście zorientowane terytorialnie (place based-policy) – teoria i praktyka polityki regionalnej. *Rozwój Regionalny i Polityka Regionalna*, 41:31–50.
- Churski, P. (2018b). Rola wielkich miast w rozwoju społeczno-gospodarczym Polski – poznańskie refleksje. w: J. Szlachta, J. Woźniak, (red.), *Rola wielkich miast w rozwoju społeczno-gospodarczym Polski*, *Studia KPZK*, 186:129–143.

- Cicha-Nazarczuk, M. Nazarczuk, J. M. (2017). Oddziaływanie specjalnych stref ekonomicznych na sytuację na rynku pracy. *Studia Prawno-Ekonomiczne*, CII:173–185.
- Correa-Onel, S., Whitworth, A., Piller, K. (2016). Assessing the Generalised Structure Preserving Estimator (GSPREE) for Local Authority Population Estimates by Ethnic Group in England. Office for National Statistics.
- Daas, P., Ossen, S., Vis-Visschers, R. J. W. M., Arends-Tóth, J. (2009). Checklist for the quality evaluation of administrative data sources. *Statistics Netherlands Discussion Paper*, 9042.
- Dehnel, G. (2003). *Statystyka małych obszarów jako narzędzie oceny rozwoju ekonomicznego regionów*. Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań.
- Dehnel, G. (2015). Rejestr podatkowy oraz rejestr ZUS jako źródło informacji dodatkowej dla statystyki gospodarczej – możliwości i ograniczenia. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 384:51–59.
- Dehnel, G. Kowalewski, J. (2015). Źródła administracyjne a krótkookresowa statystyka przedsiębiorstw – wyzwania i perspektywy. *Wiadomości Statystyczne. The Polish Statistician*, LX(11):16–29.
- Dehnel, G. Wawrowski, Ł. (2020). Robust estimation of wages in small enterprises: The application to Poland's districts. *Statistics in Transition*, 21(1):137–157.
- Deming, W. E. Stephan, F. F. (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics*, 11(4):427–444.
- Dolecka, M. Raczkiewicz, D. (2014). Wielkość bezrobocia w Polsce według spisu ludności w porównaniu z pozostałymi danymi statystyki publicznej. *Annales Universitatis Mariae Curie-Skłodowska, Sectio H-Oeconomia*, XLVIII(1):79–89.
- Dostał, L., Gabler, S., Ganninger, M., Münnich, R. (2016). Frame Correction Modelling with Applications to the German Register-Assisted Census 2011. *Scandinavian Journal of Statistics*, 43(3):904–920.
- Dziewoński, K. (1987). Strefa podmiejska - próba ujęcia teoretycznego. *Przegląd Geograficzny*, 59(1-2):55–63.

- European Council (2007). Leipzig Charter on Sustainable European Cities. European Union.
- European Council (2011). Territorial Agenda of the European Union. Towards an Inclusive, Smart and Sustainable Europe of Diverse Regions. Agreed at the Informal Ministerial Meeting of Ministers responsible for Spatial Planning and Territorial Development on 19th May 2011. European Union.
- Eurostat (2012). ESSnet on Small Area Estimation. Report on Workpackage 6 Guidelines. Guidelines for the application of the small area estimation methods in NSI sample surveys. Eurostat.
- Eurostat (2017). Methodological Manual on City Statistics. 2017 Edition. Eurostat, Luxembourg.
- Eurostat (2019). Methodological manual on territorial typologies. 2018 Edition. Eurostat, Luxembourg.
- Filas-Przybył, S. Klimanek, T. (2019). Employment-related commuting to provincial capital cities and their functional areas (FUAs) in Poland. COSTNET19 Conference, Bilbao, October, 9–11.
- Friedmann, J. Miller, J. (1965). The urban field. *Journal of the American Institute of Planners*, 31(4):312–320.
- Fuller, W. A. (1999). Environmental surveys over time. *Journal of Agricultural, Biological and Environmental Statistics*, 4(4):331–345.
- Gierańczyk, W. Ryczkowski, M. (2018). Public Statistics Resources as a Source for Research in Geography of Industry. *Studies of the Industrial Geography Commission of the Polish Geographical Society*, 32(4):240–251.
- Głuszczyk, D. Raszowski, A. (2016). Stopa bezrobocia-mankamenty pomiaru i próba ich eliminacji. *Ekonomia XXI Wieku*, 2(10):22–29.
- Gołata, E. (2010). Estymacja pośrednia aktywności ekonomicznej na potrzeby spisu opartego na rejestrach. *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Poznaniu*, 149:85–104.

- Gołata, E. (2018). *Koniec ery tradycyjnych spisów ludności*. Wydawnictwo Uniwersytetu Ekonomicznego we Poznaniu, Poznań.
- Gołata, E. Dehnel, G. (2012). Rejestry administracyjne w analizie przedsiębiorczości. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 242/2012:202–211.
- Gołata, E. Dehnel, G. (2013). Rozbieżności szacunków NSP 2011 i BAEL. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 278/2013:120–130.
- Gosh, M. (2020). Small area estimation: its evolution in five decades. *Statistics in Transition*, 21(4):1–22.
- Groß, M., Rendtel, U., Schmid, T., Schmon, S., Tzavidis, N. (2017). Estimating the density of ethnic minorities and aged people in Berlin: Multivariate kernel density estimation applied to sensitive geo-referenced administrative data protected via measurement error. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 180(1):161–183.
- GUS (2012). Raport z wyników. Narodowy Spis Powszechny Ludności i Mieszkań 2011. Główny Urząd Statystyczny, Warszawa.
- GUS (2013). Aktywność ekonomiczna ludności Polski. Narodowy Spis Powszechny Ludności i Mieszkań 2011. Główny Urząd Statystyczny, Warszawa.
- GUS (2014). Gospodarstwa domowe i rodziny. Charakterystyka demograficzna. Narodowy Spis Powszechny Ludności i Mieszkań 2011. Główny Urząd Statystyczny, Warszawa.
- GUS (2018a). Aktywność ekonomiczna ludności Polski, IV kwartał 2017 r. Główny Urząd Statystyczny, Warszawa.
- GUS (2018b). Bezrobocie rejestrowane, I-IV kwartał 2017 r. Główny Urząd Statystyczny, Warszawa.
- GUS (2018c). Pracujący w gospodarce narodowej w 2017 r. Główny Urząd Statystyczny, Warszawa.
- GUS (2018d). Zeszyt metodologiczny. Badanie Aktywności Ekonomicznej Ludności. Główny Urząd Statystyczny, Warszawa.

- GUS (2018e). Zeszyt metodologiczny. Statystyka rynku pracy i wynagrodzeń. Główny Urząd Statystyczny, Warszawa.
- GUS (2019a). Praca nierejestrowana w Polsce w 2017 r. Główny Urząd Statystyczny, Warszawa.
- GUS (2019b). Przepływy ludności związane zatrudnieniem w 2016 r. Główny Urząd Statystyczny, Warszawa.
- GUS (2019c). Wybrane aspekty rynku pracy w 2018 r. Główny Urząd Statystyczny, Warszawa.
- GUS (2019d). Zeszyt metodologiczny. Badania przedsiębiorstw niefinansowych. Główny Urząd Statystyczny, Warszawa.
- GUS (2019e). Zeszyt metodologiczny. Miesięczny meldunek o działalności gospodarczej przedsiębiorstw. Główny Urząd Statystyczny, Warszawa.
- GUS (2020a). Stratyfikacja dochodowa mieszkańców miast. Główny Urząd Statystyczny, Warszawa.
- GUS (2020b). Zeszyt metodologiczny. Pracujący w gospodarce narodowej. Główny Urząd Statystyczny, Warszawa.
- GUS (2020c). Zeszyt metodologiczny. Pracujący w gospodarce narodowej według źródeł administracyjnych. Główny Urząd Statystyczny, Bydgoszcz.
- GUS (2020d). Zeszyt metodologiczny. Wynagrodzenia w gospodarce narodowej. Warszawa, Główny Urząd Statystyczny.
- GUS (2021). Instrukcja organizacyjna do Narodowego Spisu Powszechnego Ludności i Mieszkań w 2021 . Centralne Biuro SPisowe, Główny Urząd Statystyczny.
- Hay, D. Hall, P. (1977). *Urban Regionalisation of Great Britain*. Department of Geography.
- Heffner, K. Gibas, P. (2013). Obszary funkcjonalne ośrodków regionalnych w Polsce. w: A. Nowakowska (red.), *Zrozumieć terytorium. Idea i praktyka*. Wydawnictwo Uniwersytetu Łódzkiego, Łódź, s:217–240.

- Heffner, K. Klemens, B. (2016). Wpływ polityki miejskiej i miejskich obszarów funkcjonalnych na rozwój obszarów wiejskich. Ocena problemu na podstawie badań literaturowych. *Studia KPZK*.
- Herbst, M. Wójcik, P. (2013). Delimitacja dyfuzji rozwoju z miast metropolitalnych z wykorzystaniem korelacji przestrzennej. *Studia Regionalne i Lokalne*, 4(54):5–21.
- Hidiroglou, M. Patak, Z. (2009). An application of small area estimation techniques to the Canadian Labour Force Survey. w: *Proceedings of the Survey Methods Section, Statistical Society of Canada*.
- Hidiroglou, M. A., Beaumont, J.-F., Yung, W. (2019). Development of a small area estimation system at Statistics Canada. *Survey Methodology*, 45(1):101–126.
- Huff, D. L. (1973). The delineation of a national system of planning regions on the basis of urban spheres of influence. *Regional Studies*, 7(3):323–329.
- Ireland, C. T. Kullback, S. (1968). Contingency tables with given marginals. *Biometrika*, 55(1):179–188.
- Isidro, M., Haslett, S., Jones, G. (2016). Extended structure preserving estimation (ESPREE) for updating small area estimates of poverty. *Annals of Applied Statistics*, 10(1):451–476.
- Isidro, M. C. (2010). *Intercensal Updating of Small Area Estimates, unpublished PhD thesis*. PhD thesis, Massey University, New Zealand.
- Janukowicz, P. (2010). Bezrobocie rejestrowane a bezrobocie według BAEL. *Polityka społeczna*, 37(1):18–20.
- Józefowski, T. Młodak, A. (2009). Observation of flows of population in polish statistics—problems and challenges. *Assistance for the Younger Generation. Statistics and Planning in Big Agglomerations*, Institut für Angewandte Demographie IFAD, Berlin, s:61–76.
- Józefowski, T. Młodak, A. (2017). Special economic zones in the context of regional development. *Europa XXI*, 33:33–47.

- Józefowski, T. Rynarzewska-Pietrzak, B. (2010). Ocena możliwości wykorzystania rejestru PESEL w spisie ludności. w: Gołata, E., redaktor, *Pomiar i informacja w gospodarce. Zeszyt Naukowy WIGE*. Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.
- Józefowski, T. Szymkowiak, M. (2013). Zastosowanie estymatora typu SPREE w szacowaniu liczby osób bezrobotnych w przekroju podregionów. *Studia Oeconomica Posnaniensia*, 1(10):120–136.
- Józefowski, T. Szymkowiak, M. (2015). Wykorzystanie uogólnionej miary odległości do porządkowania liniowego powiatów województwa podkarpackiego w świetle funkcjonowania specjalnej strefy ekonomicznej Euro-Park Mielec. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 384/2015:135–143.
- Kachniarz, M. Babczuk, A. (2015). Ekspertyza „Ocena podziału terytorialnego państwa z uwzględnieniem efektywności funkcjonowania urzędów organów jednostek samorządu terytorialnego – wnioski i rekomendacje”. Narodowy Instytut Samorządu Terytorialnego.
- Kaczmarek, T. (2015). Functional Urban Areas as the Focus of Development Policy in Poland. *Rozwój Regionalny i Polityka Regionalna*, 29:9–19.
- Kaczmarek, T. (2017). Implementacja Zintegrowanych Inwestycji Terytorialnych (ZIT) w miejskich obszarach funkcjonalnych. Przykład metropolii Poznań. *Rozwój Regionalny i Polityka Regionalna*, 40:9–24.
- Kaczmarek, T. Bul, R. (2014). Statystyka obszarów metropolitalnych – stan i potrzeby. *Biuletyn KPZK*.
- Klimanek, T. (2017). Estymacja pośrednia niepełnosprawności – badanie symulacyjne. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 469:91–99.
- Kłosowski, F., Pytel, S., Runge, A., Sitek, S., Zuzańska-Żyśko, E. (2013). Rynek pracy w subregionie centralnym województwa śląskiego. Podręcznik ”dobrych praktyk”. Uniwersytet Śląski, Sosnowiec.
- Kociuba, D. (2015). Miejskie obszary funkcjonalne - wyzwania planistyczne. *Studia Miejskie*, 18:39–53.

- Kociuba, D. (2017a). Delimitacja miejskich obszarów funkcjonalnych ośrodków wojewódzkich w realizacji zintegrowanych inwestycji terytorialnych w Polsce - teoria versus praktyka. *Studia Regionalne i Lokalne*, 69(3):54–78.
- Kociuba, D. (2017b). Zintegrowane Inwestycje Terytorialne jako narzędzie zarządzania terytorialnego w miejskich obszarach funkcjonalnych. *Studia KPZK*, 174:144–153.
- Kociuba, D. (2018a). Implementation of Integrated Territorial Investments in Poland - Rationale, Results, and Recommendations. *Quaestiones Geographicae*, 4:81–98.
- Kociuba, D. (2018b). Teoria i praktyka wsparcia zrównoważonego rozwoju miejskich obszarów funkcjonalnych w Polsce. *Biuletyn KPZK PAN*, 272:316–327.
- Koebe, T., Arias-Salazar, A., Rojas-Perilla, N., Schmid, T. (2021). Intercensal updating using structure-preserving methods and satellite imagery. *arXiv preprint arXiv:2103.03834*.
- Korcelli, P. (1981). Regiony miejskie w systemie osadniczym Polski. w: K. Dziewoński, P. Korcelli (red.), *Studia nad migracjami i przemianami systemu osadniczego w Polsce, Prace Geograficzne IGiPZ PAN*, 140:189–212.
- Korcelli-Olejniczak, E. (2012). *Region metropolitalny – pojęcie, struktura przestrzenna, dynamika (vol. 235)*. IGiPZ PAN.
- Korzeniewski, J. (2012). *Metody selekcji zmiennych w analizie skupień. Nowe Procedury*. Wydawnictwo Uniwersytetu Łódzkiego.
- Kuć-Czajkowska, K. (2019). *Obszary funkcjonalne miast wojewódzkich w Polsce. Przestrzeń współpracy i konkurencji samorządów terytorialnych*. Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej, Lublin.
- Kuźnik, F. (2015). Miejskie obszary funkcjonalne a polityka miejska. *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, 250:7–24.
- Kwiatkowski, E. Kwiatkowska, E. (2020). Zróżnicowanie poziomu i charakteru bezrobocia w przekroju powiatów w Polsce. *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie*, 3(987):7–29.

- Luna, A., Zhang, L.-C., Whitworth, A., Piller, K. (2014). Small area estimates of the population distribution by ethnic group in England: A proposal using structure preserving estimators. *Statistics in Transition new series and Survey Methodology Joint Issue: Small Area Estimation 2014*, 16(4):585–602.
- Luna Hernandez, A. (2016). *Multivariate Structure Preserving Estimation for Population Compositions*. Thesis for the degree of doctor of philosophy, University of Southampton.
- Madras-Kobus, B. Rogowski, J. (2013). Analiza wyników badania stóp bezrobocia według wykształcenia w województwie podlaskim. *Optimum. Studia ekonomiczne*, 6(66):98–115.
- Marchetti, S., Giusti, C., Pratesi, M. (2016). The use of Twitter data to improve small area estimates of households' share of food consumption expenditure in Italy. *ASTA Wirtschafts- und Sozialstatistisches Archiv*, 10(2):79–93.
- Marchetti, S., Giusti, C., Pratesi, M., Salvati, N., Giannotti, F., Pedreschi, D., Rinzivillo, S., Pappalardo, L., Gabrielli, L. (2015). Small area model-based estimators using big data sources. *Journal of Official Statistics*, 31(2):263–281.
- Marker, D. A. (1999). Organization of Small Area Estimators Using a Generalized Linear Regression Framework. *Journal of Official Statistics*, 15(1):1–24.
- Markowicz, I. (2015). Statystyczna analiza przestrzennego zróżnicowania bezrobocia w Polsce. *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, 223:247–256.
- Markowicz, I. (2017). Labour market in the zachodniopomorskie voivodeship on the Polish background. *Studia i Prace WNEiZ US*, 50(1):191–204.
- Markowski, T. Marszał, T. (2006). *Metropolie. Obszary metropolitalne. Metropolizacja: Problemy i pojęcia podstawowe*. Komitet Przestrzennego Zagospodarowania Kraju PAN, Warszawa.
- McKenzie, R. D. (1933). *The metropolitan community*. McGraw-Hill, New York.
- MIiR (2015a). *Krajowa Polityka Miejska 2023*. Ministerstwo Infrastruktury i Rozwoju, Warszawa.

- MIiR (2015b). Strategie ponadregionalne–wymiar terytorialny polityki rozwoju. Ministerstwo Infrastruktury i Rozwoju, Warszawa.
- MIiR (2017). Programowanie perspektywy finansowej 2014–2020. Umowa Partnerstwa. Ministerstwo Infrastruktury i Rozwoju, Warszawa.
- MIiR (2019). Krajowa Strategia Rozwoju Regionalnego 2030. Ministerstwo Inwestycji i Rozwoju.
- Młodak, A. (2006). Polish experiences and possibilities in realisation of the URBAN AUDIT programme. w: *In 25th SCORUS Conference on Regional and Urban Statistics and Research*, „Globalization Impact on Regional and Urban Statistics”, volume 30, Wrocław.
- Młodak, A. (2012). Statystyka metropolii polskich - problemy i perspektywy. *Studia Regionalne i Lokalne*, 2(48):20–38.
- Młodak, A. (2014). Zastosowanie metody bootstrapowej do oceny precyzji szacunków danych z badania reprezentacyjnego NSP 2011 na przykładzie dojazdów do pracy. *Wiadomości Statystyczne. The Polish Statistician*, 59(12):38–57.
- Mojena, R. (1977). Hierarchical grouping methods and stopping rules: an evaluation. *The Computer Journal*, 20(4):359–363.
- Morawiecki, M. (2017). Strategia na rzecz Odpowiedzialnego Rozwoju do roku 2020 (z perspektywą do 2030 r.). Rada Ministrów, Warszawa.
- MRR (2010). Regiony, miasta, obszary wiejskie. Ministerstwo Rozwoju Regionalnego, Warszawa.
- MRR (2012). Założenia krajowej polityki miejskiej do roku 2020. Ministerstwo Rozwoju Regionalnego, Warszawa.
- National Skills Commission (2021). Small Area Labour Markets. Australian Government.
- Nelder, J. A. Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society A*, 135(3):370–384.
- Noble, A., Haslett, S., Arnold, G. (2002). Small Area Estimation via Generalized Linear Models. *Journal of Official Statistics*, 18(1):45–60.

- Noworól, A. (2014). Functional urban area as the city of the future. *Czasopismo Techniczne, Architektura*:143–157.
- Noworól, A. (2016). Rekomendacje dotyczące zarządzania strategicznego w samorządzie terytorialnym. Fundacja Rozwoju Demokracji Lokalnej, Warszawa.
- Paradysz, J. (2009). Ocena dobroci estymacji dla małych obszarów. *Zeszyty Naukowe/Uniwersytet Ekonomiczny w Poznaniu*, 128:19–28.
- Pfeffermann, D. (2013). New Important Developments in Small Area Estimation. *Statistical Science*, 28(1):40–68.
- Piórek, A. (2019). Administracyjne źródła danych w badaniach statystycznych rynku pracy w Polsce. Referat wygłoszony na konferencji 230 lat Statystyki Publicznej w Polsce, Łódź, Marzec, 18–19.
- Popiński, W. (2006). Development of the Polish Labour Force Survey. *Statistics in Transition*, 7(5):1009–1030.
- Potrykowska, A. (1989). Funkcjonalne regiony miejskie w krajowym systemie osadniczym. *Prace Geograficzne IGiPZ PAN*, 152:55–75.
- Purcell, N. J. Kish, L. (1980). Postcensal Estimates for Local Areas (Or Domains). *International Statistical Review / Revue Internationale de Statistique*, 48(1):3–18.
- Qiao, C. (2005). Combining Administrative and Survey Data to Derive Small-area Estimates Using Loglinear Modelling. *Labour*, 19(4):767–800.
- Rada Ministrów (2012). Uchwała nr 239 Rady Ministrów z dnia 13 grudnia 2011 r. w sprawie przyjęcia koncepcji Przestrzennego Zagospodarowania Kraju 2030 (M.P. 2012, poz. 252). Rada Ministrów, Warszawa.
- Rada Ministrów (2016). Rozporządzenie Rady Ministrów z dnia 28 lipca 2016 r. w sprawie programu badań statystycznych statystyki publicznej na rok 2017. Dz.U. 2016 poz. 1426, Warszawa.
- Rao, J. M. K. Molina, I. (2015). *Small Area Estimation*. John Wiley and Sons, New Jersey, 2nd edition.

- Rogalińska, D. (2017). Monitorowanie polityki na poziomie regionalnym przez statystykę publiczną. *Biuletyn KPZK PAN*, 268:71–82.
- Roszka, W. (2012). System statystyki publicznej oparty na zintegrowanych źródłach danych. *Przegląd Statystyczny*, 59(numer specjalny 2):205–221.
- Roszka, W. (2013). *Statystyczna integracja danych w badaniach społeczno - ekonomicznych*. PhD thesis, Poznań University of Economics and Business.
- Ryczkowski, M. Stopiński, P. (2018). Labour Market Areas in Poland. *Przegląd Statystyczny*, LXV(3):350–368.
- Särndal, C. E. Lundström, S. (2005). *Estimation in surveys with nonresponse*. John Wiley & Sons.
- Schaible, W. L., redaktor (2013). *Indirect estimators in US federal programs*, volume 108. Springer Science & Business Media.
- Schmid, T., Bruckschen, F., Salvati, N., Zbiranski, T. (2017). Constructing socio-demographic indicators for National Statistical Institutes using mobile phone data: estimating literacy rates in Senegal. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180(4):1163–1190.
- Serafin, P. (2018). Zmiany w typologii demograficznej miejskich obszarów funkcjonalnych Polski w latach 2002–2017. *Biuletyn KPZK PAN*, 272:328–343.
- Sierak, J. Bitner, M. (2016). Planowanie finansowe i inwestycyjne w zakresie realizacji usług publicznych. *Realizacja usług publicznych w jednostkach samorządu terytorialnego – ograniczenia, możliwości, rekomendacje*.
- Śleszyński, P. (2013a). Delimitacja Miejskich Obszarów Funkcjonalnych stolic województw. *Przegląd Geograficzny*, 85(2):173–197.
- Śleszyński, P. (2013b). Kryteria delimitacji Miejskich Obszarów Funkcjonalnych Ośrodków Wojewódzkich. Ministerstwo Rozwoju Regionalnego, Warszawa.
- Śleszyński, P. (2015). Problemy delimitacji miejskich obszarów funkcjonalnych w Polsce. *Rozwój Regionalny i Polityka Regionalna*, 29:37–53.

- Śleszyński, P. (2018). *Polska średnich miast. Założenia i koncepcja deglomeracji w Polsce*. Klub Jagielloński, Warszawa.
- Śleszyński, P., Bański, J., Degórski, M., Komornicki, T. (2017). *Delimitacja obszarów strategicznej interwencji państwa: obszarów wzrostu i obszarów problemowych*. IGiPZ PAN, Warszawa.
- Śliwicki, D. i in. (2018). *Opracowanie metodologii i oszacowanie liczby pracujących w gospodarce narodowej według miejsca zamieszkania i głównego miejsca pracy na poziomie NTS 4, stopy bezrobocia rejestrowanego na poziomie NTS 5 oraz miar wynagrodzeń brutto na poziomie NTS 4*. Centrum Badań i Edukacji Statystycznej GUS, Warszawa.
- Słodowa-Hełpa, M. (2014). Dokąd zmierzamy - odnowa czy od nowa? Nadzieje, obawy i dylematy związane z instytucjonalnym stymulowaniem zintegrowanego rozwoju w Polsce. *Nierówności społeczne a wzrost gospodarczy*, 37(1/2014):20–42.
- Słupińska, M. (2013). Rozwój ukierunkowany terytorialnie w nowej polityce spójności. *Acta Universitatis Lodzianis, Folia Oeconomica*, 289:9–18.
- Stephan, F. F. (1942). An iterative method of adjusting sample frequency tables when expected marginal totals are known. *The Annals of Mathematical Statistics*, 13(2):166–178.
- Sudra, P. (2018). Ewolucja kryteriów delimitacji wielkomiejskich układów osadniczych w Polsce. *Przegląd Geograficzny*, 90(2):181–208.
- Szafranek, E. (2015). Zintegrowane Inwestycje Terytorialne jako narzędzie wspierające kształtowanie centrów rozwoju regionalnego. *Rozwój Regionalny i Polityka Regionalna*, 32:81–94.
- Szafranek, E. (2017a). Idea a praktyka wdrażania zintegrowanych inwestycji terytorialnych w Polsce. *Studia KPZK*, 174:135–143.
- Szafranek, E. (2017b). Miejskie obszary funkcjonalne a kształtowanie spójności terytorialnej. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 467:113–132.
- Szafranek, E., Kociuba, D. (2018). Development of urban areas in the conditions of territorial-oriented policy—theoretical assumptions and experience in functional areas of Polish cities. *Studia Ekonomiczne*, 361:53–75.

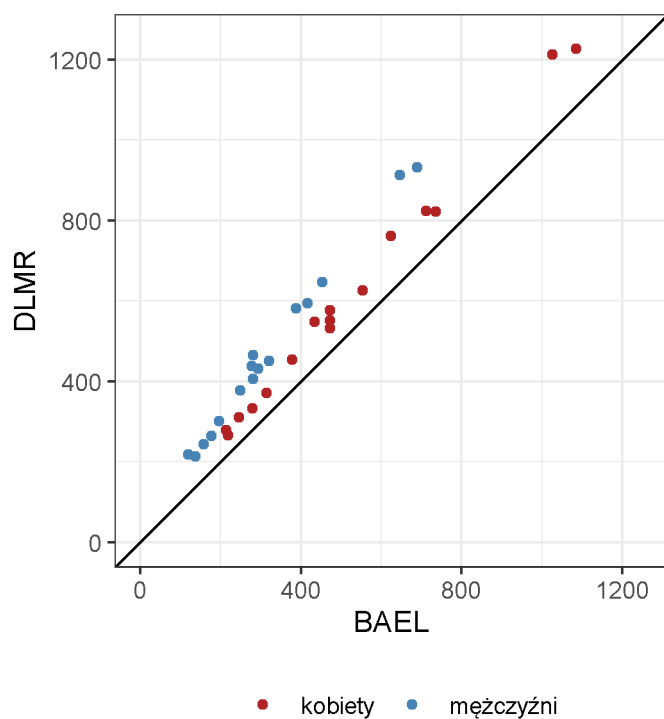
- Szlachta, J. (2017). Obszary funkcjonalne w Polsce – perspektywa europejska. *Studia KPZK*.
- Szreder, M. (2015). Big data wyzwaniem dla człowieka i statystyki. *Wiadomości Statystyczne. The Polish Statistician*, LX(8):1–11.
- Szreder, M. (2017). Nowe źródła informacji i ich wykorzystywanie w podejmowaniu decyzji. *Wiadomości Statystyczne. The Polish Statistician*, LXII(7):5–17.
- Szreder, M. (2019). Istotność statystyczna w czasach big data. *Wiadomości Statystyczne. The Polish Statistician*, 64(11):42–57.
- Szymkowiak, M. (2014). Estymatory kalibracyjne stosowane w Narodowym Spisie Powszechnym w 2011 roku. *Wiadomości Statystyczne. The Polish Statistician*, LIX(11):1–16.
- Szymkowiak, M., Józefowski, T. (2014). Konstrukcja i praktyczne wykorzystanie estymatorów typu SPREE na przykładzie dwuwymiarowych tabel kontyngencji. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu. Taksonomia*, 327:195–201.
- Szymkowiak, M., Józefowski, T., Klimanek, T. (2017a). Zaprojektowanie makiet tablic wynikowych dotyczących niepełnosprawności w przekroju domen o większej szczegółowości niż publikowane na podstawie bieżących badań reprezentacyjnych. Urząd Statystyczny w Poznaniu, Poznań.
- Szymkowiak, M., Młodak, A., Wawrowski, Ł. (2017b). Mapping poverty at the level of subregions in Poland using indirect estimation. *Statistics in Transition*, 18(4):609–635.
- Truskolaski, T., Busłowska, A., Kurza, A., Mierzyńska, D., Perło, D. (2016). Monitorowanie strategiczne miejskich obszarów funkcjonalnych realizujących Zintegrowane Inwestycje Terytorialne w Polsce. Raport 3. Wskaźniki strategiczne na poziomie miejskich obszarów funkcjonalnych. Stowarzyszenie Białostockiego Obszaru Funkcjonalnego, Białystok.
- Truskolaski, T., Busłowska, A., Waligóra, K. (2017). Formation of Functional Urban Areas: The Case of Białystok Functional Area. *Optimum. Studia Ekonomiczne*, 89(5):175–186.

- Tzavidis, N., Zhang, L. C., Luna, A., Schmid, T., Rojas-Perilla, N. (2018). From start to finish: a framework for the production of small area official statistics. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 181(4):927–979.
- Upton, G. Cook, I. (2008). *Oxford Dictionary of Statistics*. Oxford University Press, 2nd rev. edition.
- Wagner, J., Münnich, R., Hill, J., Stoffels, J., Udelhoven, T. (2017). Non-parametric small area models using shape-constrained penalized B-splines. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180(4):1089–1109.
- Wałachowski, K. Król, S. (2019). *Uciekające metropolie. Ranking 100 polskich miast*. Klub Jagielloński, Kraków.
- Walesiak, M. (2008). Procedura analizy skupień z wykorzystaniem programu komputerowego clustersim i środowiska R [Cluster analysis procedure with clustersim computer programme and R environment], *Taksonomia 15. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, 7(1207):44–56.
- Walesiak, M. Gatnar, E. (2009). *Statystyczna analiza danych z wykorzystaniem programu R*. Wydawnictwo Naukowe PWN.
- Wallgren, A. Wallgren, B. (2014). *Register-based Statistics: Statistical Methods for Administrative Data*. John Wiley & Sons, Second edition.
- Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244.
- Wesołowski, J. Tarczyński, J. (2016). Podstawy matematyczne technik imputacyjnych. *Wiadomości Statystyczne. The Polish Statistician*, LXI(9):7–54.
- Wilak, K. (2017). *Estymacja bezrobocia pozornego i biernego z wykorzystaniem strukturalnych modeli szeregów czasowych*. PhD thesis, Poznań University of Economics and Business.
- Wilak, K. (2020). Ocena skali bezrobocia biernego w Polsce. *Wiadomości Statystyczne*, 65(6):11–38.

- Witkowski, J. (2017). Badanie Aktywności Ekonomicznej Ludności jako podstawa modernizacji statystyki pracy. *Wiadomości Statystyczne*, 12(12):7–22.
- Wojciechowski, E., Podgórnika-Krzykacz, A., Dolewka, Z., Wojciechowski, M. (2014). Samorząd terytorialny w Polsce. Uniwersytet Łódzki, Łódź.
- Woźniak, J. (2015). Wybrane aspekty zintegrowanego planowania rozwoju – perspektywa regionalna. *Studia KPZK*.
- Wu, C. Lu, W. (2016). Calibration weighting methods for complex surveys. *International Statistical Review*, 84(1):79–98.
- Yoon, K. P. Hwang, C.-L. (1981). *Multiple attribute decision-making: Methods and applications*. Springer.
- Zaucha, J., Brodzicki, T., Ciołek, D., Komornicki, T., Mogiła, Z., Szlachta, J., Zaleski, J. (2015). *Terytorialny wymiar wzrostu i rozwoju*. Difin, Warszawa.
- Zgierska, A. (2017). 25 lat Badania Aktywności Ekonomicznej Ludności w Polsce — kamienie milowe i perspektywy rozwoju. *Wiadomości Statystyczne. The Polish Statistician*, 12(12).
- Zhang, L. C. (2009). Estimates for small area compositions subjected to informative missing data. *Survey Methodology*, 35(2):191–201.
- Zhang, L. C. Chambers, R. L. (2004). Small area estimates for cross-classifications. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 66(2):479–496.
- ZIT (2015a). Strategia Zintegrowanych Inwestycji Terytorialnych Miejskiego Obszaru Funkcjonalnego Gorzowa Wlkp. Gorzów Wlkp.
- ZIT (2015b). Strategia Zintegrowanych Inwesycji Terytorialnych dla Krakowskiego Obszaru Funkcjonalnego. Stowarzyszenie Metropolia Krakowska, Kraków.
- ZIT (2016a). Strategia Miejskiego Obszaru Funkcjonalnego Olsztyna. Geoprofit, Warszawa-Olsztyn.
- ZIT (2016b). Strategia Zintegrowanych Inwestycji Terytorialnych dla Bydgosko-Toruńskiego Obszaru Funkcjonalnego. Bydgoszcz.

- Zuzańska-Żyśko, E. (2012). Funkcje metropolitalne w Górnośląskim Obszarze Metropolitalnym. *Studia Regionalne i Lokalne*, 13(48):39–61.
- Zuzańska-Żyśko, E. (2017). Region metropolitalny Silesia w świetle koncepcji funkcjonalnych. *Studia KPZK PAN*, 174:116–125.
- Zwara, W. i in. (2014). *Opracowanie metodologii i oszacowanie danych dot. pracujących w gospodarce narodowej na poziomie NTS 4*. Centrum Badań i Edukacji Statystycznej GUS, Bydgoszcz.

Aneks



Rysunek A.1: Porównanie liczby osób biernych zawodowo w podziale na płeć na poziomie województw w 2017 roku

Źródło: Opracowanie własne.

Tabela A.1: Nazwy miejskich obszarów funkcjonalnych objętych instrumentem ZIT

Nazwa obszaru funkcjonalnego	Miasto/a centralne
MOF Subregionu Centralnego	Katowice-Gliwice
Warszawski OF	Warszawa
MOF Gdańsk-Gdynia-Sopot	Gdańsk-Gdynia-Sopot
Łódzki OM	Łódź
Krakowski OF	Kraków
Wrocławski OF	Wrocław
MOF Poznania	Poznań
Bydgosko-Toruński OF	Bydgoszcz-Toruń
Szczeciński OM	Szczecin
Lubelski OF	Lublin
Białostocki OF	Białystok
Rzeszowski OF	Rzeszów
Kielecki OF	Kielce
MOF Aglomeracji Opolskiej	Opole
MOF Olsztyna	Olsztyn
MOF Zielonej Góry	Zielona Góra
MOF Gorzowa Wielkopolskiego	Gorzów Wielkopolski
MOF Aglomeracji Wałbrzyskiej	Wałbrzych
MOF Aglomeracji Jeleniogórskiej	Jelenia Góra
Agglomeracji Kalisko-Ostrowskiej	Kalisz
MOF Elbląga	Elbląg
MOF Elku	Elk
MOF Subregionu Południowego	Bielsko-Biała
MOF Subregionu Północnego	Częstochowa
MOF Subregionu Zachodniego	Rybnik
Koszalińsko-Koło-brzesko-Białogardzki OF	Koszalin-Koło-brzeg-Białogard

Źródło: Opracowanie własne.

Tabela A.2: Oszacowania parametrów β rozpatrywanych wariantów estymatorów GLSM oraz GLSMM dla przekroju ogółem

Estymator	Parametr β
$GLSM_{2,W}$	0,959
$GLSMM_{2,W}$	0,907

Źródło: Opracowanie własne.

Tabela A.3: Oszacowania parametrów β rozpatrywanych wariantów estymatorów GLSM oraz GLSMM w przekroju płci

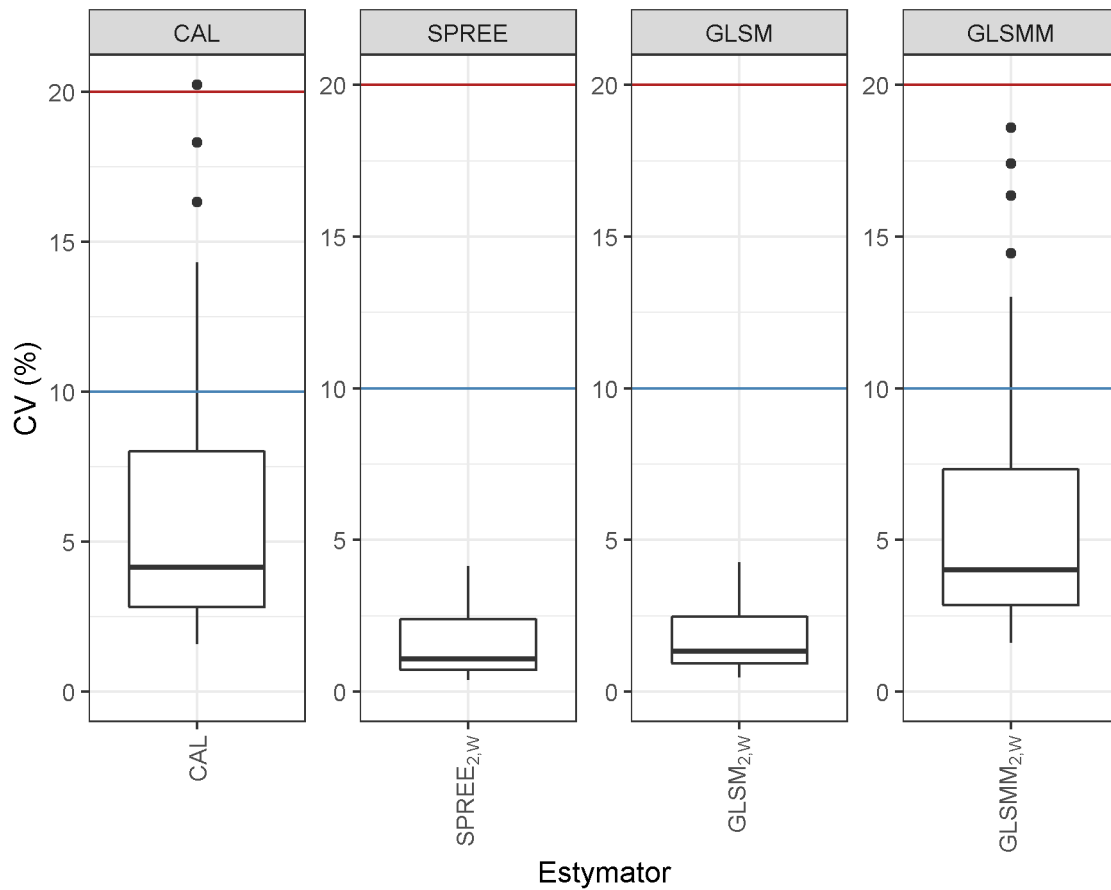
Estymator	Warstwa	Parametr β
$GLSM_{2,W}$	–	0,957
$GLSM_{2,L}$	–	1,047
$GLSM_{3P,L}$	–	0,958
$GLSMM_{2,W}$	–	0,824
$GLSMM_{2,L}$	–	0,967
$GLSMM_{3P,L}$	–	0,894
$GLSM_{3S,L}$	Kobiety	0,964
	Mężczyźni	0,952
$GLSMM_{3S,L}$	Kobiety	0,815
	Mężczyźni	0,872

Źródło: Opracowanie własne.

Tabela A.4: Oszacowania parametrów β rozpatrywanych wariantów estymatorów GLSM oraz GLSMM w przekroju grup wieku

Estymator	Warstwa	Parametr β
$GLSM_{2,W}$	–	0,951
$GLSM_{2,L}$	–	1,256
$GLSM_{3P,L}$	–	0,959
$GLSMM_{2,W}$	–	0,884
$GLSMM_{2,L}$	–	1,039
$GLSMM_{3P,L}$	–	0,878
$GLSM_{3S,L}$	15–29	0,853
	30–39	0,966
	40–49	0,914
	50+	1,007
$GLSMM_{3S,L}$	15–29	0,628
	30–39	0,974
	40–49	0,828
	50+	0,957

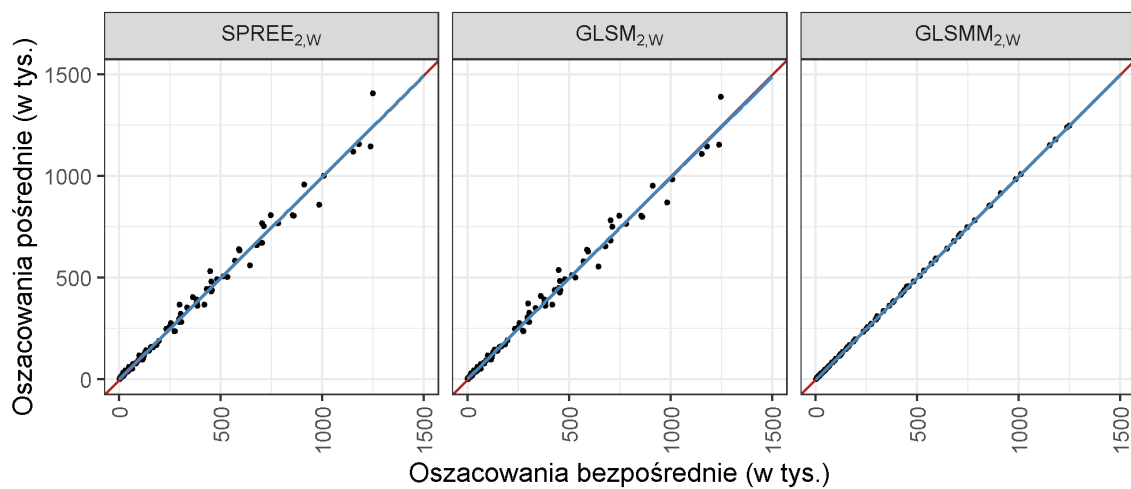
Źródło: Opracowanie własne.



Rysunek A.2: Rozkład wartości współczynnika precyzji dla kategorii rynku pracy w układzie tabel ogółem

Objaśnienia: Czerwoną linią zaznaczono wartość CV na poziomie 20%. Z kolei poziom CV = 10% wyznacza linia niebieska.

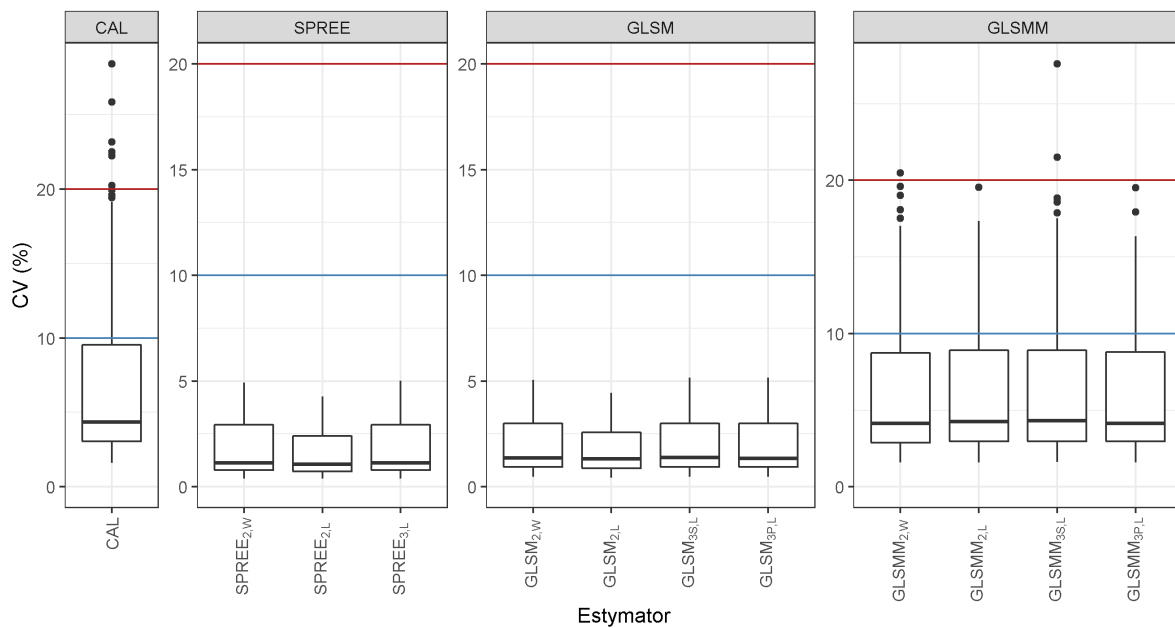
Źródło: Opracowanie własne.



Rysunek A.3: Porównanie szacunków pośrednich i kalibracyjnych w układzie tabel ogółem – $CV \leq 20\%$

Objaśnienia: Punkty układające się wzdłuż prostej oznaczonej kolorem czerwonym o równaniu $y = x$ świadczą o zgodności oszacowań wariantu estymatora pośredniego z kalibracyjnym. Natomiast niebieska linia jest prostą regresji dopasowaną do punktów na korelacyjnym diagramie rozrzutu.

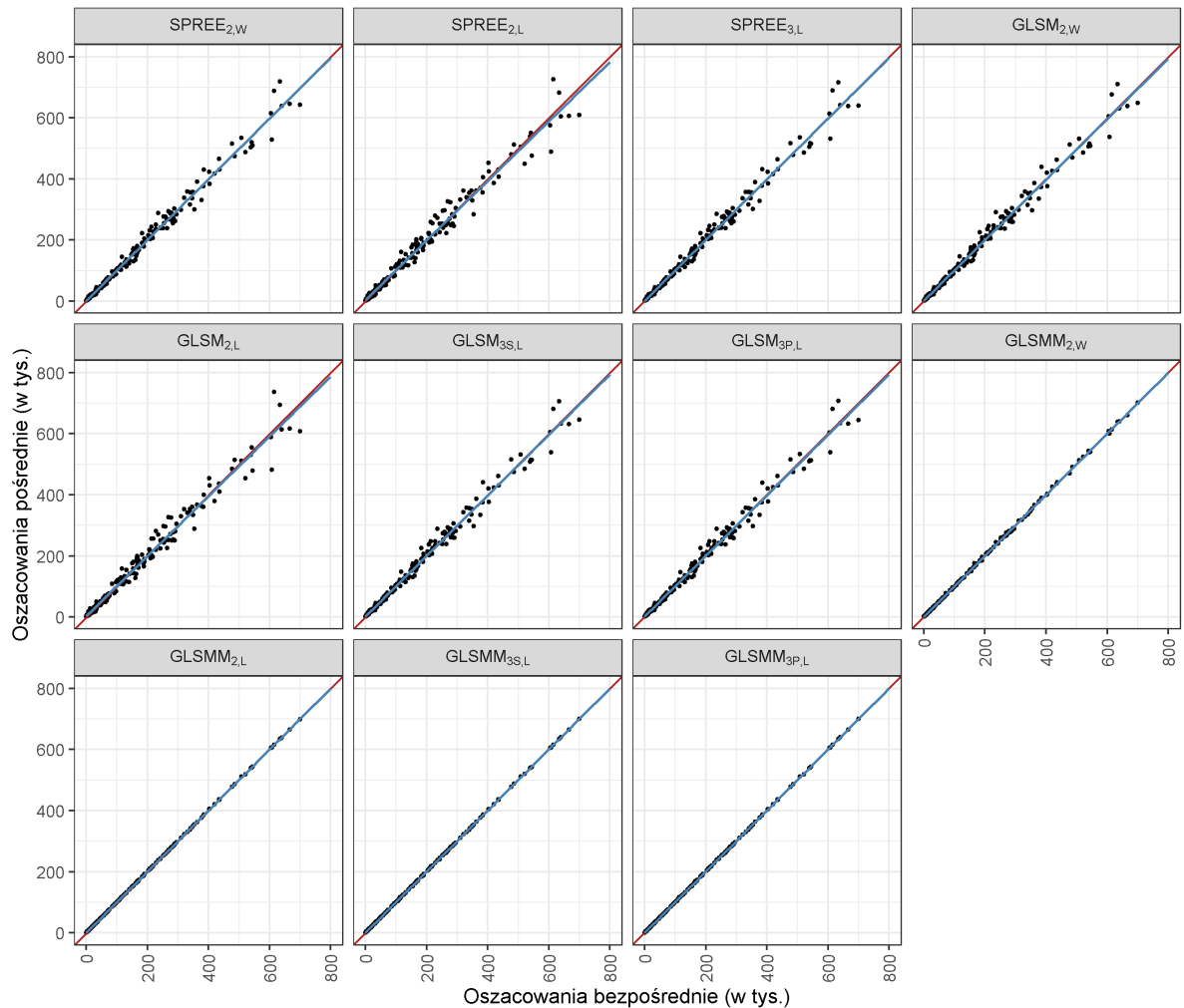
Źródło: Opracowanie własne.



Rysunek A.4: Rozkład wartości współczynnika precyzji dla kategorii rynku pracy w układzie tabel uwzględniających płeć

Objaśnienia: Czerwoną linią zaznaczono wartość CV na poziomie 20%. Z kolei poziom CV = 10% wyznacza linia niebieska.

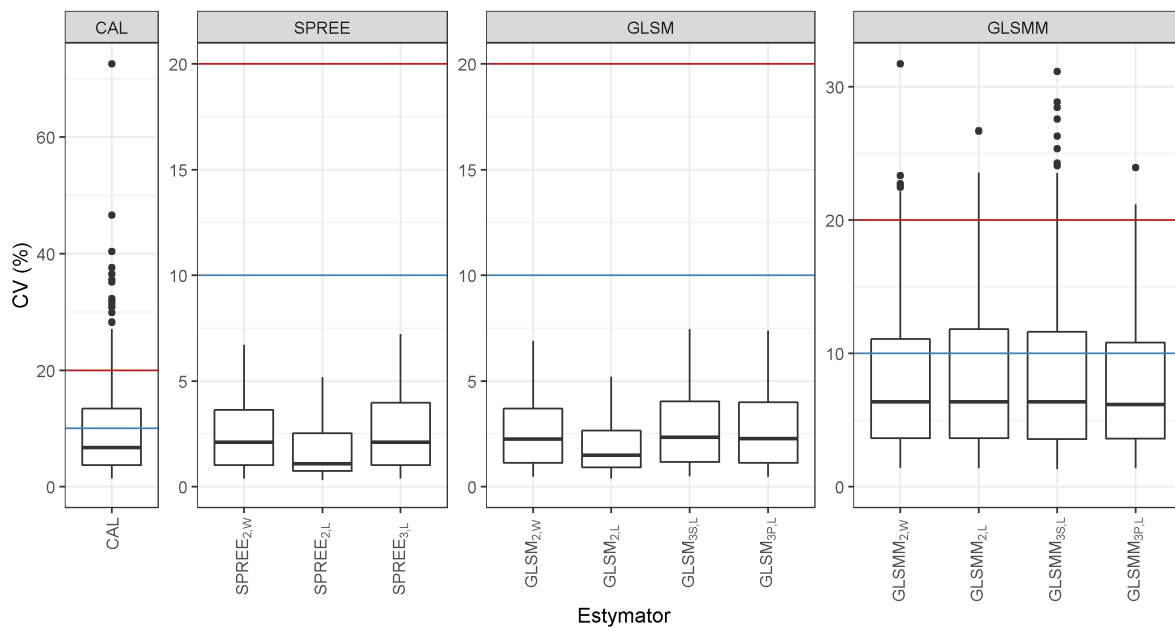
Źródło: Opracowanie własne.



Rysunek A.5: Porównanie szacunków pośrednich i kalibracyjnych w układzie tabel uwzględniających płeć – $CV \leq 20\%$

Objaśnienia: Punkty układające się wzdłuż prostej oznaczonej kolorem czerwonym o równaniu $y = x$ świadczą o zgodności oszacowań wariantu estymatora pośredniego z kalibracyjnym. Natomiast niebieska linia jest prostą regresji dopasowaną do punktów na korelacyjnym diagramie rozrzutu.

Źródło: Opracowanie własne.

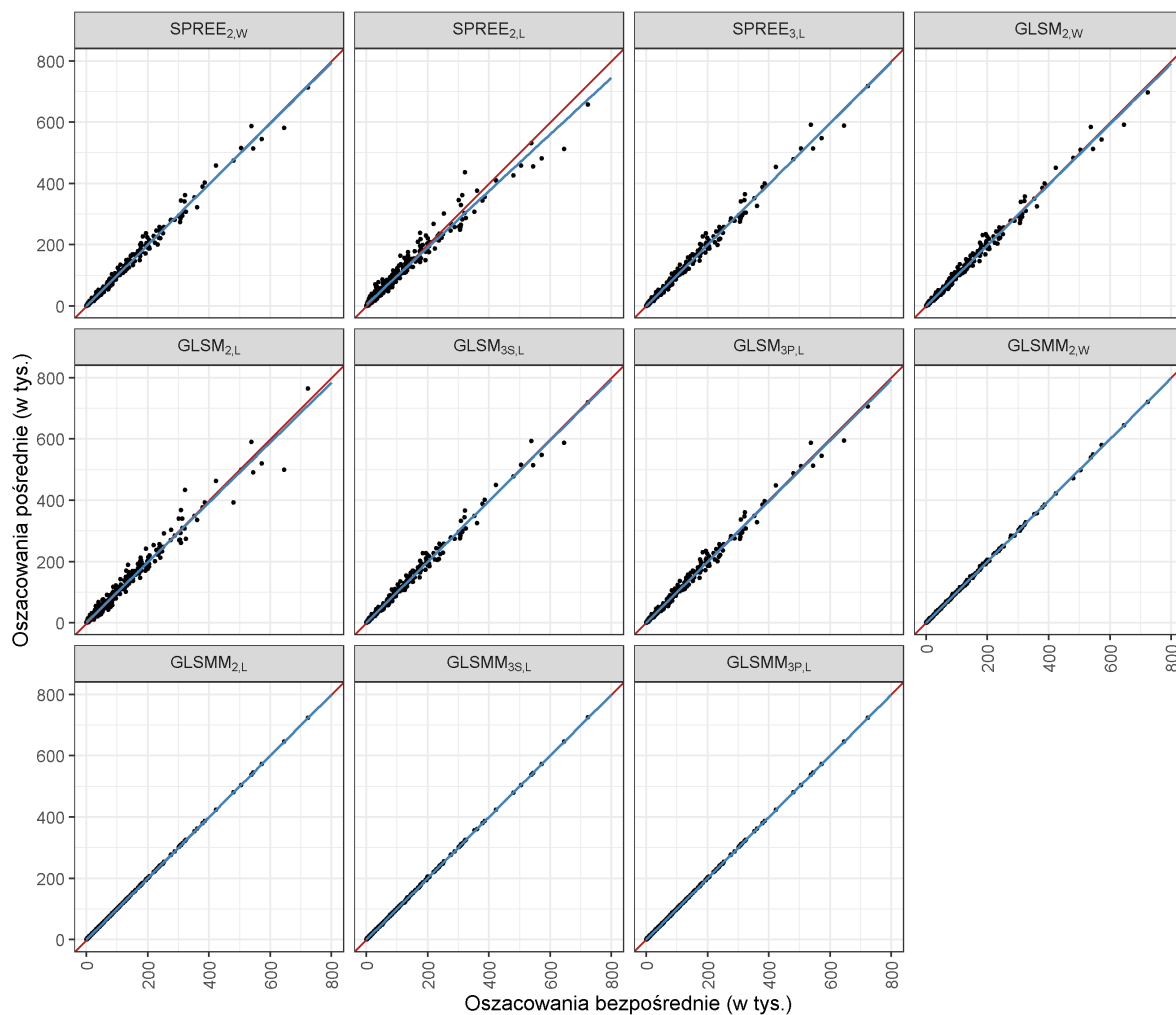


Rysunek A.6: Rozkład wartości współczynnika precyzji dla kategorii rynku pracy w układzie tabel uwzględniających grupy wieku

Uwaga: Przyjęte skale do oceny współczynnika CV są różne dla poszczególnych grup wariantów estymatorów.

Objaśnienia: Czerwoną linią zaznaczono wartość CV na poziomie 20%. Z kolei poziom CV = 10% wyznacza linia niebieska.

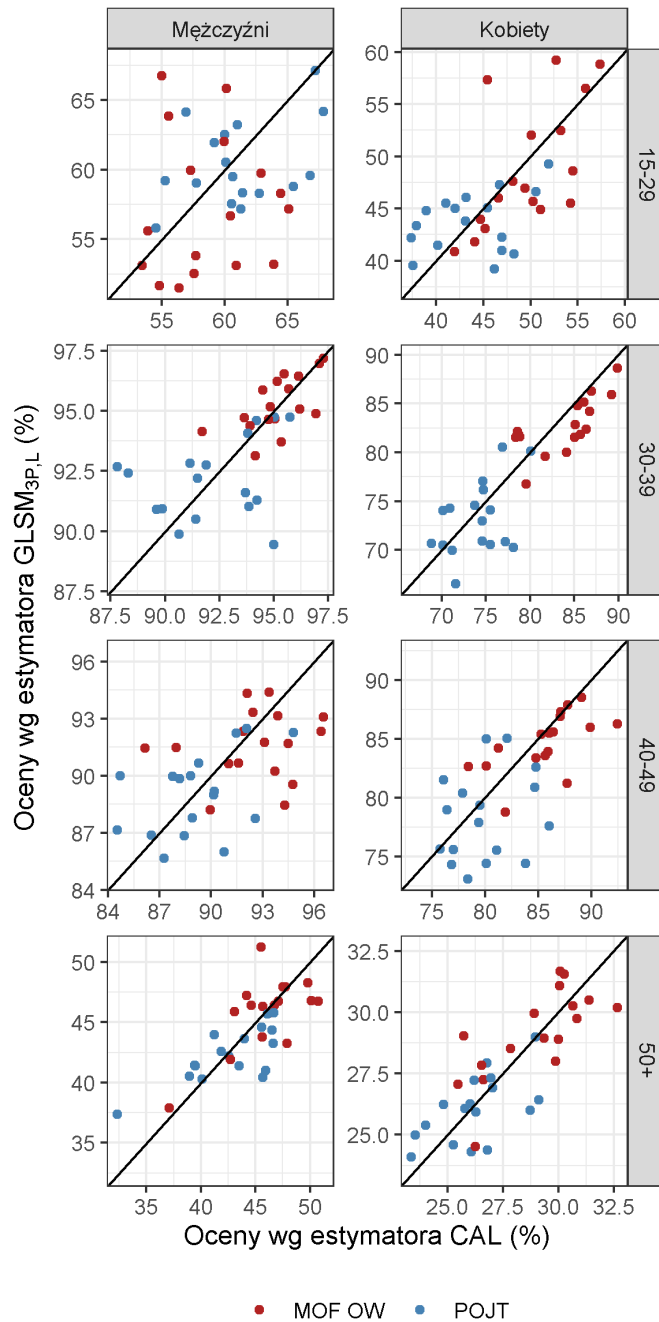
Źródło: Opracowanie własne.



Rysunek A.7: Porównanie szacunków pośrednich i kalibracyjnych w układzie tabel uwzględniających grupy wieku – $CV \leq 20\%$

Objaśnienia: Punkty układające się wzdłuż prostej oznaczonej kolorem czerwonym o równaniu $y = x$ świadczą o zgodności oszacowań wariantu estymatora pośredniego z kalibracyjnym. Natomiast niebieska linia jest prostą regresji dopasowaną do punktów na korelacyjnym diagramie rozrzutu.

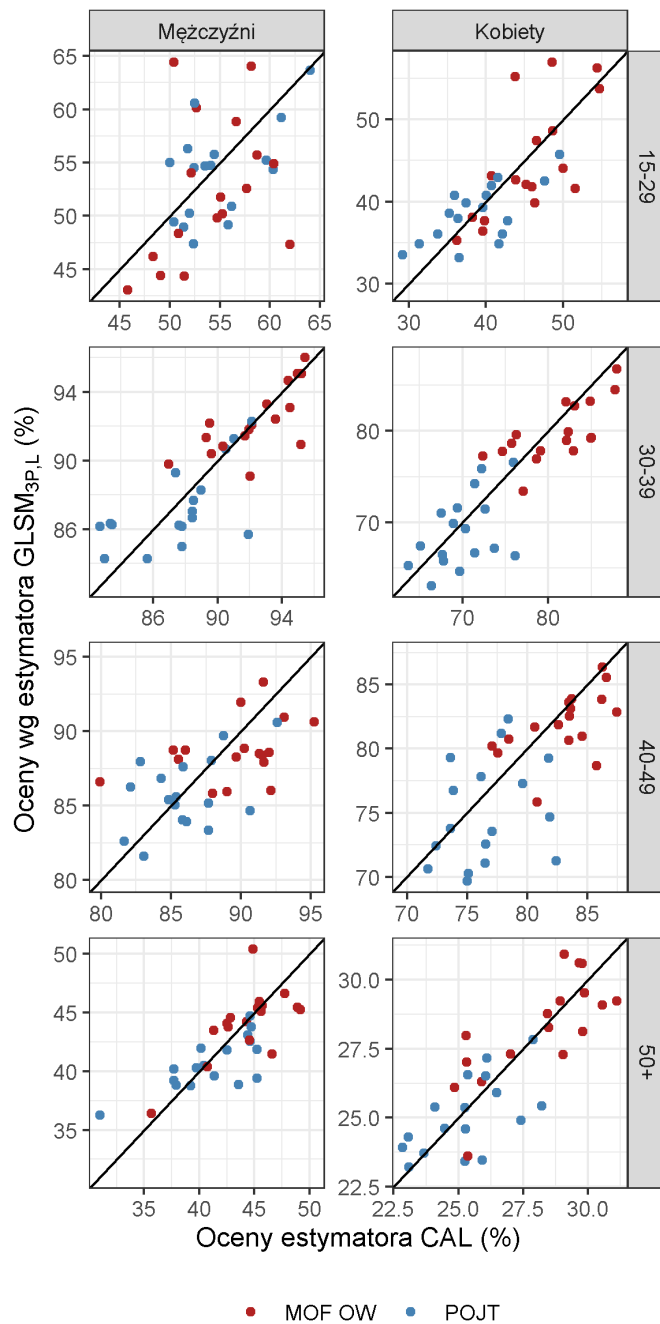
Źródło: Opracowanie własne.



Rysunek A.8: Porównanie szacunków wskaźnika aktywności ekonomicznej wyznaczonych z wykorzystaniem estymatora kalibracyjnego oraz $GLSM_{3P,L}$ dla docelowego przekroju tabeli – $CV \leq 20\%$

Objaśnienia: Punkty układające się wzdłuż prostej $y = x$ świadczą o zgodności oszacowań wskaźnika aktywności ekonomicznej wyznaczonych z wykorzystaniem estymatora kalibracyjnego oraz estymatora $GLSM_{3P,L}$.

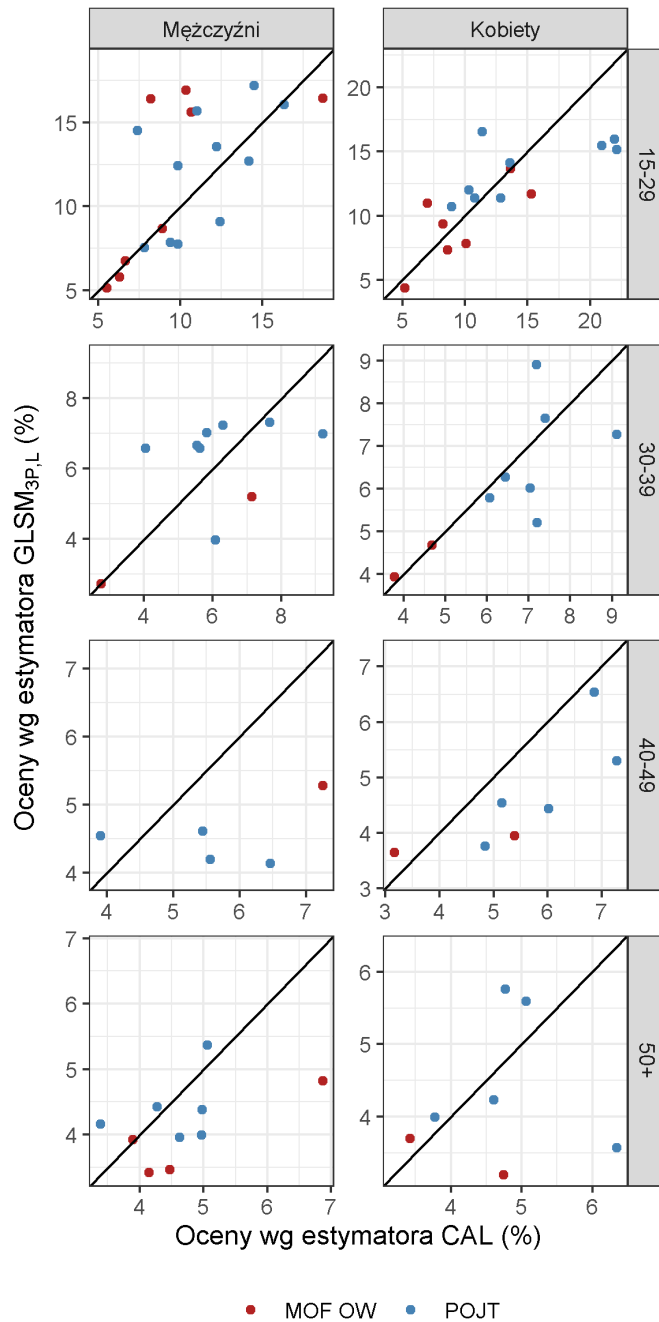
Źródło: Opracowanie własne.



Rysunek A.9: Porównanie szacunków wskaźnika zatrudnienia wyznaczonych z wykorzystaniem estymatora kalibracyjnego oraz $GLSM_{3PL}$ dla docelowego przekroju tabeli – $CV \leq 20\%$

Objaśnienia: Punkty układające się wzdłuż prostej $y = x$ świadczą o zgodności oszacowań wskaźnika zatrudnienia wyznaczonych z wykorzystaniem estymatora kalibracyjnego oraz estymatora $GLSM_{3PL}$.

Źródło: Opracowanie własne.



Rysunek A.10: Porównanie szacunków stopy bezrobocia wyznaczonych z wykorzystaniem estymatora kalibracyjnego oraz $GLSM_{3P,L}$ dla docelowego przekroju tabeli – $CV \leq 20\%$

Objaśnienia: Punkty układające się wzdłuż prostej $y = x$ świadczą o zgodności oszacowań stopy bezrobocia wyznaczonych z wykorzystaniem estymatora kalibracyjnego oraz estymatora $GLSM_{3P,L}$.

Źródło: Opracowanie własne.

Tabela A.5: Oszacowane wartości wskaźników wyznaczonych z wykorzystaniem estymatora $GLSM_{3P,L}$

JT	Wartość wskaźnika						Wartość współczynnika CV					
	MOF OW			POJT			MOF OW			POJT		
	r ^a	r ^e	r ^u	r ^a	r ^e	r ^u	r ^a	r ^e	r ^u	r ^a	r ^e	r ^u
Ogółem												
02	62,2	60,8	2,4	52,0	49,4	5,0	0,3	0,4	3,0	0,3	0,3	2,1
04	57,2	54,5	4,8	55,0	50,7	7,9	0,3	0,4	2,7	0,3	0,3	2,2
06	57,2	53,7	6,1	53,7	49,9	7,1	0,3	0,4	2,8	0,3	0,3	2,2
08G	57,4	55,7	2,9	54,4	51,6	5,1	0,6	0,6	2,9	0,6	0,6	2,2
08Z	57,2	55,2	3,6	52,0	49,4	5,1	0,5	0,5	2,7	0,5	0,5	2,1
10	58,2	55,0	5,5	57,1	54,5	4,4	0,3	0,4	2,7	0,3	0,3	2,2
12	61,8	59,9	3,0	55,7	53,2	4,5	0,3	0,3	2,8	0,3	0,3	2,2
14	63,4	61,6	2,9	58,2	54,4	6,6	0,3	0,3	2,8	0,3	0,3	2,1
16	54,8	52,2	4,7	50,7	48,2	4,9	0,4	0,4	2,6	0,3	0,4	2,1
18	59,3	55,1	7,1	53,0	49,0	7,7	0,3	0,4	2,9	0,3	0,3	2,2
20	56,8	53,0	6,6	53,2	49,8	6,4	0,3	0,4	2,8	0,3	0,3	2,2
22	58,2	56,3	3,3	54,5	51,8	5,0	0,3	0,3	2,7	0,3	0,3	2,1
24	54,7	52,3	4,3	53,7	51,8	3,5	0,4	0,4	2,6	0,3	0,3	2,4
26	56,2	52,4	6,8	52,2	48,7	6,8	0,4	0,4	2,8	0,3	0,3	2,2
28	58,9	56,0	4,9	52,9	48,7	7,9	0,3	0,4	2,6	0,3	0,3	2,2
30	63,9	62,8	1,8	58,3	56,4	3,3	0,3	0,3	3,4	0,3	0,3	2,5
32	54,8	52,8	3,7	51,9	48,2	7,1	0,4	0,4	2,6	0,3	0,4	2,1
Kobiety												
02	54,7	53,4	2,4	43,3	41,1	5,2	0,5	0,5	4,0	0,5	0,5	2,8
04	49,5	46,8	5,4	45,7	41,6	9,1	0,5	0,5	3,8	0,4	0,5	2,9
06	50,0	47,3	5,5	45,0	42,2	6,3	0,5	0,5	3,8	0,4	0,4	2,9
08G	49,8	48,2	3,1	45,2	42,6	5,7	0,8	0,8	3,9	0,7	0,8	2,9
08Z	49,9	47,9	3,9	43,1	40,6	5,8	0,7	0,7	3,9	0,7	0,7	2,8
10	50,6	48,1	5,0	48,4	46,2	4,4	0,4	0,5	3,8	0,4	0,4	2,9
12	54,7	53,1	2,9	47,7	45,6	4,5	0,4	0,5	3,9	0,4	0,4	2,9
14	57,4	55,9	2,6	50,0	47,0	6,0	0,4	0,4	4,0	0,4	0,4	2,8
16	47,0	44,5	5,2	41,5	39,2	5,5	0,6	0,6	3,8	0,5	0,5	2,8
18	51,7	48,3	6,6	44,0	40,8	7,4	0,5	0,6	3,9	0,5	0,5	2,9
20	49,5	46,8	5,5	43,7	41,3	5,5	0,5	0,5	3,8	0,5	0,5	2,8
22	51,4	49,5	3,8	45,7	43,0	5,9	0,5	0,5	3,9	0,5	0,5	2,8
24	47,2	45,0	4,8	45,6	43,9	3,7	0,5	0,6	3,8	0,4	0,5	3,0
26	48,7	45,8	6,1	43,4	40,6	6,4	0,5	0,6	3,9	0,4	0,5	2,8
28	52,1	49,6	4,8	43,9	40,2	8,4	0,5	0,5	3,8	0,5	0,5	2,9
30	56,2	55,1	2,0	49,1	47,2	3,8	0,4	0,5	4,2	0,4	0,4	3,0
32	48,3	46,4	4,0	43,6	40,2	7,8	0,5	0,6	3,8	0,5	0,5	2,8
Mężczyźni												
02	70,6	69,0	2,3	61,5	58,5	4,9	0,3	0,4	3,9	0,3	0,4	2,9
04	66,0	63,2	4,3	64,9	60,4	7,0	0,4	0,4	3,6	0,3	0,4	2,8
06	65,7	61,2	6,8	62,9	58,1	7,7	0,4	0,4	3,7	0,3	0,4	2,9

08G	65,8	64,1	2,7	64,2	61,1	4,8	0,6	0,7	4,0	0,5	0,6	3,0
08Z	65,3	63,1	3,4	61,5	58,7	4,5	0,5	0,5	3,7	0,5	0,5	3,0
10	67,4	63,4	6,0	66,4	63,4	4,4	0,3	0,4	3,5	0,3	0,3	3,0
12	69,8	67,6	3,1	64,1	61,2	4,5	0,3	0,4	3,7	0,3	0,3	3,1
14	70,6	68,4	3,1	66,9	62,2	7,0	0,3	0,3	3,6	0,3	0,4	2,9
16	63,5	60,8	4,3	60,7	57,9	4,6	0,4	0,4	3,6	0,4	0,4	3,0
18	67,5	62,5	7,5	62,5	57,6	7,9	0,4	0,4	3,9	0,3	0,4	2,9
20	65,1	60,1	7,6	63,1	58,6	7,1	0,4	0,5	3,8	0,3	0,4	2,9
22	65,7	63,8	2,9	63,8	61,0	4,4	0,4	0,4	3,8	0,3	0,4	3,0
24	62,8	60,4	3,9	62,5	60,4	3,3	0,4	0,4	3,6	0,3	0,4	3,3
26	64,4	59,7	7,3	61,6	57,2	7,1	0,4	0,5	3,7	0,3	0,4	2,9
28	66,7	63,4	5,0	62,4	57,7	7,5	0,4	0,4	3,6	0,3	0,4	2,8
30	72,6	71,5	1,5	68,2	66,2	2,9	0,3	0,4	4,4	0,3	0,3	3,5
32	62,1	59,9	3,6	60,8	56,8	6,6	0,4	0,5	3,6	0,4	0,4	2,8

15–29 lat

02	62,9	60,6	3,6	52,3	47,9	8,4	0,7	0,7	4,8	0,7	0,7	3,4
04	51,7	46,7	9,7	54,4	46,8	14,0	0,8	1,0	4,3	0,6	0,8	3,2
06	47,1	40,3	14,4	48,8	40,8	16,4	1,0	1,2	4,3	0,7	0,9	3,3
08G	49,5	46,6	5,9	51,9	47,5	8,5	0,9	1,0	4,4	0,7	0,8	3,4
08Z	49,7	45,7	8,0	51,7	47,2	8,6	0,9	1,0	4,3	0,7	0,8	3,4
10	56,0	50,7	9,4	55,8	51,3	8,1	0,7	0,9	4,3	0,6	0,7	3,4
12	60,1	56,9	5,3	53,7	48,8	9,0	0,7	0,8	4,5	0,6	0,7	3,4
14	60,4	57,5	4,7	55,7	48,6	12,7	0,7	0,8	4,5	0,6	0,7	3,3
16	52,4	47,7	9,1	50,8	46,4	8,8	0,8	0,9	4,3	0,7	0,7	3,4
18	49,7	41,9	15,8	49,2	41,4	15,8	0,9	1,2	4,3	0,7	0,9	3,2
20	48,1	41,2	14,3	50,7	43,9	13,5	0,9	1,2	4,4	0,7	0,8	3,3
22	55,4	51,7	6,5	53,2	48,2	9,3	0,7	0,8	4,4	0,7	0,7	3,3
24	54,2	50,0	7,9	55,5	52,2	6,0	0,8	0,9	4,3	0,6	0,7	3,7
26	46,1	39,1	15,2	49,9	42,4	15,1	1,0	1,3	4,3	0,7	0,9	3,2
28	48,9	43,4	11,2	49,9	42,9	14,0	0,9	1,1	4,3	0,7	0,8	3,2
30	61,5	59,5	3,2	58,6	55,1	5,9	0,7	0,7	4,9	0,6	0,7	3,7
32	48,7	45,2	7,1	49,2	43,0	12,7	0,9	1,0	4,3	0,7	0,8	3,2

30–39 lat

02	90,8	89,3	1,6	81,0	76,9	5,0	0,3	0,3	5,7	0,4	0,5	3,8
04	88,5	85,0	3,9	83,6	77,1	7,8	0,3	0,4	5,3	0,4	0,5	3,8
06	89,1	84,8	4,8	84,8	79,2	6,6	0,3	0,4	5,4	0,4	0,5	3,9
08G	87,6	85,6	2,3	81,1	77,0	5,1	0,4	0,4	5,5	0,5	0,6	3,8
08Z	88,2	85,8	2,8	80,5	76,6	4,9	0,3	0,4	5,4	0,5	0,5	3,8
10	91,3	87,3	4,3	87,6	83,8	4,3	0,2	0,3	5,4	0,3	0,4	3,9
12	90,9	89,1	2,1	83,6	80,2	4,1	0,3	0,3	5,5	0,4	0,4	3,9
14	92,8	90,8	2,1	87,7	82,2	6,3	0,2	0,3	5,5	0,3	0,4	3,8
16	85,5	82,5	3,6	78,1	74,5	4,6	0,4	0,5	5,3	0,5	0,6	3,9
18	90,3	85,9	4,9	81,5	75,6	7,2	0,3	0,4	5,4	0,4	0,5	3,8
20	88,2	83,6	5,1	83,1	78,2	5,9	0,3	0,5	5,5	0,4	0,5	3,9
22	88,6	86,4	2,4	81,0	77,2	4,6	0,3	0,3	5,4	0,4	0,5	3,8
24	89,1	86,2	3,3	84,0	81,3	3,2	0,3	0,4	5,4	0,4	0,4	4,1
26	88,4	83,8	5,2	83,9	78,5	6,4	0,3	0,4	5,4	0,4	0,5	3,8

28	88,5	85,4	3,6	81,2	75,1	7,5	0,3	0,4	5,4	0,4	0,6	3,8
30	91,5	90,3	1,3	86,0	83,4	3,0	0,2	0,3	5,8	0,3	0,4	4,1
32	86,5	84,0	2,9	80,1	74,6	6,8	0,4	0,4	5,4	0,5	0,6	3,8

40–49 lat

02	88,9	87,1	2,0	81,7	78,7	3,7	0,3	0,4	6,3	0,4	0,5	4,5
04	87,9	84,7	3,7	84,7	79,6	6,0	0,4	0,4	6,1	0,4	0,5	4,5
06	89,3	85,4	4,3	86,7	83,2	4,1	0,3	0,4	6,0	0,3	0,4	4,6
08G	86,4	84,7	2,0	80,6	77,4	3,9	0,4	0,5	6,4	0,5	0,5	4,5
08Z	87,1	84,9	2,6	80,9	78,0	3,6	0,4	0,4	6,2	0,5	0,5	4,5
10	90,4	86,0	4,9	88,8	86,1	3,1	0,3	0,4	6,0	0,3	0,3	4,6
12	89,5	87,4	2,4	85,1	82,8	2,8	0,3	0,4	6,2	0,4	0,4	4,6
14	91,3	89,1	2,5	88,7	84,7	4,5	0,3	0,3	6,2	0,3	0,4	4,5
16	83,6	80,9	3,2	79,6	76,9	3,4	0,5	0,6	6,1	0,5	0,5	4,6
18	90,3	85,9	4,8	83,6	79,5	4,9	0,3	0,4	6,0	0,4	0,5	4,5
20	87,2	83,3	4,5	83,5	80,1	4,1	0,4	0,5	6,1	0,4	0,4	4,6
22	86,4	84,3	2,5	81,8	79,0	3,4	0,4	0,5	6,2	0,4	0,5	4,6
24	87,1	84,2	3,3	84,5	82,3	2,5	0,4	0,5	6,1	0,4	0,4	4,7
26	88,5	84,2	4,8	85,6	81,9	4,3	0,4	0,5	6,0	0,3	0,4	4,5
28	88,5	85,5	3,4	81,5	76,8	5,8	0,4	0,4	6,1	0,4	0,5	4,5
30	90,6	89,4	1,3	86,9	85,0	2,2	0,3	0,3	6,7	0,3	0,4	4,8
32	84,7	82,2	2,9	80,1	75,7	5,4	0,5	0,5	6,1	0,5	0,6	4,5

50 lat i więcej

02	37,0	36,0	2,6	31,6	30,4	4,0	0,7	0,7	5,1	0,7	0,7	4,2
04	35,2	34,0	3,5	33,7	32,0	5,0	0,7	0,7	5,0	0,6	0,6	4,2
06	36,9	35,4	4,0	33,3	32,1	3,6	0,7	0,7	5,0	0,6	0,6	4,3
08G	36,6	35,7	2,4	34,8	33,4	3,9	0,7	0,8	5,1	0,6	0,7	4,2
08Z	36,6	35,6	2,7	31,4	30,2	3,8	0,7	0,7	5,1	0,7	0,7	4,2
10	36,1	34,3	4,9	35,0	33,9	3,1	0,7	0,7	5,0	0,6	0,6	4,3
12	37,3	36,2	2,8	34,0	33,1	2,8	0,7	0,7	5,1	0,6	0,6	4,4
14	38,6	37,5	2,9	36,5	35,0	4,2	0,7	0,7	5,1	0,6	0,6	4,2
16	34,1	32,8	3,9	31,7	30,5	3,8	0,7	0,8	5,0	0,7	0,7	4,2
18	38,1	36,3	4,8	32,9	31,5	4,3	0,7	0,7	5,0	0,6	0,7	4,2
20	36,9	35,2	4,8	34,1	32,8	3,8	0,7	0,7	5,1	0,6	0,6	4,2
22	34,9	34,0	2,5	34,1	32,9	3,5	0,7	0,7	5,1	0,6	0,6	4,2
24	30,4	29,2	3,8	30,5	29,6	2,9	0,9	0,9	5,0	0,7	0,7	4,3
26	36,4	34,7	4,5	31,6	30,4	3,9	0,7	0,7	5,1	0,7	0,7	4,2
28	38,6	37,2	3,7	33,5	31,6	5,5	0,7	0,7	5,0	0,6	0,7	4,2
30	39,8	39,1	1,6	35,5	34,7	2,3	0,7	0,7	5,4	0,6	0,6	4,5
32	34,0	32,8	3,3	32,9	31,2	5,3	0,7	0,8	5,0	0,6	0,7	4,2

Kobiety 15–29 lat

02	59,2	56,9	3,8	45,1	40,8	9,5	1,0	1,1	6,7	1,2	1,3	4,8
04	46,9	41,8	11,0	45,5	38,0	16,5	1,3	1,5	6,2	1,2	1,3	4,6
06	41,8	36,4	13,0	39,5	33,5	15,1	1,6	1,8	6,1	1,4	1,5	4,6
08G	46,0	43,1	6,3	43,8	39,3	10,3	1,4	1,5	6,4	1,2	1,3	4,8
08Z	45,5	41,6	8,6	43,3	38,6	10,9	1,4	1,5	6,2	1,2	1,3	4,8
10	52,0	47,4	8,8	46,6	42,5	8,7	1,2	1,3	6,2	1,2	1,2	4,9
12	56,5	53,7	4,9	44,8	40,8	9,0	1,1	1,2	6,5	1,2	1,3	4,9

14	58,8	56,2	4,4	47,3	41,9	11,4	1,0	1,1	6,6	1,2	1,3	4,7
16	47,6	42,6	10,4	42,2	37,7	10,7	1,3	1,5	6,2	1,3	1,3	4,8
18	44,0	37,7	14,3	39,2	33,2	15,5	1,5	1,7	6,1	1,4	1,5	4,6
20	43,1	38,1	11,7	41,0	36,0	12,0	1,5	1,7	6,1	1,3	1,4	4,7
22	52,5	48,6	7,4	45,0	39,9	11,4	1,2	1,3	6,3	1,2	1,3	4,7
24	48,6	44,0	9,3	46,1	42,9	6,8	1,3	1,4	6,2	1,2	1,2	5,1
26	40,8	35,3	13,7	40,6	34,9	14,1	1,6	1,9	6,1	1,3	1,4	4,7
28	44,9	39,8	11,4	41,5	34,9	16,0	1,4	1,6	6,1	1,3	1,4	4,6
30	57,3	55,2	3,7	49,3	45,7	7,1	1,1	1,1	6,7	1,1	1,2	5,0
32	45,6	42,1	7,8	42,2	36,0	14,5	1,4	1,5	6,3	1,3	1,4	4,6

Kobiety 30–39 lat

02	84,8	83,2	1,9	70,8	67,2	5,2	0,5	0,5	7,5	0,8	0,9	5,3
04	81,6	77,7	4,7	74,0	67,4	8,9	0,6	0,7	7,2	0,7	0,9	5,3
06	82,8	79,0	4,6	76,1	71,6	6,0	0,6	0,7	7,2	0,6	0,7	5,3
08G	80,0	77,8	2,7	70,2	66,3	5,6	0,7	0,7	7,3	0,8	0,9	5,3
08Z	81,8	79,2	3,2	70,5	66,5	5,7	0,6	0,7	7,3	0,8	0,9	5,3
10	86,3	82,8	4,1	80,1	76,6	4,4	0,5	0,5	7,3	0,6	0,6	5,3
12	85,1	83,2	2,2	74,3	71,0	4,4	0,5	0,5	7,4	0,7	0,7	5,3
14	88,6	86,8	2,1	80,5	75,9	5,8	0,4	0,4	7,4	0,5	0,6	5,3
16	76,8	73,4	4,4	66,5	63,1	5,2	0,8	0,9	7,2	1,0	1,0	5,3
18	84,2	79,9	5,1	70,9	65,8	7,3	0,5	0,6	7,2	0,8	0,9	5,3
20	81,5	77,8	4,5	73,0	69,3	5,1	0,6	0,7	7,2	0,7	0,8	5,3
22	82,1	79,6	3,1	70,5	66,7	5,5	0,6	0,6	7,3	0,8	0,9	5,3
24	81,8	78,6	3,9	74,1	71,4	3,6	0,6	0,7	7,3	0,7	0,7	5,4
26	81,6	77,3	5,3	74,5	69,9	6,3	0,6	0,7	7,2	0,7	0,8	5,3
28	82,4	79,3	3,8	70,6	65,2	7,7	0,6	0,7	7,3	0,8	0,9	5,3
30	85,9	84,5	1,7	77,0	74,3	3,6	0,5	0,5	7,5	0,6	0,7	5,4
32	79,6	77,0	3,3	70,0	64,7	7,6	0,7	0,7	7,3	0,8	0,9	5,3

Kobiety 40–49 lat

02	85,5	83,6	2,2	75,5	72,5	4,0	0,5	0,6	8,6	0,7	0,8	6,0
04	84,2	80,7	4,1	79,4	73,5	7,4	0,6	0,7	8,3	0,6	0,7	6,0
06	86,3	82,9	4,0	82,6	79,3	4,0	0,5	0,6	8,3	0,5	0,6	6,0
08G	83,6	81,7	2,3	74,4	71,1	4,5	0,6	0,7	8,5	0,8	0,8	6,0
08Z	83,4	81,0	2,9	74,4	71,3	4,2	0,6	0,7	8,4	0,8	0,8	6,0
10	87,9	83,9	4,5	85,0	82,3	3,2	0,5	0,6	8,2	0,5	0,5	6,1
12	86,0	83,8	2,5	80,4	77,8	3,2	0,5	0,6	8,5	0,6	0,6	6,1
14	88,5	86,4	2,4	85,0	81,2	4,5	0,4	0,5	8,5	0,5	0,5	6,0
16	78,8	75,9	3,7	73,1	70,2	3,9	0,9	0,9	8,3	0,8	0,8	6,0
18	87,3	83,1	4,7	77,9	73,8	5,3	0,5	0,6	8,2	0,7	0,7	6,0
20	83,9	80,7	3,9	77,6	74,7	3,8	0,6	0,7	8,3	0,7	0,7	6,1
22	82,6	80,2	2,9	75,6	72,4	4,2	0,7	0,7	8,4	0,7	0,8	6,0
24	82,7	79,6	3,6	79,0	76,7	2,8	0,7	0,7	8,3	0,6	0,7	6,2
26	85,6	81,8	4,4	80,9	77,3	4,4	0,5	0,7	8,3	0,6	0,6	6,0
28	85,4	82,5	3,4	75,6	70,6	6,5	0,5	0,6	8,3	0,7	0,8	6,0
30	86,9	85,6	1,5	81,5	79,3	2,7	0,5	0,5	8,8	0,6	0,6	6,2
32	81,2	78,7	3,1	74,3	69,6	6,2	0,7	0,8	8,4	0,8	0,9	6,0

Kobiety 50 lat i więcej

02	29,7	29,1	2,2	24,6	23,7	3,6	1,0	1,0	7,5	0,9	0,9	5,9
04	27,2	26,3	3,4	25,4	23,9	5,8	1,0	1,0	7,4	0,9	0,9	5,8
06	30,2	29,2	3,2	26,2	25,4	3,2	1,0	1,0	7,4	0,8	0,9	5,9
08G	28,9	28,3	2,3	26,9	25,9	3,8	1,0	1,0	7,5	0,8	0,9	5,9
08Z	28,9	28,1	2,6	24,1	23,2	3,7	1,0	1,0	7,4	0,9	0,9	5,9
10	28,5	27,3	4,2	27,3	26,5	3,0	1,0	1,0	7,4	0,8	0,9	6,0
12	30,2	29,5	2,4	27,9	27,2	2,7	1,0	1,0	7,4	0,8	0,9	6,0
14	31,7	30,9	2,4	29,0	27,8	4,0	1,0	1,0	7,5	0,8	0,9	5,9
16	27,1	26,1	3,6	24,3	23,4	3,7	1,0	1,0	7,4	0,9	0,9	5,9
18	30,5	29,2	4,2	26,0	24,9	4,2	1,0	1,0	7,4	0,8	0,9	5,8
20	29,9	28,8	3,9	26,3	25,4	3,4	1,0	1,0	7,4	0,8	0,9	5,9
22	28,0	27,3	2,5	26,4	25,4	3,7	1,0	1,0	7,4	0,8	0,9	5,9
24	24,5	23,6	3,7	25,0	24,3	2,8	1,1	1,1	7,4	0,9	0,9	6,0
26	29,0	28,0	3,6	24,4	23,5	3,7	1,0	1,0	7,4	0,9	0,9	5,9
28	31,6	30,6	3,1	26,1	24,6	5,6	1,0	1,0	7,4	0,8	0,9	5,8
30	31,1	30,6	1,6	27,2	26,5	2,5	1,0	1,0	7,6	0,8	0,9	6,0
32	27,8	27,0	2,9	25,9	24,6	5,1	1,0	1,0	7,4	0,9	0,9	5,8

Mężczyźni 15–29 lat

02	66,7	64,4	3,5	59,1	54,5	7,7	0,8	0,9	6,6	0,7	0,8	4,6
04	56,7	51,7	8,7	62,5	54,7	12,4	1,0	1,1	6,0	0,6	0,8	4,3
06	52,5	44,3	15,6	57,2	47,3	17,2	1,1	1,4	5,8	0,8	1,1	4,2
08G	53,1	50,2	5,5	59,6	55,2	7,3	1,1	1,1	6,2	0,7	0,8	4,6
08Z	53,8	49,8	7,4	59,2	55,0	7,1	1,0	1,1	6,1	0,7	0,8	4,6
10	59,9	54,1	9,8	64,2	59,2	7,7	0,9	1,0	6,0	0,6	0,8	4,6
12	63,8	60,2	5,8	61,9	56,3	9,1	0,8	0,9	6,2	0,7	0,8	4,4
14	62,0	58,8	5,1	63,2	54,7	13,5	0,9	0,9	6,3	0,6	0,9	4,2
16	57,2	52,5	8,1	58,8	54,4	7,5	1,0	1,1	6,0	0,7	0,8	4,6
18	55,6	46,2	16,9	58,3	49,0	16,1	1,0	1,3	5,7	0,7	1,0	4,2
20	53,1	44,4	16,4	59,5	50,9	14,5	1,1	1,4	5,8	0,7	0,9	4,2
22	58,3	54,9	5,8	60,5	55,8	7,9	0,9	1,0	6,2	0,7	0,8	4,5
24	59,7	55,7	6,8	64,1	60,6	5,5	0,9	1,0	6,1	0,6	0,7	4,9
26	51,5	43,0	16,4	58,3	49,2	15,7	1,2	1,5	5,8	0,7	1,0	4,2
28	53,2	47,3	11,0	57,6	50,2	12,7	1,1	1,2	5,9	0,7	0,9	4,3
30	65,8	64,0	2,7	67,1	63,7	5,1	0,8	0,9	6,8	0,6	0,7	4,9
32	51,7	48,3	6,5	55,8	49,4	11,4	1,1	1,2	6,2	0,8	0,9	4,3

Mężczyźni 30–39 lat

02	96,4	95,1	1,4	91,0	86,7	4,8	0,2	0,3	8,2	0,4	0,5	5,4
04	95,2	92,1	3,2	92,7	86,3	7,0	0,3	0,4	7,8	0,3	0,5	5,2
06	95,1	90,4	4,9	92,7	86,2	7,0	0,3	0,5	7,7	0,3	0,5	5,2
08G	94,1	92,2	2,1	90,5	86,2	4,7	0,4	0,4	8,0	0,5	0,5	5,4
08Z	94,7	92,4	2,4	90,9	87,0	4,3	0,4	0,4	7,9	0,4	0,5	5,4
10	96,2	91,8	4,6	94,7	90,7	4,3	0,2	0,4	7,7	0,2	0,4	5,4
12	96,5	94,7	1,9	92,8	89,3	3,8	0,2	0,3	8,0	0,3	0,4	5,5
14	97,2	95,1	2,2	94,6	88,3	6,7	0,2	0,3	7,9	0,3	0,5	5,3
16	93,7	90,9	3,0	89,5	85,7	4,2	0,4	0,5	7,8	0,5	0,6	5,4
18	95,9	91,3	4,7	91,6	85,0	7,2	0,3	0,5	7,7	0,4	0,6	5,2
20	94,4	89,1	5,6	92,2	86,1	6,6	0,4	0,6	7,7	0,4	0,5	5,3

22	94,9	93,1	1,9	91,3	87,7	4,0	0,3	0,4	8,0	0,4	0,5	5,5
24	95,9	93,3	2,7	94,1	91,3	3,0	0,3	0,3	7,8	0,3	0,3	5,7
26	94,7	89,8	5,2	92,4	86,3	6,6	0,3	0,5	7,7	0,4	0,5	5,3
28	94,6	91,4	3,4	90,9	84,3	7,3	0,4	0,4	7,8	0,4	0,6	5,2
30	97,0	96,0	1,0	94,7	92,3	2,6	0,2	0,2	8,4	0,2	0,3	5,8
32	93,1	90,8	2,5	89,9	84,3	6,3	0,5	0,5	7,9	0,5	0,6	5,3

Mężczyźni 40–49 lat

02	92,3	90,6	1,8	87,8	84,6	3,5	0,4	0,4	8,7	0,5	0,5	6,3
04	91,7	88,8	3,2	90,0	85,6	4,9	0,4	0,5	8,5	0,4	0,5	6,3
06	92,3	88,1	4,6	90,6	86,8	4,2	0,4	0,6	8,3	0,4	0,5	6,3
08G	89,5	87,9	1,8	86,9	83,9	3,4	0,6	0,6	8,8	0,5	0,6	6,4
08Z	90,7	88,5	2,4	86,8	84,1	3,2	0,5	0,5	8,6	0,5	0,6	6,4
10	93,1	88,3	5,2	92,5	89,7	3,0	0,4	0,6	8,3	0,3	0,4	6,4
12	93,1	90,9	2,3	89,8	87,6	2,5	0,4	0,4	8,6	0,4	0,4	6,5
14	94,3	92,0	2,5	92,2	88,0	4,5	0,3	0,4	8,6	0,3	0,4	6,3
16	88,5	86,0	2,8	86,0	83,4	3,1	0,6	0,7	8,5	0,6	0,6	6,4
18	93,3	88,7	4,9	89,2	85,1	4,6	0,3	0,6	8,3	0,4	0,5	6,3
20	90,6	85,9	5,2	89,0	85,2	4,3	0,5	0,7	8,3	0,4	0,5	6,3
22	90,2	88,4	2,0	87,8	85,4	2,7	0,5	0,6	8,7	0,5	0,5	6,4
24	91,5	88,7	3,0	90,0	88,0	2,2	0,4	0,5	8,5	0,4	0,4	6,5
26	91,4	86,6	5,3	90,0	86,2	4,1	0,4	0,7	8,3	0,4	0,5	6,3
28	91,7	88,6	3,4	87,1	82,6	5,2	0,4	0,5	8,4	0,5	0,6	6,3
30	94,4	93,3	1,2	92,3	90,6	1,8	0,3	0,3	9,1	0,3	0,3	6,7
32	88,2	85,8	2,7	85,7	81,6	4,7	0,6	0,7	8,5	0,6	0,7	6,3

Mężczyźni 50 lat i więcej

02	46,8	45,4	2,9	40,5	38,8	4,3	0,8	0,8	6,6	0,8	0,8	5,4
04	45,8	44,2	3,5	43,9	41,9	4,5	0,8	0,8	6,5	0,7	0,7	5,4
06	46,3	44,1	4,8	42,2	40,5	4,0	0,8	0,8	6,5	0,7	0,8	5,4
08G	46,8	45,6	2,4	44,3	42,6	4,0	0,8	0,8	6,7	0,7	0,7	5,4
08Z	46,4	45,1	2,8	40,3	38,7	3,9	0,8	0,8	6,6	0,8	0,8	5,4
10	47,2	44,6	5,5	44,6	43,1	3,3	0,8	0,8	6,5	0,7	0,7	5,5
12	46,7	45,3	3,1	41,4	40,2	2,9	0,8	0,8	6,6	0,7	0,8	5,6
14	48,3	46,6	3,4	45,8	43,8	4,4	0,8	0,8	6,5	0,7	0,7	5,4
16	43,2	41,5	4,1	41,0	39,4	3,9	0,8	0,9	6,5	0,7	0,8	5,4
18	47,9	45,4	5,2	41,4	39,6	4,4	0,8	0,8	6,5	0,7	0,8	5,4
20	46,4	43,8	5,7	43,6	41,8	4,2	0,8	0,8	6,5	0,7	0,7	5,4
22	43,8	42,7	2,5	43,3	41,8	3,3	0,8	0,9	6,7	0,7	0,7	5,5
24	37,9	36,4	3,9	37,4	36,2	3,0	1,0	1,1	6,5	0,9	0,9	5,5
26	45,9	43,5	5,2	40,4	38,8	4,0	0,8	0,8	6,5	0,8	0,8	5,4
28	48,0	45,9	4,2	42,6	40,3	5,4	0,8	0,8	6,5	0,7	0,8	5,3
30	51,2	50,4	1,6	45,7	44,7	2,1	0,8	0,8	6,9	0,7	0,7	5,8
32	41,9	40,4	3,7	41,4	39,2	5,4	0,9	0,9	6,5	0,7	0,8	5,3

Źródło: Opracowanie własne.

Spis tablic

4.1. Wykaz zmiennych i ich kategorii	117
4.2. Podkategorie zmiennych	118
4.3. Opis zaimplementowanych wariantów estymatorów	119
4.4. Oszacowania parametrów β rozpatrywanych wariantów estymatorów GLSM oraz GLSMM dla przekroju płci i grup wieku	133
A.1. Nazwy miejskich obszarów funkcjonalnych objętych instrumentem ZIT	192
A.2. Oszacowania parametrów β rozpatrywanych wariantów estymatorów GLSM oraz GLSMM dla przekroju ogółem	193
A.3. Oszacowania parametrów β rozpatrywanych wariantów estymatorów GLSM oraz GLSMM w przekroju płci	193
A.4. Oszacowania parametrów β rozpatrywanych wariantów estymatorów GLSM oraz GLSMM w przekroju grup wieku	193
A.5. Oszacowane wartości wskaźników wyznaczonych z wykorzystaniem estymatora $GLSM_{3P,L}$	203

Spis rysunków

1.1. Miejskie obszary funkcjonalne objęte instrumentem ZIT w Polsce w perspektywie finansowej Unii Europejskiej 2014–2020	28
1.2. Porównanie miejskich obszarów funkcjonalnych wyznaczonych dla miast wojewódzkich w ramach programu Urban Audit do realizacji instrumentu ZIT . . .	35
2.1. Zakres informacji pozyskiwanych w BAEL dla poszczególnych zbiorowości . .	50
2.2. Przebieg procedury przygotowującej zbiory liczby osób pracujących oraz bezrobotnych w głównym miejscu pracy na podstawie AZD	57
2.3. Porównanie liczby osób pracujących w podziale na płeć na poziomie województw w 2017 roku pomiędzy poszczególnymi źródłami danych	70
2.4. Porównanie liczby osób bezrobotnych w podziale na płeć na poziomie województw w 2017 roku pomiędzy poszczególnymi źródłami danych	72
4.1. Schemat procedury badawczej przeprowadzonej w ramach niniejszej rozprawy	112
4.2. Przekroje tabel ogółem w podziale na kategorie rynku pracy	124
4.3. Przekroje tabel uwzględniające płeć w podziale na kategorie rynku pracy . . .	125
4.4. Przekroje tabel uwzględniające grupy wieku w podziale na kategorie rynku pracy	127
4.5. Przekroje tabel uwzględniające płeć i grupy wieku w podziale na kategorie rynku pracy	128
4.6. Rozkład wartości współczynnika precyzji dla kategorii rynku pracy w układzie tabel uwzględniających płeć i grupy wieku	130

4.7. Porównanie szacunków pośrednich i kalibracyjnych w układzie tabel uwzględniających płeć i grupy wieku – $CV \leq 20\%$	132
4.8. Względne różnice pomiędzy ocenami estymatora kalibracyjnego, a poszczególnymi wariantami estymatorów typu SPREE w przyjętych przekrojach	135
5.1. Porównanie ocen precyzji współczynnika aktywności ekonomicznej	141
5.2. Porównanie ocen precyzji wskaźnika zatrudnienia	143
5.3. Porównanie ocen precyzji dla stopy bezrobocia	144
5.4. Zestawienie zysku na precyzji dla poszczególnych wskaźników w analizowanych przekrojach wyznaczonych z wykorzystaniem estymatora $GLSM_{3P,L}$ w porównaniu z oszacowaniami za pomocą estymatora kalibracyjnego	145
5.5. Procentowe różnice pomiędzy oszacowaniami wskaźników uzyskanymi z wykorzystaniem estymatora $GLSM_{3P,L}$ i kalibracyjnego	147
5.6. Procentowe różnice pomiędzy oszacowaniami stopy bezrobocia uzyskanymi z wykorzystaniem estymatora $GLSM_{3P,L}$ i kalibracyjnego	148
5.7. Porównanie rozkładu współczynnika aktywności ekonomicznej pomiędzy MOF OW a POJT w ramach poszczególnych przekrojów	151
5.8. Przestrzenne zróżnicowanie współczynnika aktywności ekonomicznej w analizowanych przekrojach	152
5.9. Porównanie rozkładu wskaźnika zatrudnienia pomiędzy MOF OW a POJT w ramach poszczególnych przekrojów	154
5.10. Przestrzenne zróżnicowanie wskaźnika zatrudnienia ludności w analizowanych przekrojach	155
5.11. Porównanie rozkładu stopy bezrobocia pomiędzy MOF OW a POJT w ramach poszczególnych przekrojów	157
5.12. Przestrzenne zróżnicowanie stopy bezrobocia w analizowanych przekrojach	158
5.13. Ranking miejskich obszarów funkcjonalnych ośrodków wojewódzkich – metoda TOPSIS	161
5.14. Grupowanie miejskich obszarów funkcjonalnych ośrodków wojewódzkich metodą Warda ze względu na sytuację panującą na rynku pracy	162
5.15. Grupowanie miejskich obszarów funkcjonalnych ośrodków wojewódzkich metodą Warda – ujęcie przestrzenne	162

A.1. Porównanie liczby osób biernych zawodowo w podziale na płeć na poziomie województw w 2017 roku	191
A.2. Rozkład wartości współczynnika precyzji dla kategorii rynku pracy w układzie tabel ogółem	194
A.3. Porównanie szacunków pośrednich i kalibracyjnych w układzie tabel ogółem – $CV \leq 20\%$	195
A.4. Rozkład wartości współczynnika precyzji dla kategorii rynku pracy w układzie tabel uwzględniających płeć	196
A.5. Porównanie szacunków pośrednich i kalibracyjnych w układzie tabel uwzględniających płeć – $CV \leq 20\%$	197
A.6. Rozkład wartości współczynnika precyzji dla kategorii rynku pracy w układzie tabel uwzględniających grupy wieku	198
A.7. Porównanie szacunków pośrednich i kalibracyjnych w układzie tabel uwzględniających grupy wieku – $CV \leq 20\%$	199
A.8. Porównanie szacunków wskaźnika aktywności ekonomicznej wyznaczonych z wykorzystaniem estymatora kalibracyjnego oraz $GLSM_{3P,L}$ dla docelowego przekroju tabeli – $CV \leq 20\%$	200
A.9. Porównanie szacunków wskaźnika zatrudnienia wyznaczonych z wykorzystaniem estymatora kalibracyjnego oraz $GLSM_{3P,L}$ dla docelowego przekroju tabeli – $CV \leq 20\%$	201
A.10. Porównanie szacunków stopy bezrobocia wyznaczonych z wykorzystaniem estymatora kalibracyjnego oraz $GLSM_{3P,L}$ dla docelowego przekroju tabeli – $CV \leq 20\%$	202