

Wydział Informatyki i Gospodarki Elektronicznej



UNIWERSYTET EKONOMICZNY
W POZNANIU

Praca doktorska

**DELIMITACJA LOKALNYCH RYNKÓW PRACY W POLSCE
Z WYKORZYSTANIEM MODELOWANIA
WIELOPOZIOMOWEGO**

mgr Hanna Gruchociak

Promotor

dr hab. Elżbieta Gołata, prof. nadzw. UEP

Promotor pomocniczy

dr Tomasz Klimanek

Uniwersytet Ekonomiczny w Poznaniu

Wydział Informatyki i Gospodarki Elektronicznej

Katedra Statystyki

Poznań 2013

Spis treści

WSTĘP	5
Rozdział 1. Wprowadzenie do problematyki lokalnego rynku pracy	15
1.1. Definicje lokalnego rynku pracy	15
1.2. Wprowadzenie do delimitacji lokalnych rynków pracy.....	18
1.2.1. Cel i uzasadnienie potrzeby delimitacji lokalnych rynków pracy.....	18
1.2.2. Koncepcje delimitacji lokalnych rynków pracy	20
1.2.3. Pożądane cechy lokalnego rynku pracy.....	22
1.3. Dojazdy do pracy w Polsce	25
1.3.1. Charakterystyka źródeł danych	25
1.3.2. Struktura i przestrzenne zróżnicowanie dojazdów do pracy	29
1.4. Charakterystyka lokalnych rynków pracy w Polsce	34
Rozdział 2. Metody delimitacji lokalnych rynków pracy	42
2.1. Metody delimitacji ich klasyfikacja	42
2.2. Stan badań związanych z delimitacją lokalnych rynków pracy w Polsce.....	48
2.3. Europejski algorytm regionalizacji	51
2.4. Algorytm Kristensena	60
2.5. Podejście taksonomiczne.....	62
Rozdział 3. Modelowanie wielopoziomowe	69
3.1. Idea, korzyści i założenia modelowania wielopoziomowego	69
3.2. Klasyczny model wielopoziomowy	77
3.3. Model wielopoziomowy z dwoma kryteriami grupowania.....	97
3.4. Estymacja parametrów modelu	108
3.5. Kryteria oceny dopasowania modelu	117

Rozdział 4. Oszacowanie macierzy migracji dla roku 2011	122
4.1. Oszacowanie wartości na diagonalu	123
4.2. Oszacowanie wartości poza diagonalą.....	153
4.3. Ocena jakości oszacowania elementów macierzy migracji	183
Rozdział 5. Delimitacja lokalnych rynków pracy w Polsce	188
5.1. Delimitacja lokalnych rynków pracy w Polsce dla roku 2006.....	188
5.1.1. Zastosowanie europejskiego algorytmu regionalizacji.....	188
5.1.2. Zastosowanie algorytmu Kristensena	197
5.1.3. Zastosowanie podejścia taksonomicznego	200
5.2. Próba delimitacji lokalnych rynków pracy w Polsce dla roku 2011	207
5.3. Podsumowanie otrzymanych wyników.....	216
ZAKOŃCZENIE	221
LITERATURA	228
SPIS TABEL	241
SPIS SCHEMATÓW	248
SPIS WYKRESÓW	249
Aneks	251
Aneks A.1. Opis trudności obliczeniowych i kody źródłowe	251
Aneks A.2. Skład wyznaczonych lokalnych rynków pracy	258

WSTĘP

Uzasadnienie podjęcia tematyki badawczej

Delimitacja lokalnych rynków pracy jest zagadnieniem niezmiernie istotnym z punktu widzenia wspomaganie zrównoważonego rozwoju regionów. Rozwój samorządności lokalnej powoduje wzrost zapotrzebowania na informacje niezbędne dla opracowania realizacji strategii rozwoju gospodarczego i społecznego, efektywnego zarządzania, a także szeroko rozumianej poprawy warunków życia ludności. Znajomość struktury lokalnych rynków pracy może okazać się kluczowa przy planowaniu inwestycji, podejmowaniu decyzji o kierunkach rozwoju infrastruktury czy pozyskiwaniu funduszy (np. unijnych). W wielu krajach europejskich o rozwiniętej gospodarce rynkowej delimitacja lokalnych rynków pracy przeprowadzana jest regularnie, a jej wyniki są powszechnie wykorzystywane w polityce regionalnej. Jak podaje Organizacja Współpracy Gospodarczej i Rozwoju, tylko pięć państw członkowskich nie definiowało lokalnych rynków pracy (por. OECD, 2002; Casado-Diaz, Coombes, 2011).

Opisywane w literaturze i stosowane w innych krajach algorytmy służące do delimitacji lokalnych rynków pracy różnią się pod wieloma względami, jednak wszystkie mają przynajmniej jedną cechę wspólną, aby je zastosować konieczna jest znajomość macierzy przepływów ludności związanych z zatrudnieniem. Jest to macierz kwadratowa o wymiarze równym liczbie jednostek bazowych, która zawiera na przecięciu *i-tego* wiersza i *j-tej* kolumny informację o liczbie osób zamieszkałych na terenie *i-tej* jednostki bazowej i jednocześnie zatrudnionych na terenie *j-tej* jednostki, tj. przyjeżdżających o pracy z jednostki *i-tej* do jednostki *j-tej*. Na diagonalu tej macierzy, są więc informacje o liczbie osób pracujących na terenie badanej jednostki bazowej i jednocześnie zamieszkujących na jej terenie.

Zapoczątkowana w 1989 roku transformacja polskiej gospodarki, doprowadziła do wielu istotnych zmian na naszym rynku pracy. Prywatyzacja przedsiębiorstw państwowych, likwidacja lub restrukturyzacja nierentownych przedsiębiorstw, a przede wszystkim zmiany strukturalne w polskiej gospodarce doprowadziły do wzrostu bezrobocia. Nasiliły się problemy związane ze zbilansowaniem popytu i podaży pracy. Wzrost przedsiębiorczości Polaków oraz ekspansja kapitału zagranicznego doprowadziły

do powstania tysięcy nowych przedsiębiorstw o różnej wielkości. Wzrost gospodarczy oraz rozwój różnych form lokalnej samorządności spowodowały zwiększone zapotrzebowanie na bieżącą informację oraz monitorowanie zmian na niskim poziomie agregacji terytorialnej. Trzeba jednak pamiętać, że brak odpowiednich danych uniemożliwił przeprowadzenie, od lat osiemdziesiątych ubiegłego stulecia, kompleksowej delimitacji lokalnych rynków pracy, delimitacji opartej na informacji o rzeczywistych dojazdach do pracy, ich natężeniu i kierunku.

W 2010 roku Urząd Statystyczny w Poznaniu opublikował wyniki unikatowego badania przepływów ludności. Badania dotyczyły zatrudnienia w 2006 r. i wykorzystywały dane pozyskane z zasobów rejestrów podatkowych Ministerstwa Finansów (por. Dojazdy do pracy w Polsce, red. K. Kruszka, 2010; Filas-Przybył, Stachowiak, 2010). Do szacunków włączono informacje o płatnikach podatku (zarówno osobach fizycznych, jak i zatrudniających je przedsiębiorstwach), zgromadzone w kilku różnych rejestrach podatkowych systemu POLTAX. W szczególności, integracja danych z formularzy PIT-11/8B oraz PIT-40, pozwoliła oszacować macierz przepływów pracowników najemnych, którzy dojeżdżają do pracy poza gminę, w której zamieszkują. Wyniki badania, które poprzedzone było kilkuletnimi pracami studialnymi w Urzędzie Statystycznym w Poznaniu, zostały w pełni opublikowane w 2011 roku. Udostępnione dane o dojazdach do pracy, umożliwiają przeprowadzenie pierwszej - od wielu lat - delimitacji lokalnych rynków pracy w Polsce, odpowiadającej standardom europejskim.

Od początku transformacji, wspomniane badanie przepływów ludności związanych z zatrudnieniem, obejmujące informację o kierunkach dojazdów do pracy zostało przeprowadzone tylko raz. Z tego powodu, możliwe jest przeprowadzenie delimitacji lokalnych rynków pracy w Polsce aktualnej dla roku 2006. W pracy podjęto próbę wykorzystania metodologii modelowania wielopoziomowego do estymacji macierzy migracji. Uzyskanie zadowalających rezultatów umożliwiłoby predykcję macierzy przepływów ludności związanych z zatrudnieniem także dla lat późniejszych.

Cele pracy

Głównym celem badawczym pracy doktorskiej jest przeprowadzenie delimitacji lokalnych rynków pracy w Polsce. Wynik delimitacji określony jest przez wybór algorytmu, modelu, zmiennych, ich natężenia oraz skali występujących między nimi relacji. Punktem wyjściowym pracy jest przegląd metod delimitacji, ich porównanie oraz

ocena. Pozwoliło to na wybór takiego podejścia metodologicznego, które byłoby najlepiej dostosowane do polskich realiów. W związku z tak określonym celem głównym sformułowano następujące cele szczegółowe:

1. Dokonanie przeglądu i oceny metod delimitacji lokalnych rynków pracy.
2. Dostosowanie europejskiego algorytmu regionalizacji do warunków polskich oraz zaproponowanie autorskiego podejścia optymalnego dla realiów polskich.
3. Przedstawienie idei modelowania wielopoziomowego i skonstruowanie modelu wielopoziomowego w celu oszacowania dojazdów do pracy w roku 2011.
4. Estymacja macierzy przepływów ludności związanych z zatrudnieniem.
5. Zastosowanie wybranych podejść zaczerpniętych z literatury oraz autorskiego modelu do delimitacji lokalnych rynków pracy w Polsce.

Wskazane cele szczegółowe wiążą się z poszczególnymi etapami prowadzonego badania. Realizacja pierwszego z wymienionych celów związana jest, z przeprowadzonymi wnikliwymi studiami literaturowymi. Pozwoliły one na zaprezentowanie stosowanych w innych państwach metod delimitacji lokalnych rynków pracy. Na podstawie literatury, opisującej dotychczasowe zastosowania poszczególnych modeli oraz ich wstępnej weryfikacji empirycznej, stwierdzono potrzebę ich dostosowania do polskich warunków. Doświadczenia płynące z tego etapu badań pozwoliły do przejścia do drugiego etapu i zaproponowanie autorskiego algorytmu łączącego zalety proponowanych w literaturze rozwiązań. Dzięki takiemu podejściu przeprowadzona delimitacja, pozwoli w pełni uwzględnić specyfikę polskiego rynku pracy.

Zastosowane podejścia, zarówno zaczerpnięte z literatury jak i zaproponowane przez autorkę, bazują na danych macierzy przepływów związanych z zatrudnieniem, nazywanej także macierzą migracji. Informacje takie dostępne są dla roku 2006, na podstawie wyników wspomnianego wyżej badania. Brak dostępności macierzy migracji dla lat późniejszych, był inspiracją dla zaproponowania metody jej estymacji przy pomocy modelu wielopoziomowego. Tak więc kolejnym, trzecim celem szczegółowym pracy jest przedstawienie idei oraz sposobu konstrukcji modeli wielopoziomowych. Czwarty cel szczegółowy badań polegał na empirycznym oszacowaniu macierzy przepływów związanych z zatrudnieniem dla roku 2011, przy wykorzystaniu modelowania wielopoziomowego. Ostatni wskazany cel szczegółowy ma również charakter empiryczny. Jest nim delimitacja lokalnych rynków pracy w Polsce. Delimitacja zostanie

przeprowadzona na podstawie dwóch podejść zaczerpniętych z literatury oraz autorskiego algorytmu, bezpośrednio na podstawie danych z badania przepływów ludności związanych z zatrudnieniem dla roku 2006. Następnie podjęta zostanie próba przeprowadzenia delimitacji również dla roku 2011. W tym celu wykorzystane zostaną szacunki macierzy migracji otrzymane za pomocą modelu wielopoziomowego.

Hipotezy badawcze

W trakcie realizacji powyższych celów szczegółowych sformułowano następujące hipotezy badawcze:

1. Lokalne rynki pracy można dobrze scharakteryzować przy pomocy danych o dojazdach do pracy, które mogą stanowić podstawę delimitacji tych rynków.
2. Zastosowanie modelowania wielopoziomowego umożliwi oszacowanie macierzy przepływów związanych z zatrudnieniem.
3. W polskich warunkach, bardziej uzasadniona jest delimitacja lokalnych rynków pracy niewymagająca wyodrębniania ośrodków centralnych.
4. Duża liczba parametrów, w stosowanych algorytmach, stwarza możliwość subiektywnego kształtowania wyników delimitacji.
5. Delimitacja lokalnych rynków pracy powinna być przeprowadzana dla całego kraju, a nie tylko w odniesieniu do poszczególnych województw czy regionów.

Przygotowana rozprawa ma charakter teoretyczno-empiryczny. Wśród wykorzystanych metod badawczych znalazły się studia literaturowe, metody statystyki wielowymiarowej w tym taksonomiczne, a w szczególności delimitacji lokalnych rynków pracy, metody modelowania wielopoziomowego oraz wnioskowania statystycznego.

Zakres przestrzenny i czasowy, źródła danych oraz zastosowane metody badawcze

Przedmiotem rozprawy jest delimitacja lokalnych rynków pracy w Polsce dla roku 2006 oraz oszacowanie macierzy migracji umożliwiającej delimitację dla roku 2011. Jako bazową jednostkę terytorialną, będącą podstawą określenia lokalnego rynku pracy, przyjęto gminę z uwzględnieniem podziału na część miejską i wiejską, tj. jednostkę najniższego szczebla podziału administracyjnego kraju.

Głównym źródłem danych statystycznych były wyniki badania przepływów ludności związanych z zatrudnieniem. Przeprowadzono je w Ośrodku Statystyki Miast Urzędu

Statystycznego w Poznaniu. Dane pozyskano przeprowadzając integrację zasobów systemu POLTAX dla 2006 roku. Wykorzystano również inne, zaczerpnięte z banku danych lokalnych, informacje dotyczące aktywności ekonomicznej ludności oraz sytuacji gospodarczej w przekroju lokalnym dla lat 2006 i 2011. Były to między innymi: badanie podmiotów gospodarki narodowej wpisanych do rejestru REGON; pracujących według miejsca pracy; warunków pracy; nakładów na działalność innowacyjną w przedsiębiorstwach; dane o liczbie bezrobotnych zarejestrowanych w urzędach pracy; badanie ludności oraz ruchu naturalnego; informacje o przeciętnych wynagrodzeniach; transporcie kolejowym; liczbie studentów i absolwentów szkół wyższych; komunikacji miejskiej oraz drogach publicznych powiatowych i gminnych.

Wśród zastosowanych w rozprawie metod badawczych wymienić można:

1. Metody delimitacji lokalnych rynków pracy (empiryczna weryfikacja modeli proponowanych w literaturze):
 - europejski algorytm regionalizacji, opracowany przez Coombesa
 - algorytm zaproponowany przez Kristensena
2. Autorski model delimitacji lokalnych rynków pracy.
3. Modelowanie wielopoziomowe:
 - klasyczny model wielopoziomowy
 - model wielopoziomowy z dwoma kryteriami grupowania

W pracy przeprowadzono studia literaturowe w zakresie lokalnych rynków pracy i ich delimitacji. Dokonano analizy porównawczej stosowanych modeli, ich założeń i uwarunkowań oraz otrzymanych w warunkach polskich rezultatów. Na tej podstawie zaproponowano autorski model uwzględniający specyfikę lokalnego rynku pracy w Polsce.

Oszacowanie macierzy przepływów związanych z zatrudnieniem postanowiono oprzeć na możliwie jak najszerszym wachlarzu charakterystyk rynku pracy, stopnia rozwoju regionów oraz stanie ich infrastruktury. Informacje z tego zakresu publikowane są na różnych poziomach agregacji (gmina, powiat, województwo). W celu rozwiązania problemu szacowania wielkości dla gmin, z wykorzystaniem informacji z poziomu

powiatów oraz województw, zaproponowano zastosowanie modelowania wielopoziomowego. Modele te, uwzględniając relacje pomiędzy czynnikami określonymi dla jednostek indywidualnych oraz na poziomach zagregowanych, stwarzają możliwość odzwierciedlenia skomplikowanej struktury dojazdów do pracy. Ponadto, uwzględniając różne przyczyny zmienności, pozwalają one uzyskać dobre dopasowanie do danych empirycznych. Z tego powodu, modele wielopoziomowe, wykorzystywane są w statystyce małych obszarów. J.N.K. Rao (2003) określa je mianem modeli dla małych domen. Uwzględniając konstrukcję macierzy migracji oraz odmienny charakter wielkości na diagonalu i poza diagonalą, skonstruowano dwa odrębne modele.

W celu estymacji wartości na diagonalu zastosowano klasyczny model trypoziomowy. Celem konstrukcji tego modelu było oszacowanie liczby osób pracujących w gminie zamieszkania w przekroju gmin. Gminy 'zagnieżdżone' są w powiatach i województwach, co zdeterminowało trypoziomową strukturę modelu. W celu oszacowania wartości poza diagonalą posłużono się modelem czteropoziomowym z dwoma kryteriami grupowania określonymi na poziomie drugim. Celem jego konstrukcji było oszacowanie przepływów związanych z zatrudnieniem w przekroju par gmin. Dobór modelu czteropoziomowego z dwoma kryteriami grupowania zdeterminowany został strukturą par gmin. Jako pierwsze kryterium grupowania przyjęto gminę z której dany przepływ związany z zatrudnieniem pochodzi, drugie kryterium grupowania określono jako gminę, do której jest on skierowany. Gminy te z kolei 'zagnieżdżone' są w powiatach oraz województwach.

Wkładem własnym autorki jest zaproponowanie algorytmu delimitacji lokalnych rynków pracy w Polsce. Oryginalnym rozwiązaniem jest także propozycja szacowania macierzy migracji związanych z zatrudnieniem przy wykorzystaniu modelowania wielopoziomowego. Jako novum praktyczne traktować można przeprowadzenie kompleksowej delimitacji lokalnych rynków pracy w Polsce przy wykorzystaniu europejskiego algorytmu regionalizacji oraz autorskiego podejścia wraz z analizą rezultatów.

Struktura rozprawy

Prezentowana rozprawa doktorska składa się z pięciu rozdziałów i aneksu oraz wstępu i zakończenia.

W **rozdziale pierwszym** omówiono podstawowe kwestie związane z lokalnym rynkiem pracy. W szczególności określono zakres pojęcia lokalnego rynku pracy. Następnie, zaprezentowana została idea delimitacji lokalnych rynków pracy, która jest wspólna dla wszystkich algorytmów przedstawionych szczegółowo w rozdziale drugim. Na podstawie literatury przedmiotu, autorka opisała czynniki kształtujące lokalny rynek pracy, wskazując między innymi te z nich, które dobrze nadają się do opisu lokalnego rynku pracy oraz oszacowania charakterystyk dojazdów do pracy. Na końcu rozdziału scharakteryzowane zostały źródła, z których pozyskane mogą zostać informacje dotyczące lokalnych rynkach pracy w warunkach polskich.

W **rozdziale drugim** omówiono metody delimitacji lokalnych rynków pracy opisane w literaturze oraz stosowane w innych państwach. Ponadto zaproponowana została autorska metoda oparta na podejściu taksonomicznym, łącząca zalety innych podejść.

Pierwszym z podejść zaczerpniętych z literatury jest algorytm opracowany przez M. G. Coombesa, A. E. Greena i S. Openshawa do delimitacji lokalnych rynków pracy w Wielkiej Brytanii (por. Coombes, Green, Openshaw, 1986). Procedura ta znana jest, jako **europejski algorytm regionalizacji** i rekomendowana przez Eurostat, jako standardowe podejście do definiowania lokalnych rynków pracy w krajach europejskich (por. Eurostat, 1992). Oprócz Wielkiej Brytanii jest ona stosowana w wielu krajach, między innymi we Włoszech, Danii, Nowej Zelandii oraz Australii.

W pierwszej fazie europejskiego algorytmu regionalizacji wytypowane zostały te jednostki bazowe, które pełnią rolę ośrodków centralnych, wokół których skupiają się lokalne rynki pracy. Następnie jednostki bazowe przydzielane są do ośrodków centralnych, z którymi wykazują istotny stopień interakcji. W trakcie działania algorytmu niektóre z wcześniej wybranych ośrodków centralnych mogą stracić ten status. Również przynależność jednostki bazowej do danego lokalnego rynku pracy może zostać negatywnie zweryfikowana. Wszystkie kryteria, zarówno służące do wyboru ośrodków centralnych, jak i do przydzielania jednostek bazowych do lokalnych rynków pracy, obliczane są tylko i wyłącznie na podstawie danych zawartych w macierzy przepływów związanych z zatrudnieniem. Ponadto należy określić wartości wielu parametrów zależnych, między innymi, od liczby i wielkości jednostek bazowych w danym kraju oraz od natężenia i łatwości przejazdów między tymi jednostkami.

Jako kolejne podejście zaczerpnięte z literatury, wykorzystano **algorytm** zaproponowany przez K. **Kristensena**, do delimitacji lokalnych rynków pracy w Danii (por. Kristensen, 1998). Algorytm ten podzielić można na trzy główne etapy. W pierwszym z nich następuje transformacja macierzy przepływów związanych z zatrudnieniem na macierz binarną o takim samym wymiarze, w której uwzględniane są wyłącznie te przepływy, które zostaną uznane za istotne. Następnie, na podstawie okrojonej, zgodnie z powyższym opisem, macierzy przepływów związanych z zatrudnieniem podjęta zostaje próba wyodrębnienia zbiorów zamkniętych. Jako zbiór zamknięty rozumiany jest taki podzbiór jednostek bazowych, z których nie istnieje istotny wyjazd ludności związany z zatrudnieniem do jednostek spoza tego podzbioru. Jeżeli zidentyfikowane zostaną zbiory zamknięte, to uznawane one zostają za lokalne rynki pracy. Pozostałe jednostki bazowe łączone są w etapie trzecim w lokalne rynki pracy według funkcji podobieństwa. Przebiega to w sposób hierarchiczny, to znaczy, jeżeli tworzenie danego lokalnego rynku pracy zostanie przerwane, nie ma już możliwości późniejszego modyfikowania jego składu. W celu zastosowania tego podejścia należy określić wartości dwóch parametrów.

Jako alternatywę dla opisanych algorytmów postanowiono zaproponować podejście łączące zalety obu z nich. Ze względu na wykorzystywaną metodologię, zaproponowana metoda nazywana dalej będzie **podejściem taksonomicznym**. Podobnie, jak w opisanych podejściach, informacją wejściową jest macierz przepływów związanych z zatrudnieniem, zaś rezultatem, podział jednostek bazowych, na rozłączne i pokrywające całą zbiorowość, lokalne rynki pracy. W opracowanym algorytmie nie ma jednak konieczności określania żadnych parametrów wejściowych, a w trakcie postępowania nie są wyodrębniane ośrodki centralne. Ponadto, uwzględniane są zarówno przepływy pomiędzy jednostkami bazowymi jak i grupami jednostek bazowych, a procedura tworzenia lokalnych rynków pracy przebiega równolegle, czyli w trakcie trwania algorytmu do każdego z wcześniej wyodrębnionych lokalnych rynków pracy dołączona może zostać kolejna jednostka bazowa lub inny lokalny rynek pracy. Algorytm ten opiera się na łączeniu tych jednostek bazowych, pomiędzy którymi występuje silny przepływ związany z zatrudnieniem.

Rozdział trzeci, podobnie jak dwa wcześniejsze, ma charakter teoretyczny. Zawiera on opis metodologii modelowania wielopoziomowego, pozwalającej na uwzględnienie zróżnicowania poziomu badanych zmiennych w zależności od przynależności jednostek badania do jednostek (grup) wyższych poziomów. Ponadto, dzięki podzieleniu badanej

populacji na grupy, możliwe jest wyjaśnienie części zmienności szacowanej cechy za pomocą charakterystyk z drugiego (lub wyższego) poziomu. W przypadku estymowania zmiennych określonych na populacji charakteryzującej się strukturą wielopoziomową, zastosowanie metodologii modelowania wielopoziomowego w istotnym stopniu poprawia precyzję szacunku, pod warunkiem, że zmienne te również charakteryzują się strukturą wielopoziomową (por. Goldstein, 2003; Hox 2010; Raudenbush, Bryk, 2002, Twisk, 2010).

W rozprawie wyszczególnione zostanie podejście klasyczne modelowania wielopoziomowego, w którym na drugim (i wyższych) poziomach występuje jedna zmienna grupująca. Podejście takie pozwala na szacowanie zmiennych dotyczących zbiorowości gmin. Zastosowanie metodologii modelowania wielopoziomowego pozwoli na uwzględnienie możliwości wpływu przynależności do jednostki drugiego (lub wyższego) poziomu, zarówno na wartość zmiennej objaśnianej określonej dla jednostek poziomu pierwszego, jak i rodzaj zależności pomiędzy szacowaną zmienną a zmiennymi objaśniającymi z pierwszego poziomu. Dodatkowo objaśnia się zmienność szacowanej cechy na poziomie pierwszym za pomocą zmiennych określonych na poziomach wyższych. Tak więc, przyjmuje się, że współczynniki kierunkowe dla zmiennych objaśniających z pierwszego poziomu nie tylko różnią się zależnie od przynależności do jednostek z wyższych poziomów, ale są objaśniane przez zmienne z wyższych poziomów (por. Goldstein, 2003; Hox, 2010; Raudenbush, Bryk, 2002; Twisk, 2010; Gruchociak, 2012a; Gruchociak 2012b).

W podrozdziale 3.3 opisano model wielopoziomowy z dwoma kryteriami grupowania (por. Biecek, 2011). Korzyści płynące z zastosowania takiego podejścia są podobne, jak dla klasycznego modelu wielopoziomowego, jednak dzięki uwzględnieniu dwóch niezależnych podziałów jednostek badania na grupy, możliwe jest na przykład wykorzystanie zmiennych objaśniających z wyższych poziomów odpowiadających obu podziałom. Model taki znajdzie zastosowanie podczas szacowania, w kolejnym rozdziale, przepływów pomiędzy parami gmin. W podrozdziale 3.4 opisane zostaną także kryteria oceny dopasowania modelu.

Rozdział czwarty ma charakter empiryczny. Przy pomocy metodologii modelowania wielopoziomowego opisanej w rozdziale 3 przeprowadzono szacunki macierzy przepływów ludności związanych z zatrudnieniem aktualnej dla roku 2011. Oszacowania

przeprowadzono przy założeniu, że relacje pomiędzy przepływami ludności związanymi z zatrudnieniem oraz czynnikami je determinującymi były w roku 2011 takie, jak w roku 2006. Znajomość estymowanej w tym rozdziale macierzy migracji niezbędna jest do przeprowadzenia delimitacji lokalnych rynków pracy w Polsce aktualnej dla roku 2011.

W **rozdziale piątym** przeprowadzono delimitację lokalnych rynków pracy w Polsce. W tym celu zastosowano dwa algorytmy zaczerpnięte z literatury (por. Coombes, Green, Openshaw, 1986; Kristensen, 1998) oraz samodzielnie opracowane podejście taksonomiczne. Najpierw delimitację przeprowadzono z wykorzystaniem danych z badania przepływów ludności związanych z zatrudnieniem. W wyniku otrzymano lokalne rynki pracy aktualne dla roku 2006. Następnie podjęto próbę delimitacji lokalnych rynków pracy w Polsce na podstawie oszacowanej w rozdziale czwartym macierzy przepływów ludności związanej z zatrudnieniem, aktualnej dla roku 2011. Przedstawiono także porównanie wyników otrzymanych za pomocą różnych algorytmów delimitacji oraz przy wykorzystaniu różnych danych.

W zakończeniu rozprawy doktorskiej omówiono uzyskane rezultaty. W szczególności przedstawiono wyniki weryfikacji sformułowanych we wstępie hipotez. Określono także stopień realizacji zamierzonych celów rozprawy. W wyniku dyskusji problemów rozwiązywanych w pracy, wskazane zostały kierunki dalszych badań.

W aneksie pierwszym omówiono problemy obliczeniowe powstałe podczas szacowania macierzy przepływów związanych z zatrudnieniem aktualnej dla roku 2011. Ponadto dołączono kod autorskiego algorytmu służącego do delimitacji lokalnych rynków pracy oprogramowany w R.

W aneksie drugim wyszczególniono skład wyznaczonych w rozprawie lokalnych rynków pracy w Polsce .

Autorka chciałaby bardzo podziękować Promotorowi, Pani prof. dr hab. Elżbiecie Gołacie, za inspirację, cierpliwość oraz wsparcie merytoryczne okazane w trakcie przygotowywania niniejszej rozprawy doktorskiej.

Rozdział 1. Wprowadzenie do problematyki lokalnego rynku pracy

1.1. Definicje lokalnego rynku pracy

W teorii ekonomii pojęciem rynku określa się miejsce, gdzie spotykają się podaż i popyt oraz kształtuje się cena towaru (por. Jarmołowicz (red.), 2008). Jeżeli towarem, którym handluje się na tak zdefiniowanym rynku jest praca, mamy do czynienia z rynkiem pracy. Praca może być rozumiana jako usługa, to znaczy fakt świadczenia pracy, lub jako potencjał pracy, czyli inaczej kapitał ludzki. Jest ona towarem nietypowym, ponieważ po zakupie” przez pracodawcę nie staje się jego własnością (por. Unolt, 1999). Co więcej towar, czyli praca jest nierozzerwalnie związany ze stroną podażową, czyli osobami deklarującymi chęć wykonywania pracy. W tej sytuacji również strona sprzedająca pracę może mieć zdecydowane preferencje określające komu praca ta zostanie sprzedana (por. Góra, Sztanderska, 2006).

W klasycznej ekonomii rynek nie jest umiejscowiony w przestrzeni. Jednak w przypadku rynku pracy jego umiejscowienie oraz przestrzenny zakres oddziaływania wydają się być niezmiernie istotne (por. Gołata, 2004). Trudno bowiem zakładać, że dla osoby szukającej pracy nie ma znaczenia, gdzie miałaby podjąć zatrudnienie oraz, że dostępność siły roboczej jest wszędzie taka sama.

Rynek pracy jest bardzo zróżnicowany pod wieloma względami, nie tylko terytorialnym. Nie należy lekceważyć faktu, że „towar” którym handluje się na rynku pracy, czyli praca, jest bardzo zróżnicowany. Na rynku pracy, nie każde stanowisko jest odpowiednie dla każdej poszukującej pracy osoby. Oczywiście istnieje możliwość przekwalifikowania się pracowników spowodowana zmianą sytuacji na rynku pracy. Jednak wiąże się to z pewnym wysiłkiem oraz często wysokimi kosztami ze strony osób poszukujących pracy i zazwyczaj podejmowane jest dopiero w ostateczności. Ponadto o ile w części przypadków, dotyczących szczególnie stanowisk nie wymagających zbyt wyrafinowanych kwalifikacji, przekwalifikowanie się jest stosunkowo proste, np. dzięki ukończeniu kursów na operatorów różnych maszyn itp., to w innych sytuacjach jest teoretycznie możliwe, jednak mocno utrudnione. Zmiana kwalifikacji oznacza często konieczność pogłębienia swojego wykształcenia, czy zdobycia nowych umiejętności zupełnie od podstaw, co w pewnych sytuacjach jest nawet niemożliwe, np. ze względu

na stan zdrowia czy kondycję fizyczną (por. Góra, Sztanderska, 2006). Specyfika ta uwidacznia się w szczególności w przypadku kategorii specjalistów, takich jak dziennikarze, politycy, muzycy, rzeźbiarze, czy śpiewacy operowi, którzy w celu podjęcia zatrudnienia często zmuszeni są podejmować dojazdy do pracy na dużych odległościach (por. Gołata, 2004). Zatem przyjęcie ograniczeń przestrzennych byłoby inne dla lokalnego rynku pracy robotników fizycznych oraz muzyków, przy czym drugi z nich byłby znacznie większy niż pierwszy. Delimitacja lokalnych rynków pracy dla każdej grupy zawodowej osobno byłaby zapewne interesująca dla osób do niej należących (tak pracowników jak i pracodawców), jednak z punktu widzenia władz lokalnych trudno byłoby rozważać każdą grupę oddzielnie. Z punktu widzenia realizacji zadań strategicznych związanych z rozwojem gospodarczym, przedmiotem zainteresowania władz samorządowych, byłyby raczej lokalne rynki pracy wyznaczone dla wszystkich osób pracujących łącznie. Takie podejście stosowane jest w literaturze (por. Casado-Diaz, Coombes, 2011) i przyjęte zostało również w prezentowanej pracy. Tak więc za lokalny rynek pracy przyjęto możliwie najmniejszy obszar na terenie którego równoważą się popyt i podaż na pracę, bez wyróżniania różnych jej rodzajów.

Administracyjne granice bardzo często nie odpowiadają rzeczywistej strukturze lokalnych rynków pracy (por. Casado-Diaz, Coombes, 2011; Prodromidis, 2010; Szumlicz, 1993). W wielu krajach, między innymi w Polsce, kształt granic administracyjnych w dużym stopniu wynika z tradycji i topografii terenu, co nie czyni wyznaczonych jednostek administracyjnych odpowiednimi obszarami do analizy relacji pomiędzy popytem i podażą pracy¹. Z drugiej strony w Polsce, ze względu na brak danych o dojazdach do pracy do niedawna nie było możliwości dokładnego wyznaczenia lokalnych rynków pracy, dlatego w wielu pracach przyjmowano powiat jako lokalny rynek pracy (por. Gołata, 2004; Góra, Sztanderska, 2006; Kryńska, 2010; Misiak, Tokarski, 2012). Warto przy tym jednak podkreślić, że autorzy prac w których przyjmowano takie podejście, często podkreślali jego niedoskonałość i wskazywali, że właściwsza byłaby dezagregacja przestrzenna faktycznie odzwierciedlająca lokalny charakter relacji na rynku pracy (por. Gołata, 2004; Góra, Sztanderska, 2006).

¹ Warto zaznaczyć, że również w Polsce, w roku 1993, przy okazji dyskusji o powiatowej strukturze terytorialnej, wysuwano postulat, aby budowa struktury powiatowej odpowiadała sieci lokalnych rynków pracy (por. Szumlicz, 1993)

W ostatnich dekadach na świecie podjętych zostało wiele prób zdefiniowania lokalnego rynku pracy. Jedną z pierwszych definicji sformułowaną została w latach czterdziestych przez komisję do spraw planowania siły roboczej podczas drugiej wojny światowej w USA (The War Manpower Commission). W celu obiektywnej identyfikacji jednostek administracyjnych, jako lokalny rynek pracy określono te jednostki „w ramach których osoby pracujące mogą zmienić pracę bez konieczności zmiany miejsca zamieszkania”. Jest to nadal podstawa definicji przyjmowanej przez Urząd Pracy w USA, zgodnie z którą lokalny rynek pracy to „ekonomicznie zintegrowany obszar geograficzny, w ramach którego osoby mogą zamieszkać i znaleźć zatrudnienie w możliwej do przyjęcia odległości oraz którego mieszkańcy mogą zmienić zatrudnienie nie zmieniając miejsca zamieszkania”

(por. US Department of Labor, 2003). Podobną definicję przyjmują C.M.Tolbert i M.Sizer (1996), określając lokalny rynek pracy jako „obszar obejmujący miejsca zamieszkania oraz miejsca pracy lokalnej populacji”. Z kolei R.I.Buntig (1962)² pod pojęciem lokalnego rynku pracy rozumie „możliwie najmniejszy obszar, który równocześnie: (a) zawiera miejsca zamieszkania ludności uważanej przez znajdujące się na jego terenie firmy za ich potencjalną siłę roboczą oraz (b) zawiera w swoich granicach większość firm, które zamieszkująca go ludność uważa za ewentualne potencjalne miejsca pracy”.

J.F.B.Goodman (1970), określając jak definiować lokalne rynki pracy, ustalił dwa postulaty możliwe do sprawdzenia w praktyce: (a) granice lokalnych rynków pracy powinny być możliwie rzadko przekraczane w ramach codziennych dojazdów do pracy oraz (b) w ramach całego lokalnego rynku pracy powinno występować duże nasilenie codziennych dojazdów do pracy, w wyniku czego danego obszaru nie da się już podzielić na mniejsze rynki pracy. Te dwa postulaty powinny być podstawą każdej próby wyodrębnienia lokalnych rynków pracy.

W literaturze polskiej T.Szumlicz (1993) wskazuje, że lokalny rynek pracy powinien odpowiadać zakresem zasięgowi terytorialnemu wzajemnego oddziaływania podaży i popytu na siłę roboczą. Ponadto ten sam autor wskazuje, że kryterium delimitacji

² Buntig R.I. (1962), *Employer Orientation in Local Labour Markets*, Chapel Hill, University of Carolina Press, cyt. za Casado-Diaz, Coombes, 2011.

powinny być izochrony dogodnych dojazdów do pracy³ (obszary obiektywnych rynków pracy)

albo wyznaczone powinny być w sposób administracyjny (teren działalności rejonowej agencji pośrednictwa pracy). Z kolei M.Góra i U.Sztanderska (2006) definiują lokalny rynek pracy jako obszar, w którego obrębie czas dojazdu do pracy jest tak krótki, że nie stanowi istotnej bariery zatrudnienia, jednak ze względu na brak danych o dojazdach do pracy obszar ten utożsamiają z powiatem .

W niniejszej pracy przyjęto definicję zgodną z przepisami szkolenia zawodowego *Job Trainning Partnership Act*, według której lokalny rynek pracy określany jest jako ekonomicznie zintegrowany obszar geograficzny, w ramach którego zamieszkująca ludność może znaleźć zatrudnienie bądź je zmienić bez konieczności zmiany miejsca zamieszkania (por. Gołata 2004).

1.2. Wprowadzenie do delimitacji lokalnych rynków pracy

1.2.1. Cel i uzasadnienie potrzeby delimitacji lokalnych rynków pracy

Cele i sposoby wykorzystania delimitacji lokalnych rynków pracy zostały przedstawione między innymi w publikacji Organizacji Współpracy Ekonomicznej i Rozwoju, do której Polska należy od roku 1996 (por. OECD, 2002). Praca ta powstała na podstawie ankiety przeprowadzonej wśród 22 krajów członkowskich. Wykorzystano w niej pojęcie regionu funkcjonalnego rozumianego, jako jednostka terytorialna, wynikająca ze struktury relacji społecznych i ekonomicznych, której granice nie odzwierciedlają szczegółów geograficznych lub wydarzeń historycznych. Jest to zatem funkcjonalny podobszar terytorium. Najbardziej typowym pojęciem, wykorzystywanym przy definiowaniu regionu funkcjonalnego jest właśnie pojęcie rynku pracy. Spośród 22 krajów członkowskich OECD, które odpowiedziały na ankietę, tylko 5 (Japonia, Meksyk, Korea, Turcja i Hiszpania) nie definiowało regionów funkcjonalnych w postaci rynków pracy. Można, zatem stwierdzić, że większość krajów członkowskich OECD definiuje i przeprowadza delimitację regionów funkcjonalnych w postaci lokalnych rynków pracy. Należy przy tym zauważyć, że delimitacja lokalnych rynków pracy w większości krajów oparta jest na tej samej zasadzie, dotyczącej warunków związanych z dojazdami do pracy.

³ Izochrona – linia na mapie łącząca punkty, w których występuje to samo zjawisko w tym samym czasie. Innymi słowy izochrony wyznaczają obszary o jednakowej osiągalności czasowej.

Zupełnie inaczej sytuacja wygląda w Polsce, która jako lokalne rynki pracy zadeklarowała po prostu powiaty (por. OECD, 2002). W większości krajów lokalne rynki pracy są delimitowane poprzez dołączanie całych gmin. Dzięki zastosowaniu tego rozwiązania, są zazwyczaj w pełni kompatybilne z jednostkami podziału administracyjnego, dla których gromadzone są dane statystyczne. W rezultacie istotne statystyki i wskaźniki, wykorzystywane w analizie terytorialnej są dostępne także na poziomie lokalnych rynków pracy.

Większość krajów, które określały lokalne rynki pracy, wykorzystuje je jako podstawę w badaniach struktury rynków pracy, analizach terytorialnego zróżnicowania rozwoju społeczno-ekonomicznego, szacowaniu regionalnych dysproporcji. W wielu krajach, takich jak Finlandia, Francja, Niemcy czy Włochy, właśnie lokalne rynki pracy wykorzystywane są do zidentyfikowania obszarów, które otrzymują wsparcie z funduszy krajowych lub unijnych (por. OECD, 2002).

W czterech krajach członkowskich (Austria, Kanada, Dania i Szwajcaria) lokalne rynki pracy były wykorzystywane, bezpośrednio lub częściowo, do wdrożenia pewnych kierunków działań, odnoszących się głównie do rynków pracy i transportu (por. OECD, 2002). Jako przykładowe wykorzystania lokalnych rynków pracy w polityce regionalnej podać można między innymi kontrolę przeludnienia obszarów czy redukcję kosztów osobistych i społecznych czasochłonnych dojazdów do pracy (por. METREX, 2010; Cervero, 1995; Van der Laan, Schalke, 2001; Vliegen, 2003). Jeden z przykładów praktycznego wykorzystania wyznaczonych lokalnych rynków pracy pochodzi z Wielkiej Brytanii, gdzie już w połowie ubiegłego wieku wyznaczono lokalną stopę bezrobocia. Departament Handlu i Przemysłu Wielkiej Brytanii użył lokalnych rynków pracy także, jako jednostek, z których utworzono regiony NUTS 2⁴ (por. Coombes, 2002).

Australijskie Biuro Transportu i Ekonomii Lokalnej (Bureau of Transport and Regional Economics) w przekroju lokalnych rynków pracy analizuje regionalną strukturę przemysłu. W Danii wyniki przeprowadzonej delimitacji stały się podstawą do dyskusji o reorganizacji samorządności lokalnej (por. Andersen, 2002). Wyniki przeprowadzonych analiz mogą zostać użyte również w pracach naukowych. Przykładem takiego

⁴ **NUTS** (*Nomenclature of Units for Territorial Statistics*) Klasyfikacja Jednostek Terytorialnych do Celów Statystycznych - standard rozwinięty w Unii Europejskiej na potrzeby identyfikowania statystycznych jednostek terytorialnych (NUTS 2 – jednostki o liczebności populacji 0.8 - 3 mln, w Polsce określone jako województwa)

wykorzystania może być analiza migracji w Nowej Zelandii przeprowadzona w przekroju lokalnych rynków pracy (por. Papps, Newell, 2002).

W literaturze polskiej również podkreśla się, że w celu prowadzenia jakiegokolwiek gospodarki lokalnej konieczna jest znajomość społecznego systemu terytorialnego danej przestrzeni, którego jednostki charakteryzują się cechami lokalnych rynków pracy (por. Parysek, 1996; Parysek, 2001). Również P.Korcelli, A.Potrykowska i D.Bodzak w swojej pracy zaznaczają, że badania w zakresie dojazdów do pracy mają istotne znaczenie dla rozbudowy teorii systemów osadniczych, dokładniejszego poznania przemian wewnętrznej struktury miast oraz rozwoju planistycznych modeli osadnictwa (por. Korcelli, Potrykowska, Bodzak, 1981). Z kolei P.Śleszyński (2012) wskazuje, że rozpoznanie powiązań przestrzennych w zakresie miejsca zamieszkania i pracy stanowi podstawę do bardziej racjonalnego zagospodarowania przestrzennego, w tym polityki transportowej i ogólnie organizacji transportu, zwłaszcza publicznego. Wyniki delimitacji lokalnych rynków pracy będą mogły w warunkach gospodarki rynkowej stanowić także wsparcie dla pracodawców i lokalnych samorządów w decyzjach dotyczących szeroko pojętej polityki zatrudnienia.

1.2.2. Koncepcje delimitacji lokalnych rynków pracy

W podrozdziale 1.1 stwierdzono, że większość badaczy traktuje jako lokalny rynek pracy obszar, w ramach którego zamieszkująca ludność może podjąć pracę bez konieczności zmiany miejsca zamieszkania. Tak więc jest to obszar, na terenie którego mieszkańcy są skłonni dojeżdżać do pracy. W konsekwencji przyjęcia powyższej definicji konieczne jest określenie sposobu badania gotowości ludności do podjęcia dojazdów do pracy. Ważnym byłoby uwzględnienie kierunków i odległości, struktury demograficznej i społecznej dojeżdżających do pracy, środków transportu, czasu itd.

Można przyjąć, że gotowość ta zdeterminowana będzie dwiema grupami charakterystyk: dotyczących dostępności i atrakcyjności pracy oraz określających łatwość codziennego przejazdu (por. Ohme, 1988, Gruchociak, 2012d). Do pierwszej grupy zaliczyć można liczbę przeprowadzanych rekrutacji oraz na jakie stanowiska poszukiwano kandydatów, ile osób poszukuje pracy na terenie danego obszaru oraz wszelkie charakterystyki warunków pracy, takie jak wynagrodzenie czy wymiar zatrudnienia, ale również wiele niemierzalnych zmiennych, określających komfort pracy. Do drugiej grupy zaliczyć

należy przede wszystkim odległość, która mierzona może być w kilometrach, zarówno w linii prostej jak i uwzględniając drogę samochodową czy kolejową lub biorąc pod uwagę czas czy koszty dojazdu (cena biletu, koszt benzyny itp.). Tak więc, aby podjąć się delimitacji lokalnych rynków pracy a priori, należałoby określić grupę czynników charakteryzujących rynek pracy oraz łatwość dojazdu. Wiele z powyżej opisanych czynników jest jednak niemierzalnych a informacje na temat większości z nich są bardzo trudno dostępne (por. Casado-Diaz, Coombes, 2011). Ponadto trudno przewidzieć dokładnie, który czynnik i w jakim stopniu zmotywuje konkretną osobę o indywidualnych preferencjach do podjęcia pracy w danym miejscu. W związku z tym zdecydowanie bardziej uzasadnione wydaje się podejście a posteriori, czyli ocena skłonności ludności do podejmowania pracy w danym miejscu na podstawie informacji o miejscu zamieszkania oraz zatrudnienia poszczególnych osób. W literaturze przedmiotu, jako podstawę delimitacji lokalnych rynków pracy, wskazuje się zazwyczaj kryterium dojazdów do pracy (por. Klassen, Paelinck, Wagenaar, 1982).

Pomimo powszechnego przeprowadzania delimitacji lokalnych rynków pracy, wśród badaczy nie ma zgodności odnośnie określenia najlepszej metody (por. Casado-Diaz, Coombes, 2011; Landre, Hakansson, 2013). Jako cechę wspólną wykorzystywanych algorytmów wskazać można przeprowadzanie delimitacji w oparciu o natężenie codziennych dojazdów związanych z zatrudnieniem. Ponadto przyjmuje się, że algorytm powinien działać na zasadzie podziału jednostek bazowych na rozłączne zestawy, pokrywające cały badany obszar (zazwyczaj kraj). W literaturze przedmiotu znaleźć można także opinie, że traktowanie lokalnych rynków pracy jako odizolowanych jednostek utrudnia wykrycie powiązań funkcjonalnych, które występują zarówno wewnątrz jak i pomiędzy nimi (por. Kallioras, Kandylis, Kromyidakis, Pantazis, 2011). W większości prac z tego zakresu, ze względów praktycznych przyjmuje się jednak, że wyznaczone lokalne rynki pracy powinny być rozłączne (por. Casado-Diaz, Coombes, 2011).

Rozłączność lokalnych rynków pracy wymagana jest przez badaczy w wielu krajach (por. Coombes, 2002; Coombes, Bond, 2008; Andersen, 2002; Bamber, Walter, 2009; Brown, Holmes, 1971; Casado-Diaz, 2000; Coombes, Green, Openshaw, 1986; Coombes, Wymer, Charlton, Bailey, Stonehouse, Openshaw, 1997; Fusco, Caglioni, 2011; Kristensen, 1998; Lagnerö, 2003; Newell, Papps, 2001; Nystuen, Dacey, 1968; Papps, Newell, 2002; Sforzi, 1987; Tolbert, Sizer, 1996; Van der Haegen, Pattyn, 1980;

Prodromidis, 2008; Prodromidis, 2010; Kallioras, Kandyliis, Kromyidakis, Pantazis, 2011; Stachowski 1993, Gruchociak 2012d). Opisywane przez nich algorytmy delimitacji lokalnych rynków pracy działają na zasadzie łączenia skończonej liczby jednostek bazowych w pokrywające całą zbiorowość grupy. Jednocześnie grupowanie to spełnia wymóg pełności i rozłączności. Łączenie jednostek bazowych przeprowadzane jest na podstawie informacji o codziennych dojazdach do pracy pomiędzy nimi. Jako jednostki bazowe przyjmuje się możliwie najmniejsze jednostki terytorialne, które są rozłączne i jednocześnie pokrywają cały obszar badanego kraju. Dokładność przeprowadzonej delimitacji jest oczywiście tym większa, im więcej wyodrębniono jednostek bazowych na badanym obszarze (w kraju). Najczęściej jedyną wymaganą informacją wejściową jest macierz przepływów związanych z zatrudnieniem, nazywana również macierzą migracji związanych z zatrudnieniem⁵. Jest to macierz kwadratowa o wymiarze równym liczbie jednostek bazowych, zawierająca na przecięciu *i-tego* wiersza oraz *j-tej* kolumny informacje o liczbie osób mieszkających na terenie *i-tej* jednostki bazowej i jednocześnie zatrudnionych na terenie *j-tej* jednostki bazowej. W szczególności na diagonalu macierzy znajdują się liczby ludności pracującej w obszarze swojego zamieszkania.

1.2.3. Pożądane cechy lokalnego rynku pracy

Pomimo braku zgodności w kwestii algorytmu, jaki należy stosować do delimitacji, w literaturze znaleźć można wskazania dotyczące cech, którymi powinny charakteryzować się wyznaczone lokalne rynki pracy (por. tab.1.1). Postulaty te zostały sformułowane przez Eurostat i są zalecane do stosowania w delimitacji lokalnych rynków pracy w krajach europejskich (por. Eurostat, 1992). Szczegółowe ich omówienie zawarto także w pracy przeglądowej J. M. Casado-Diaz i M.Coombes'a (por. Casado-Diaz, Coombes, 2011).

Dwa pierwsze, z wymienionych w tab.1.1 postulatów, uznawane są za nadrzędne i odnoszą się do celu, w jakim delimitacja jest przeprowadzana. Tak więc po pierwsze, wyznaczone lokalne rynki pracy powinny mieć zakres odpowiadający celom

⁵ W dalszej części pracy macierz przepływów związanych z zatrudnieniem będzie nazywana w skrócie macierzą migracji. Ponieważ w pracy omawiane będą wyłącznie migracje codzienne związane z zatrudnieniem, nazywane w literaturze również migracjami wahadłowymi, uproszczenie to nie powinno prowadzić do nieporozumień.

politycznym, umożliwiającą prowadzenie spójnej polityki rozwoju regionalnego przez władze lokalne.

Po drugie każdy z wyznaczonych obszarów powinien charakteryzować się cechami właściwymi dla samodzielnego rynku pracy. Tak, więc delimitacja powinna być przeprowadzona na podstawie sytuacji faktycznej (opisanej najczęściej przy pomocy danych o dojazdach do pracy) a nie na podstawie z góry narzuconych warunków, takich jak wybrane a priori ośrodki centralne czy maksymalna rozpiętość lokalnego rynku pracy. Ponadto wskazuje się, że jeżeli dwa pierwsze warunki są spełnione dla dwóch różnych podziałów, za lepszy z nich uznaje się ten, w przypadku którego wyodrębniono więcej lokalnych rynków pracy. Warunek ten często traktowany jest, jako dziesiąta zasada delimitacji lokalnych rynków pracy (por. Casado_Diaz, Coombes, 2011).

Tabela 1.1. Zasady wyznaczania lokalnych rynków pracy

Zasada		Praktyka
cele		
1	odpowiedniość	obszar statystycznie określony, odpowiedni dla celów politycznych
2	poprawność	każdy wyznaczony obszar powinien charakteryzować się cechami lokalnego rynku pracy
ograniczenia		
3	rozłączność i pokrycie całości	każda jednostka bazowa powinna zostać przydzielona do dokładnie jednego lokalnego rynku pracy
4	spójność	każdy wyznaczony lokalny rynek pracy powinien być spójny terytorialnie
kryteria		
5	autonomia	nasilenie dojazdów do pracy wewnątrz lokalnych rynków pracy powinno być jak największe a pomiędzy nimi jak najmniejsze
6	jednorodność	rozmiar lokalnych rynków powinien być jak najmniejszy
7	koherencja	granice powinny być jednoznacznie wyznaczone
8	zgodność	preferowane są granice oparte na podziale administracyjnym
podsumowanie		
9	elastyczność	metoda powinna sprawdzać się w przypadku różnych obszarów

Źródło: Casado-Diaz, Coombes, 2011

Trzecia zasada nakazuje, aby każda z jednostek bazowych została przydzielona dokładnie do jednego lokalnego rynku pracy. W ten sposób wyznaczone obszary są rozłączne i pokrywają całą przestrzeń. Zrozumiałe jest jednak, że mieszkańcy jednostek bazowych leżących na granicach lokalnych rynków pracy mogą podejmować pracę również na terenie sąsiednich jednostek. Sytuacja taka mogłaby zostać uznana za przesłankę do pominięcia zasady trzeciej (por. Kallioras, Kandylis, Kromydakis, Pantazis, 2011), jednak uwzględnienie tej zasady ma wiele zalet przy praktycznym wykorzystaniu wyników przeprowadzonej delimitacji. Ponadto warunek ten nie jest wymagany przy konstrukcji obszarów metropolitalnych, co jest podstawową różnicą pomiędzy ich wyodrębnianiem a delimitacją lokalnych rynków pracy.

Z kolei według czwartego postulatu, każdy lokalny rynek pracy powinien być spójny terytorialnie. Wskazane jest jednak, żeby spójność ta wynikała nie tyle z narzucenia zasady łączenia tylko jednostek bazowych mających wspólną granicę, a ze specyfiki kryteriów łączenia poszczególnych jednostek bazowych. W związku z powyższym w literaturze dopuszcza się niewielkie odstępstwa od powyższej zasady, które uzasadniać można na przykład większą łatwością dojazdu do dalszej jednostki bazowej, niż do jednostek sąsiednich. W wielu metodach spójność terytorialna uzyskiwana jest w wyniku narzucenia warunku wspólnej granicy. Są to między innymi niektóre z metod numerycznej taksonomii, regionalizacji niefunkcjonalnej, w tym między innymi popularna w Polsce metoda J.D.Nystuena i M.F.Daceya⁶ (1961) oraz program automatycznego podziału na strefy zaproponowany przez W.S.Openshawa i L.Rao (1995). W przypadku danych o dojazdach do pracy najsilniejsze interakcje odnotowywane są zazwyczaj między pobliskimi jednostkami administracyjnymi. W konsekwencji ich uwzględniania, wynikiem delimitacji są obszary spójne z natury (bez nakładania dodatkowego warunku), w większym stopniu niż w przypadku innych danych. W ten sposób możliwe jest uniknięcie nakładania dodatkowego warunku wspólnej granicy, podczas gdy wyznaczane są lokalne rynki pracy, w których prawie wszystkie jednostki bazowe spełniają warunek bezpośredniego sąsiedztwa.

Cztery kolejne postulaty traktowane mogą być jako wskazówki do określania kryteriów delimitacji. Tak, więc piąta zasada dotyczy autonomii wyznaczanych lokalnych rynków pracy. Oznacza to, że nasilenie dojazdów do pracy pomiędzy wyznaczanymi obszarami

⁶ Metoda Nstuyena-Daceya jest metodą grupowania hierarchcznego na podstawie relacji największego przepływu (por. Stachowski, 1993; Śleszyński, Czapiewski, 2011).

powinno być jak najmniejsze, zaś nasilenie dojazdów do pracy wewnątrz nich powinno być jak największe.

Z kolei szósty warunek zaleca jednorodność wyznaczanych obszarów. W celu spełnienia tego warunku rozmiary lokalnych rynków pracy powinny być jak najmniejsze, np. ich powierzchnia powinna zawierać się w określonych granicach.

Zasada siódma określa koherencję wyznaczonych rynków pracy. Zgodnie z nią granice powinny być możliwie jednoznacznie wyznaczone.

Zgodnie z warunkiem ósmym, pożądaną cechą delimitacji, jest w miarę możliwości zgodność z granicami opartymi na podziale administracyjnym. W związku z tym, jako jednostki bazowe preferowane są możliwie najmniejsze jednostki podziału administracyjnego.

Według postulatu dziewiątego, metoda delimitacji powinna być elastyczna, co oznacza, że powinna sprawdzać się w przypadku różnych obszarów.

W przypadku aż dziewięciu różnych zasad powstaje pole do dyskusji nad hierarchią ich ważności. W swojej pracy J.M.Casado-Diaz i M.Coombes (2011) wskazują, że decydujące znaczenie powinny mieć cztery pierwsze zasady, przy jednoczesnej maksymalizacji liczby wyznaczonych lokalnych rynków pracy.

1.3. Dojazdy do pracy w Polsce

1.3.1. Charakterystyka źródeł danych

Dostępność i gwarantowana okresowość⁷ danych umożliwiających rekonstrukcję macierzy przepływów związanych z zatrudnieniem, pozwalają na regularne przeprowadzanie delimitacji lokalnych rynków pracy (Coombes et al., 1982; Casado-Diaz, Coombes, 2011). Niestety, w warunkach polskich, ze względu na brak odpowiednich źródeł danych statystycznych, przeprowadzenie pełnej delimitacji lokalnych rynków pracy w Polsce było do niedawna niemożliwe.

⁷ W większości krajów z Organisation for Economic Co-operation and Development (OECD) dane dotyczące dojazdów do pracy publikowane są co najmniej co dziesięć lat.

W okresie bezpośrednio powojennym w Polsce regularnie przeprowadzane były Spisy Kadrowe⁸ dostarczające informacji o migracjach wahadłowych. Główny Urząd Statystyczny przeprowadził siedem takich badań⁹, ostatnie w 1983 roku. Informacje dotyczące dojazdów do pracy zbierane były również przy okazji spisu ludności w 1978, jednak dane te nie uwzględniały kierunków dojazdów. Po raz ostatni pytanie w tym zakresie pojawiło się w mikrosписie w roku 1995. Dotyczyło ono jedynie faktu dojeżdżania do pracy poza teren gminy (nie uwzględniając kierunku ani odległości). Podczas Spisu Ludności w 2002 roku tematyka ta została całkowicie pominięta. W całym okresie 20 lat transformacji gospodarczej i ogromnych przemian na rynku pracy nie były dostępne dane charakteryzujące mobilność przestrzenną pracowników. W tej sytuacji poszukiwano informacji na temat dojazdów w alternatywnych źródłach, w tym w dostępnych rejestrach administracyjnych.

Należy zaznaczyć, że pomimo braku danych na temat dojazdów do pracy, w Polsce podejmowane były próby ich oszacowania lub pozyskania informacji przy pomocy badań ankietowych. W swoich opracowaniach autorzy podkreślali jednak, że nie jest to podejście zadowalające (por. Swianiewicz, Klimska, 2005; Rosik, Stępiak, Wiśniewski, 2010).

W 2010 roku Departament Pracy Głównego Urzędu Statystycznego przeprowadził badanie pt. *Dojazdy do pracy w 2010 roku na podstawie BAEL*. Badanie to dostarczyło wielu cennych informacji o strukturze ludności dojeżdżającej do pracy, takich jak wiek, płeć czy wykształcenie, obszar zamieszkania (miasto/wieś), czas dojazdu i środek transportu. Jego wyniki nie pozwalają jednak na rekonstrukcję macierzy przepływów codziennych związanych z zatrudnieniem, ponieważ nie obejmują kierunków dojazdów oraz nie pozwalają na uogólnienie informacji na poziomie małych jednostek terytorialnych, co jak wskazano wyżej jest jednym z warunków delimitacji lokalnych rynków pracy.

Również w 2010 roku Ośrodek Statystyki Miast Urzędu Statystycznego w Poznaniu opublikował wyniki unikatowego badania przepływów ludności związanych z zatrudnieniem. Zawierały one informacje dotyczące roku 2006, które zostały pozyskane

⁸ Szerszą informację o tym źródle oraz badaniach przeprowadzonych na jego podstawie zamieszczono w podrozdziale 2.2.

⁹ Główny Urząd Statystyczny przeprowadził Spisy Kadrowe w latach: 1956, 1958, 1964, 1968, 1973, 1977 i 1988 (por. Gawryszewski, 2005, str. 70).

z zasobów rejestrów podatkowych Ministerstwa Finansów (por. Dojazdy do pracy w Polsce, red. Kruszka, 2010; Filas-Przybył, Stachowiak, 2010). Są to wyniki pierwszego, opublikowanego przez statystykę publiczną badania, dla którego informacje pozyskane zostały w wyniku integracji rejestrów administracyjnych¹⁰. Udostępnienie przez Ministerstwo Finansów danych rejestrów systemu podatkowego POLTAX, poprzedzone było kilkuletnimi pracami studialnymi w Urzędzie Statystycznym w Poznaniu. Są to pierwsze od 1988 roku dane dotyczące dojazdów do pracy. Ponadto, w roku 2011, wyniki badania dojazdów do pracy zostały uzupełnione poprzez udostępnienie tabeli przepływów pracowników najemnych, którzy dojeżdżają do pracy poza gminę swojego zamieszkania¹¹.

Dane z dotyczące dojazdów do pracy przedstawione są według stanu z dnia 31 grudnia 2006 roku. Badana zbiorowość to dojeżdżający do pracy pracownicy najemni, którzy mieszkają

w gminie innej niż ta, w której znajduje się ich miejsce pracy, i jednocześnie mają zwiększone koszty uzyskania przychodu z tytułu dojazdów do pracy (2,3 mln osób, co stanowi 24,7% ogółu pracowników najemnych). Jako jednostkę terytorialną przyjęto gminę z uwzględnieniem podziału na część miejską i wiejską. W wyniku przyjęcia takiego rozwiązania, otrzymano na terenie Polski, 3062 jednostek bazowych (por. *Ogólna charakterystyka badania przepływów ludności związanych z zatrudnieniem w 2006 r.*). Udostępnione dane dotyczą liczby osób podejmujących codzienne dojazdy do pracy dla każdej pary gmin, z podziałem na część miejską i wiejską w Polsce. W celu ochrony danych osobowych w opublikowanej tabeli przepływów pracowników najemnych, którzy dojeżdżają do pracy poza gminę swojego zamieszkania, uwzględniono wyłącznie dojazdy powyżej dziewięciu osób. W ten sposób, opublikowana została informacja na temat przepływów ludności związanych z zatrudnieniem (por. Filas-Przybył, Stachowiak, 2010). Dla każdego przepływu, w tabeli podane są takie informacje, jak kod terytorialny gminy zamieszkania podatnika, kod terytorialny gminy płatnika, nazwa gminy zamieszkania podatnika czy nazwa gminy płatnika.

Tak, więc w celu konstrukcji macierzy przepływów związanych z zatrudnieniem potrzebnej do delimitacji lokalnych rynków pracy w Polsce, utworzono macierz

¹⁰ Metody statystycznej integracji danych zostały szczegółowo opisane w języku polskim w pracy W. Roszki (2013).

¹¹ Na stronie internetowej Ośrodka Statystyki Miast Urzędu Statystycznego w Poznaniu w formie pliku Excel'a, http://www.stat.gov.pl/poznan/69_567_PLK_HTML.htm.

o wymiarze 3062x3062. Macierz ta zawiera 25375 niezerowych wartości poza diagonalą, zaś w pozostałych miejscach, z wyjątkiem diagonali przyjęto zera (pomimo, że mogą one zawierać dowolne liczby nie większe od 9) (por. Dojazdy do pracy w Polsce, red. Kruszką, 2010).

Na diagonalu macierzy przepływów powinny znaleźć się liczby osób jednocześnie mieszkających i pracujących w danej jednostce bazowej. Informacje z badania przepływów ludności związanych z zatrudnieniem dotyczą zbiorowości pracowników najemnych, którzy dojeżdżają do pracy poza gminę swojego zamieszkania, zatem nie dostarczają bezpośrednio informacji o liczbie ludności pracującej w gminie zamieszkania¹². Również inne dostępne źródła danych nie zawierają bezpośredniej informacji na ten temat. Aby wyznaczyć odpowiednie liczby pracujących w gminie zamieszkania, skorzystano z danych publikowanych w Banku Danych Lokalnych GUS w tabeli o nazwie: *Pracujący w głównym miejscu pracy*¹³. Informacje te pochodzą ze sprawozdawczości statystycznej i dostępne są dla poziomu gminy z uwzględnieniem podziału na miasto i wieś. Wyniki sprawozdawczości udostępniane są co roku. Informują one o liczbie osób zatrudnionych w danej gminie, wliczając w to również osoby mieszkające poza daną gminą. W związku z tym, wartości na diagonalu macierzy przepływów oszacowano jako różnice pomiędzy liczbą osób pracujących w danej jednostce bazowej i liczbą osób przyjeżdżających do niej do pracy. Dane o liczbie osób przyjeżdżających do pracy zaczerpnięto z wyżej omówionego *Badania przepływów ludności związanych z zatrudnieniem*. Warto zwrócić uwagę, że odjemna i odjemnik takiej różnicy pochodzą z różnych źródeł danych. Różne źródła danych, a także nieznacznie odmienne definicje osoby pracującej¹⁴, użyte do obliczenia tej różnicy mogły doprowadzić do niewielkich przekłamań. Odzwierciedleniem tej sytuacji może być

¹² Na podstawie danych uzyskanych z rejestrów POLTAX, w ramach badania przepływów ludności związanych z zatrudnieniem, oszacowano również liczby ludności pracującej w gminie zamieszkania. Jednak w wyniku późniejszego ograniczenia zbiorowości do pracowników najemnych, wyniki te nie zostały upublicznione. Warto jednak zauważyć, że w publikacji przygotowanej przez GUS autorzy posługują się informacją o pracujących w miejscu zamieszkania (por. Dojazdy do pracy w Polsce, red. Kruszką, 2010).

¹³ Badanie to realizowane jest przez GUS na formularzu Z-06 „Sprawozdanie o pracujących, wynagrodzeniach

i czasie pracy” oraz SP-3 „Sprawozdanie o działalności gospodarczej przedsiębiorstw”

¹⁴ Dane z tabeli *Pracujący w głównym miejscu pracy* BDL nie obejmują pracujących w jednostkach budżetowych działających w zakresie obrony narodowej i bezpieczeństwa publicznego, podmiotów gospodarczych o liczbie pracujących do 9 osób, SOF, duchownych pełniących obowiązki duszpasterskie oraz pracujących w gospodarstwach indywidualnych (SOF obejmuje osoby pracujące w fundacjach, stowarzyszeniach, partiach politycznych, organizacjach społecznych, organizacjach pracodawców, samorządu gospodarczego i zawodowego)

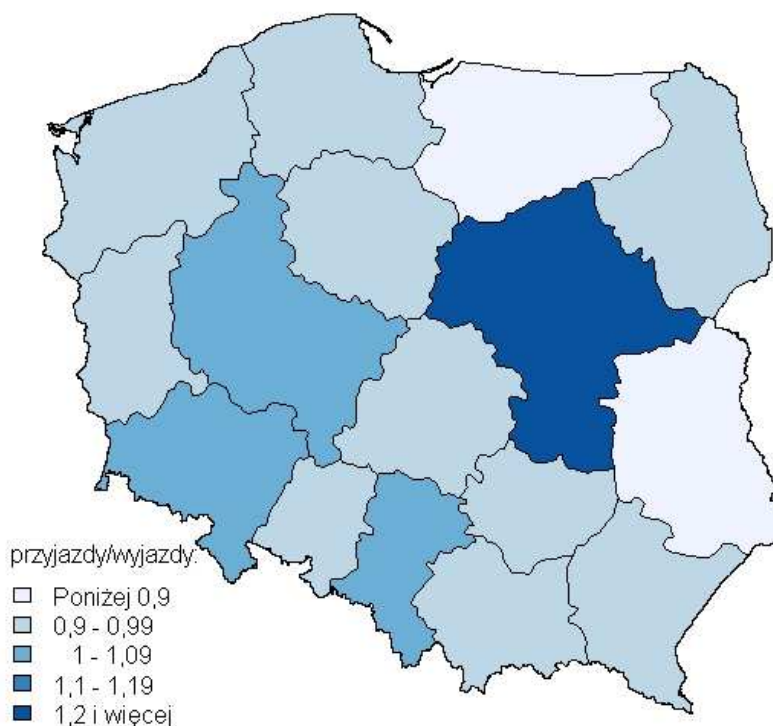
na przykład fakt, że dla 25 z 3062 gmin uzyskano na diagonalu macierzy przepływów niewielkie ujemne wartości. Dla dalszych rozważań wartości te zastąpiono zerami.

Udostępnione wyniki badania przepływów ludności związanych z zatrudnieniem umożliwiają więc przeprowadzenie pierwszej od wielu lat kompleksowej oraz niezależnej od subiektywnej opinii badacza delimitacji lokalnych rynków pracy w Polsce.

1.3.2. Struktura i przestrzenne zróżnicowanie dojazdów do pracy

Wyniki badania przepływów ludności związanych z zatrudnieniem wskazują, że dojeżdżający do pracy w Polsce w końcu 2006 roku stanowili 27,4% ogółu pracowników najemnych. Świadczy to niewątpliwie o powszechności zatrudnienia poza miejscem zamieszkania. Analizując zróżnicowanie terytorialne dojazdów do pracy w Polsce warto zwrócić również uwagę na skalę zjawiska wskazując kilka miast oraz województw, w których zyskuje ono szczególnie duże rozmiary. Przykładowo w Warszawie pracuje aż 31% wszystkich osób zatrudnionych na terenie województwa mazowieckiego, przy czym aż 51,7% osób zatrudnionych w stolicy mieszka poza jej granicami. Więcej niż co piąta osoba zatrudniona w województwie małopolskim pracuje w Krakowie. W konsekwencji 34,7% zatrudnionych na terenie tego miasta to osoby przyjeżdżające do pracy. W województwach podkarpackim, śląskim, małopolskim i wielkopolskim więcej niż co trzeci zatrudniony pracuje poza swoim miejscem zamieszkania.

Różnice pomiędzy poszczególnymi województwami nie są zbyt duże, zazwyczaj na jedną osobę wyjeżdżającą do pracy przyjeżdża do pracy również około jedna osoba. Wyjątek stanowi województwo mazowieckie, gdzie na jedną osobę wyjeżdżającą do pracy poza granice województwa przypada aż 1,3 osoby przyjeżdżającej do pracy. Odmienna sytuacja dotyczy województw lubelskiego i warmińsko-mazurskiego, gdzie na 1 osobę wyjeżdżającą przyjeżdża mniej niż 0,9 osoby. Należy jednak pamiętać, iż spowodowane jest to nieuwzględnieniem dojazdów wewnątrz województw.



Wykres 1.1. Stosunek przyjeżdżających do wyjeżdżających według województw w 2006r.

Źródło: Opracowania własne

Syntetyczną miarą ukazującą strukturę dojazdów do pracy jest relacja pomiędzy przyjeżdżającymi a wyjeżdżającymi do danych obszarów. Zróżnicowanie w przekroju województw widoczne jest na wykresie mapowym 1.1.

Różnice pomiędzy poszczególnymi województwami nie są zbyt duże, zazwyczaj na jedną osobę wyjeżdżającą do pracy przyjeżdża do pracy również około jedna osoba. Wyjątek stanowi województwo mazowieckie, gdzie na jedną osobę wyjeżdżającą do pracy poza granice województwa przypada aż 1,3 osoby przyjeżdżającej do pracy. Odmienna sytuacja dotyczy województw lubuskiego i warmińsko-mazurskiego, gdzie na 1 osobę wyjeżdżającą przyjeżdża mniej niż 0,9 osoby. Należy jednak pamiętać, iż wynika to z powodu nieuwzględnienia dojazdów wewnątrz województw.

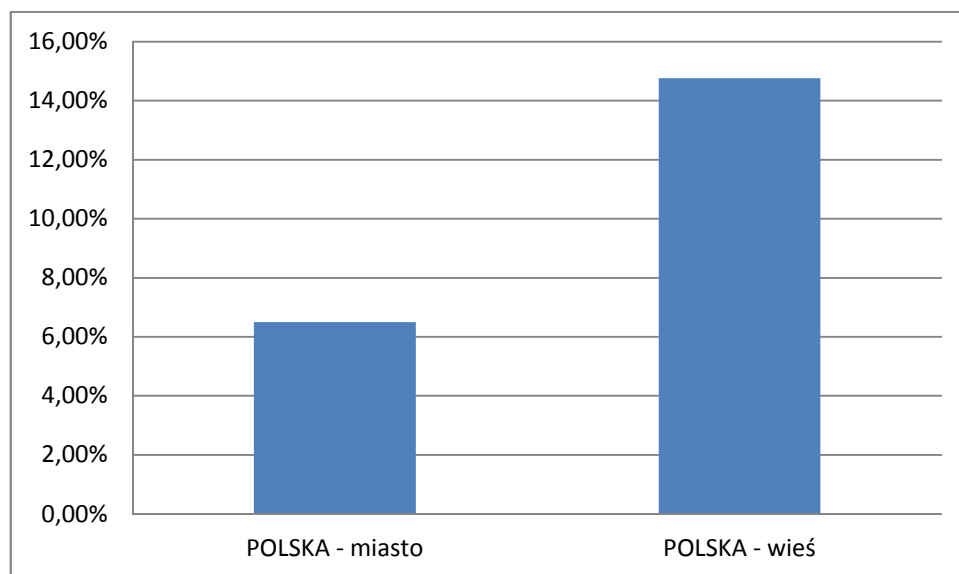
Znacznie większe zróżnicowanie można zaobserwować pomiędzy obszarami miejskimi oraz wiejskimi. Uwzględnienie takiego podziału ukazuje jak wyraźnie niezrównoważony jest rozwój części wiejskiej i miejskiej Polski. Dysproporcję tą obrazuje tabela 1:

Tabela 1.2. Wyjazdy i przyjazdy do pracy w przekroju miasto - wieś, Polska 2006 r.

Jednostka terytorialna	liczba osób wyjeżdżających do pracy	liczba osób przyjeżdżających do pracy	saldo przyjazdów i wyjazdów do pracy	liczba osób przyjeżdżających do pracy przypadająca na 1 osobę wyjeżdżającą do pracy	pracujący w głównym miejscu pracy ogółem	Ludność w wieku produkcyjnym (powyżej 17 lat) ogółem	liczba osób wyjeżdżających do pracy/ludność w wieku produkcyjnym
	2006	2006	2006	2006	2006	2006	2006
	[osoba]	[osoba]	[osoba]	[osoba]	[osoba]	[osoba]	[%]
miasto	1 002 720	1 844 666	841 946	1,84	6 723 138	15 425 442	6,50%
wieś	1 337 023	495 077	-841 946	0,37	1 315 007	9 056 228	14,76%

Źródło: Opracowanie własne na podstawie Banku Danych Regionalnych GUS

Około dwukrotnie więcej mieszkańców obszarów wiejskich niż miejskich szuka zatrudnienia poza obszarem zamieszkania (7% ludności obszarów miejskich i 15% ludności obszarów wiejskich). Zdecydowana większość osób pracujących poza swoim obszarem zamieszkania zatrudniona jest w obszarze miejskim (79%). Sytuację tą zobrazowano na wykresach 1.2 i 1.3.



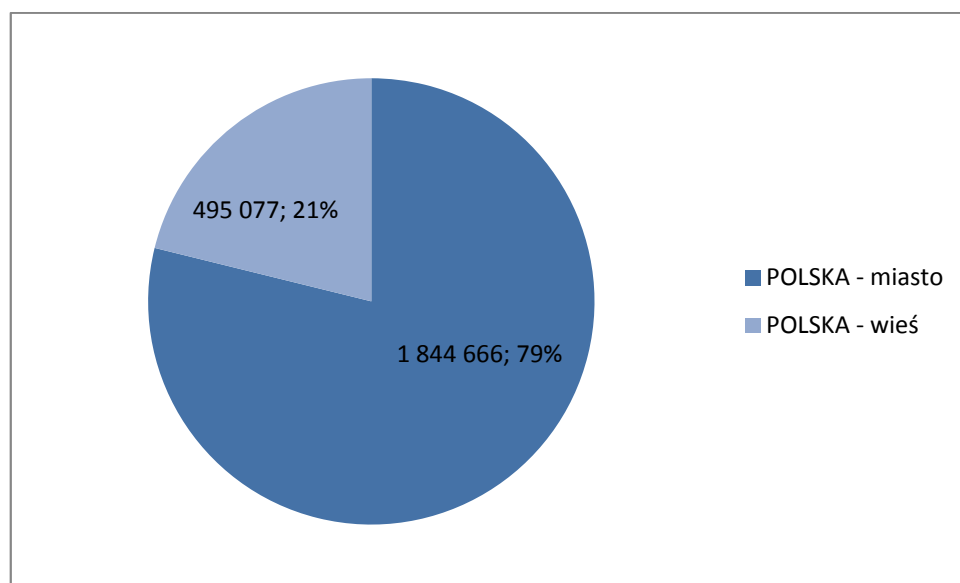
Wykres 1.2. Wyjeżdżający do pracy w stosunku do ludności w wieku produkcyjnym według miejsca zamieszkania, Polska 2006 r.

Źródło: Opracowanie własne na podstawie Banku Danych Regionalnych.

Można więc przypuszczać, że mieszkańcy zarówno miast jak i obszarów wiejskich, w sytuacji kiedy decydują się szukać pracy poza obszarem zamieszkania, szukają

jej w pierwszej kolejności w obszarach miejskich. Sytuacja taka jest po części skutkiem tego, że w miastach notuje się znacznie niższą stopę bezrobocia i wyższe średnie wynagrodzenie niż w obszarach wiejskich¹⁵.

Rozważając z kolei zróżnicowanie dojazdów do pracy w przekroju powiatów zauważmy, że z połowy powiatów mniej niż 10,1% procent ludności w wieku produkcyjnym wyjeżdża do pracy. Natomiast na terenie połowy powiatów przyjeżdżający do pracy stanowią mniej niż 7,32% zatrudnionych.



Wykres 1.3. Przyjeżdżający do pracy według obszaru zatrudnienia, Polska 2006 r.

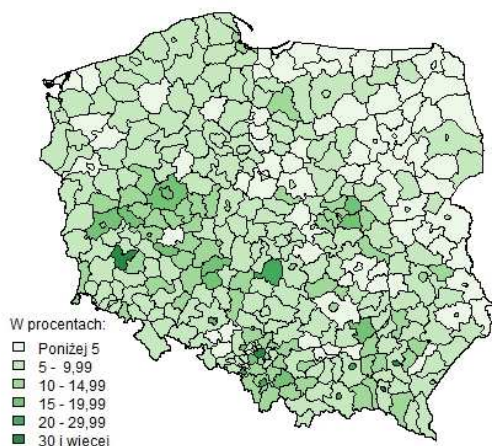
Źródło: Opracowanie własne na podstawie Banku Danych Regionalnych.

Odwołując się do wskazanych wcześniej koncepcji konstruujących lokalne rynki pracy wokół wyodrębnionych ośrodków centralnych, zaobserwować można wzrost natężenia zarówno przyjazdów, jak i wyjazdów do pracy w miarę zmniejszania się odległości od dużych miast (por. wyk. 1.4, 1.5). Spostrzeżenie to jest zgodne z koncepcją kręgów koncentrycznych Thunena określającą oddziaływanie dużych miast.

Stwierdzono również silne zróżnicowanie terytorialne natężenia dojazdów do pracy w różnych regionach Polski (por. wyk. 1.4, 1.5). Największy udział przyjeżdżających do pracy zaobserwowano dla następujących powiatów: m. Krosno, m. Katowice, których przyjeżdżający do pracy stanowili ponad 50% ludności w wieku produkcyjnym. Dla powiatów polkowickiego i m. Rzeszów udział ten przekraczał 30%. Wysoką, ponad

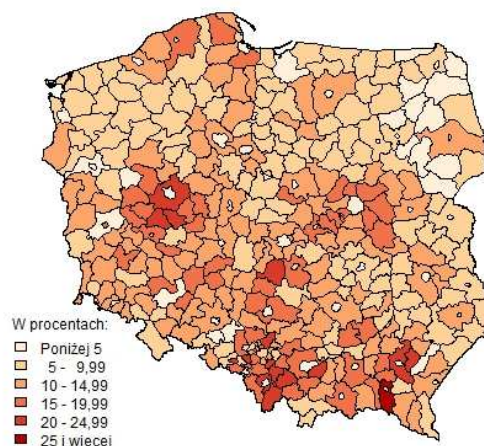
¹⁵ Dla przykładu stopa bezrobocia rejestrowanego w mieście Poznań wynosiła 5%, a w otaczającym je powiecie poznańskim 6,6%.

20 procentową wartość zaobserwowano między innymi w powiatach bełchatowskim, m. Tarnów czy m. Sopot czy poznańskim. Fakt, że wśród wymienionych powiatów nie znalazły się największe ośrodki miejskie związany jest z ich wielkością mierzona liczbą mieszkańców w wieku produkcyjnym, w stosunku do której nawet duża liczba przyjeżdżających jest względnie niska.



Wykres 1.4. Stosunek liczby przyjeżdżających do pracy do ludności w wieku produkcyjnym, przekrój powiatów, Polska 2006

Źródło: opracowanie własne na podstawie danych GUS



Wykres 1.5. Stosunek liczby wyjeżdżających do pracy do ludności w wieku produkcyjnym, przekrój powiatów, Polska 2006

Źródło: opracowanie własne na podstawie danych GUS

Udział osób wyjeżdżających do pracy w ludności w wieku produkcyjnym przekraczający 20% odnotowano w 15 powiatach. Największe natężenie tej zmiennej stwierdzono w powiatach: krośnieńskim, bielskim, będzińskim, oświęcimskim czy poznańskim (por. Gołata, Dehnel, Gruchociak, 2011).

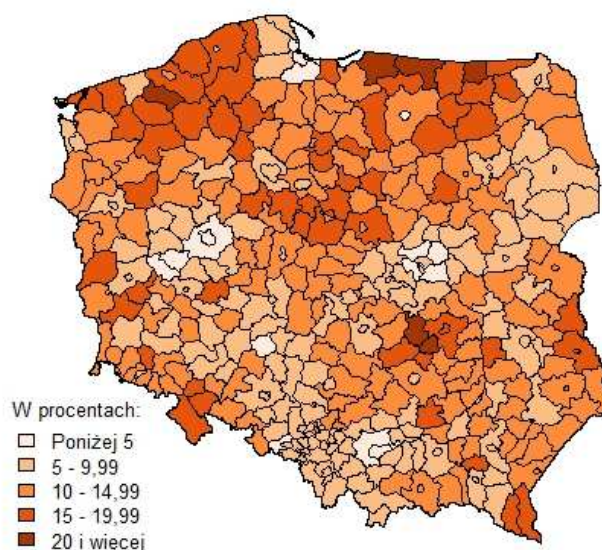
Delimitacja lokalnych rynków pracy przeprowadzona została dla 3062 jednostek bazowych zdefiniowanych jako gminy z uwzględnieniem podziału na część miejską i wiejską. Mobilność ludności w związku z zatrudnieniem w przekroju najmniejszych jednostek terytorialnych zauważmy, że w 80% gmin liczba wyjeżdżających do pracy jest większa niż liczba przyjeżdżających. Natomiast w 20% gmin wyjeżdża do pracy więcej niż połowa pracowników najemnych. Przytoczone liczby świadczą o dużej skali i powszechności zjawiska.

1.4. Charakterystyka lokalnych rynków pracy w Polsce

Jak wspomniano wcześniej, do przeprowadzenia delimitacji lokalnych rynków pracy potrzebne będą wyłącznie dane dotyczące dojazdów do pracy oraz ludności zatrudnionej w gminie zamieszkania, odpowiadające zakresem macierzy migracji. Źródła pozyskania takich danych omówione zostały w podrozdziale poprzednim, jednak zaznaczono w nim również, że najistotniejsze z nich, badanie przepływów związanych z zatrudnieniem, aktualne jest dotychczas wyłącznie dla roku 2006¹⁶. W związku z tym, w celu podjęcia próby oszacowania macierzy migracji i późniejszej delimitacji lokalnych rynków pracy dla roku 2011, konieczne było przedstawienie charakterystyki lokalnych rynków pracy w Polsce. Studia literaturowe w tym zakresie dostarczą niezbędnej wiedzy merytorycznej, która w dalszej kolejności, pozwoli wskazać zestaw potencjalnych zmiennych opisujących dojazdy do pracy.

Jak zwracają uwagę W. Gumuła, J. Socha, W. Wojciechowski (2007), migracje wahadłowe są silnie związane z sytuacją ekonomiczną, zarówno w obszarze zamieszkania jak i potencjalnego miejsca pracy osób podejmujących zatrudnienie. Zatem w procesie ich szacowania wśród zmiennych objaśniających powinny znaleźć się charakterystyki związane ze stopniem rozwoju i sytuacją na rynku pracy w obu tych obszarach. Zmienne te mogą mieć różny wpływ na natężenie dojazdów do pracy w różnych obszarach Polski, przede wszystkim dlatego, że dojazdy do pracy są w dużej mierze ograniczone przez odległość (por. Gawryszewski, Potrykowska, 1980). Mianowicie obszar o podobnej sytuacji gospodarczej w jednym województwie może być jednym z lepiej rozwiniętych i przyciągać tym samym osoby poszukujące pracy, zaś w innym województwie, charakteryzującym się lepszą sytuacją ekonomiczną, być postrzeganym jako mało atrakcyjny przez potencjalnych pracowników (por. wyk. 1.6). Wskazane różne relacje pozwala uwzględnić modelowanie wielopoziomowe. Ponadto, dzięki jego zastosowaniu możliwe jest uwzględnienie zmiennych określonych dla różnych poziomów agregacji przestrzennej.

¹⁶ W Ośrodku Statystyki Miast w Urzędzie Statystycznym w Poznaniu, trwają prace nad przygotowaniem aktualnej informacji o kierunkach dojazdów do pracy z wykorzystaniem nie tylko danych z bazy POLTAX, ale również formularzy ZUS. Badanie takie miałyby być powtarzane dwukrotnie w każdym okresie międzypisowym.



Wykres 1.6. Stosunek liczby bezrobotnych do ludności w wieku produkcyjnym, przekrój powiatów, Polska 2006

Źródło: Opracowanie własne na podstawie danych GUS

W literaturze podaje się, że w skład zestawu charakterystyk opisujących sytuację na rynku pracy, powinny wchodzić zmienne określające zarówno popyt jak i podaż pracy (por. Gołata, 2004). Podaż siły roboczej traktować można jako zbiorowość osób, zgłaszających chęć podjęcia pracy. W jej ramach wyszczególnić można podaż zrealizowaną, czyli ludność pracującą oraz podaż niezrealizowaną, czyli zbiorowość osób wyrażających chęć podjęcia pracy. W ramach niezrealizowanej strony podażowej dotyczy wyróżnić można osoby bezrobotne (około 60%), pozostałe osoby pozostające bez pracy deklarujące chęć jej podjęcia (około 7%) oraz osoby pracujące, lecz pragnące zmienić pracę (około 33%) (por. Kotowska, 2012). Wśród cech określających stronę podażową wymienić można liczbę ludności zatrudnionej i poszukującej zatrudnienia oraz jej oczekiwania co do stanowiska pracy, kwalifikacje czy wykształcenie (por. Gołata, 2004).

Popyt na siłę roboczą podzielić można na zaspokojony oraz niezaspokojony. Popyt zaspokojony odnosi się do obsadzonych miejsc pracy, zatem w pewnym sensie można utożsamiać go z podażą zrealizowaną. Popyt niezaspokojony dotyczy z kolei wolnych miejsc pracy. Jako charakterystyki strony popytowej wymienić można między innymi czas pracy, warunki pracy, kwalifikacje konieczne do objęcia zatrudnienia oraz dogodność położenia miejsca pracy. Kolejną grupą zmiennych mogą być charakterystyki dotyczące kształtowania się ceny na rynku pracy, czyli po prostu

wynagrodzenia za pracę (por. Góra, Sztanderska, 2006; Gumuła, Socha, Wojciechowski, 2007; Gruchociak, 2010).

W literaturze wśród wskaźników stopnia rozwoju regionów wymienia się stopień zatrudnienia (por. Kwiatkowski, Włodarczyk, 2012; Strahl, 1998). W warunkach polskich dane z tego zakresu zaczerpnąć można z badania *Pracujący w gospodarce narodowej*. Prezentowane w jego ramach dane dotyczą osób według miejsca pracy i dostarczane są w przekroju gmin z uwzględnieniem części miejskiej i wiejskiej.

Jedną z głównych charakterystyk potencjalnej strony podażowej lokalnego rynku pracy, bez rozróżniania podaży zrealizowanej bądź niezrealizowanej, jest liczba osób znajdujących się w wieku zdolności do pracy (por. Góra, Sztanderska, 2006). Dane z tego zakresu prezentowane są w ramach wyników ewidencji ruchu naturalnego ludności, na podstawie której w okresach międzypisowych, z zastosowaniem metody bilansowej, określany jest stan i struktura ludności według podstawowych cech demograficznych. Dane te dotyczą między innymi osób w wieku produkcyjnym według faktycznego miejsca zamieszkania. Większość badaczy w Polsce, jako wiek produkcyjny przyjmuje 18-64 lata dla mężczyzn oraz 18-59 lat dla kobiet, ponieważ od takiego wieku Polski Kodeks Pracy dopuszcza zatrudnienie bez specjalnej ochrony, górna granica zdeterminowana jest zaś uzyskaniem uprawnień emerytalnych¹⁷ (por. Góra, Sztanderska, 2006). Podejście takie zastosowano również w niniejszej pracy. Odpowiednie dane dotyczące osób w wieku produkcyjnym prezentowane są w Banku Danych Lokalnych GUS w tabeli o nazwie: *Stan ludności i ruch naturalny* w przekroju gmin z uwzględnieniem podziału na część miejską i wiejską.

Stronę podażową charakteryzują również informacje o kwalifikacjach osób chcących i mogących podjąć zatrudnienie, takie jak wykształcenie, wyuczony zawód czy doświadczenie zawodowe (por. Jarmołowicz, Kalinowska-Sufinowicz, 2012; Góra, Sztanderska, 2006). W ramach danych dotyczących kwalifikacji, najłatwiej dostępne są informacje o wykształceniu. W literaturze wskazuję się, że jest to jeden z podstawowych mierników postępu społeczno-gospodarczego regionów.

¹⁷ Z dniem 1 stycznia 2013 r. weszły w życie przepisy ustawy z dnia 11 maja 2012 r. o zmianie ustawy o emeryturach i rentach z Funduszu Ubezpieczeń Społecznych oraz niektórych innych ustaw (Dz. U. z dnia 6 czerwca 2012 r., poz. 637), nowelizujące wiele regulacji z zakresu m.in. zaopatrzenia emerytalnego. Wśród wprowadzonych zgodnie z nią zmian wyszczególnić należy stopniowe podnoszenie wieku emerytalnego, dla kobiet i mężczyzn, do 67 lat. Zmiany te nie dotyczą jednak badań prowadzonych w prezentowanej pracy, ponieważ jej zakres czasowy ograniczony jest do lat 2006-2011.

W szczególności podkreśla się udział osób z wyższym wykształceniem (por. Zagórska, Rauziński, 2011; Kotowska, 2012). W warunkach polskich dane z tego zakresu pozyskane mogą zostać z Badania Aktywności Ekonomicznej Ludności, którego wyniki prezentowane są w przekroju województw.

Za ważną charakterystykę stopnia rozwoju lokalnego rynku pracy uznawana jest także natężenie bezrobocia¹⁸ (por. Gołata, 2004; Jarmołowicz, Kalinowska-Sufinowicz, 2012; Gazińska, 1995; Kopczevska, 2010; Paradysz J., Paradysz K., 2012). Ponieważ rzetelna informacja o osobach pracujących, jednocześnie poszukujących pracy jest bardzo trudna do pozyskania, informacja o bezrobotnych jest jedyną dostępną charakterystyką niezrealizowanej podaży pracy. Informacje na temat bezrobocia w Polsce zaczerpnąć można między innymi z Narodowych Spisów Powszechnych, prowadzonego przez GUS *Badania Aktywności Ekonomicznej Ludności* (BAEL) oraz na podstawie rejestru bezrobotnych Ministerstwa Pracy i Polityki Społecznej. Informacje pochodzące ze spisów ludności oraz z BAELu definiują bezrobotnych zgodnie ze standardami międzynarodowymi, według zaleceń Międzynarodowej Organizacji Pracy (MOP)¹⁹. Mankamentem tych danych jest, w przypadku spisów, rzadkość publikacji odpowiadająca mniej więcej 10-letniemu okresowi przeprowadzania spisów. Dane BAEL publikowane są z częstotliwością kwartalną. Jednak w tym przypadku, reprezentacyjny charakter badania oraz ograniczona wielkość próby, umożliwiają szacunek jedynie na poziomie kraju, makroregionów i województw. Żadne informacje nie są publikowane na poziomie gmin, ani powiatów. Na poziomie lokalnego rynku pracy, dostępne są jedynie informacje z rejestru Ministerstwa Pracy i Polityki Społecznej gromadzone przez urzędy pracy. Są one publikowane z częstotliwością kwartalną²⁰. Dane o tzw. bezrobociu rejestrowanym dotyczą osób zarejestrowanych jako bezrobotne, które zgodnie z Ustawą z dnia 20 kwietnia 2004 r. o promocji zatrudnienia i instytucjach rynku pracy, są zdolne do pracy oraz gotowe do jej podjęcia (por. Rocznik Statystyczny Pracy, 2006).

¹⁸ Syntetyczną miarą natężenia bezrobocia jest stopa bezrobocia. Jednak ze względu na cel pracy dotyczący delimitacji lokalnych rynków pracy, przydatność tej miary jest ograniczona. Miara ta określana na poziomie powiatów jest obciążona pewnym błędem wynikającym z niezgodności sposobu określenia licznika i mianownika. Liczba bezrobotnych, – licznik – mierzona jest według stałego miejsca zamieszkania. Natomiast liczba aktywnych zawodowo – mianownik – według miejsca pracy. Z powodu tej niezgodności, GUS

nie publikuje informacji o stopie bezrobocia w przekroju gmin (szerzej na ten temat pisze E. Gołata, 2006)

¹⁹ Definicje stosowane w badaniach międzynarodowych zostały przyjęte podczas XV Międzynarodowej Konferencji Statystyków Rynku Pracy w styczniu 1993 r. (por. MOP, 1993, *Resolution concerning the International Classification of Status in Employment (ICSE)*)

²⁰ Informacja o natężeniu bezrobocia w skali kraju i województw publikowana jest w każdym miesiącu.

Informacje o liczbie bezrobotnych w przekroju gmin bez uwzględnienia podziału na część miejską i wiejską, publikowane są raz w roku.

W literaturze wskazuje się, że popyt na pracę powinien być z kolei opisany przez zapotrzebowanie pracodawców na pracowników, które przekłada się na liczbę oraz specyfikę miejsc pracy generowanych przez firmy i instytucje, a także gospodarstwa domowe

(por. Góra, Sztanderska, 2006). W ramach zatrudnienia wyróżnić można zatrudnienie w sektorze prywatnym oraz publicznym, przy czym wskazuje się, że zatrudnienie w sektorze prywatnym jest silniej związane z tempem wzrostu gospodarczego (por. Jarmołowicz, Kalinowska-Sufinowicz, 2012). Możliwe jest również rozpatrywanie zatrudnienia w podziale na różne sektory, przy czym najsilniej związane z sytuacją na rynku pracy jest zatrudnienie w sektorze usług (por. Jarmołowicz, Kalinowska-Sufinowicz, 2012). A.Kosztowniak (2011) wskazuje na wpływ kapitału zagranicznego na rozwój gospodarczy regionów. Ponieważ szczegółowe dane z tego zakresu dotyczące pojedynczych miejsc pracy są trudne do pozyskania, popyt na pracę scharakteryzowany zostanie przy pomocy liczby przedsiębiorstw. Informacje na temat liczby przedsiębiorstw zaczerpnięte mogą być z Krajowego Rejestru Urzędowego Podmiotów Gospodarki Narodowej (KRUPGN REGON). Na podstawie Rejestru Gospodarki Narodowej (REGON) dostępny jest podział podmiotów według sektorów własnościowych, w ramach którego w sektorze publicznym wyróżnia się: państwowe, samorządowe i komunalne jednostki prawa budżetowego, gospodarstwa pomocnicze, , przedsiębiorstwa komunalne, przedsiębiorstwa państwowe, spółki prawa handlowego, spółki handlowe, spółki prawa handlowego z udziałem kapitału zagranicznego, spółki handlowe z udziałem kapitału zagranicznego, oraz gospodarstwa pomocnicze, zaś w ramach sektora prywatnego: zakłady osób fizycznych prowadzących działalność gospodarczą, osoby fizyczne prowadzące działalność gospodarczą, spółki prawa handlowego z wyszczególnieniem tych z udziałem kapitału zagranicznego, spółki handlowe z wyszczególnieniem tych z udziałem kapitału zagranicznego, spółdzielnie, fundacje oraz stowarzyszenia i organizacje społeczne. Dane na temat liczby przedsiębiorstw z rejestru REGON dostępne są w przekroju gmin z uwzględnieniem podziału na część miejską i wiejską.

Oprócz powyższego podziału liczby przedsiębiorstw w gminach z podziałem na część miejską i wiejską dostępne są również według sekcji PKD²¹.

Kolejną ważną charakterystyką strony popytowej lokalnego rynku pracy są warunki pracy na oferowanych przez przedsiębiorstwa stanowiskach (por. Kotowska, 2012). Wyczerpująca informacja z tego zakresu nie jest w polskiej statystyce dostępna, jednak za pewną charakterystykę warunków pracy przyjąć można liczbę osób pracujących w warunkach zagrożenia. Dane na temat warunków pracy, czyli zespołu czynników występujących

w środowisku pracy, wynikających z procesu produkcyjnego oraz czynników związanych z wykonywaniem pracy gromadzone są corocznie przez GUS na podstawie sprawozdań o warunkach pracy realizowanych na formularzu Z-10 i publikowane w przekroju powiatów.

Ważną charakterystyką kondycji strony popytowej lokalnego rynku pracy jest również zaangażowanie przedsiębiorstw w rozwój wykorzystywanych technologii (por. Strahl, 2010). Dane z tego zakresu pozyskane mogą zostać z badania działalności innowacyjnej, którego wyniki publikowane są w przekroju województw.

Kolejną ważną charakterystyką lokalnego rynku pracy jest przeciętny poziom wynagrodzeń za pracę (por. Jarmołowicz, Kalinowska-Sufinowicz, 2012). Ze strony podażowej pracy są to wynagrodzenia otrzymywane przez pracujących, a dla bezrobotnych minimalna płaca, za którą gotowi są podjąć zatrudnienie. Ze strony popytowej należałoby uwzględnić wypłacane wynagrodzenia w przypadku obsadzonych miejsc pracy oraz oferowane płace na aktualnie nieobsadzonych stanowiskach (por. Góra, Sztanderska, 2006). Niestety dane o minimalnych wynagrodzeniach oczekiwanych przez osoby bezrobotne oraz wynagrodzeniach proponowanych przez firmy nie są dostępne. W związku z tym wyżej omówiona grupa charakterystyk reprezentowana będzie przez średnie wynagrodzenie. Warto zauważyć, że charakterystyka ta odnosi się zarówno do strony podażowej (w zakresie wynagrodzenia otrzymywanego przez osoby pracujące) jak i popytowej (w zakresie kosztów pracy ponoszonych przez przedsiębiorstwa) lokalnych rynków pracy. Dane na ten temat udostępnione przez GUS w przekroju powiatów dotyczą zbiorowości pracowników przedsiębiorstw zatrudniających więcej niż

²¹ Przynależność do sekcji określana jest w momencie rejestracji. Na poziomie gmin nie wyszczególnia się wielkości przedsiębiorstw.

9 osób a informują o średnich wynagrodzeniach w ujęciu brutto, tj łącznie z zaliczkami na poczet podatku dochodowego od osób fizycznych oraz ze składkami na obowiązkowe ubezpieczenia społeczne (emerytalne, rentowe i chorobowe) płaconymi przez ubezpieczonego pracownika.

Ważną determinantą natężenia dojazdów do pracy jest odległość pomiędzy miejscem zamieszkania osób zasilających podaż pracy oraz potencjalnymi miejscami pracy. Do pewnego stopnia pracownicy są skłonni dojeżdżać do pracy na dalszą odległość, jednak pod warunkiem wyższego wynagrodzenia i tylko w pewnych granicach (por. Gumuła, Socha, Wojciechowski, 2007). Odległość pomiędzy gminami może być mierzona w kilometrach, za pomocą czasu oraz kosztów dojazdu. O ile odległość w linii prostej podana zostanie wprost, czas i koszty dojazdu scharakteryzowane będą przy pomocy informacji publikowanych w ramach badania transportu i łączności. Uwzględnione w tym zakresie zostaną dane na temat dróg publicznych powiatowych i gminnych (w przekroju powiatów) oraz komunikacji miejskiej i transportu kolejowego (w przekroju województw). Z kolei zgodnie z „teorią ośrodków centralnych” (por. Beguin, 1992), która głosi, że regiony wokół dużych miast rozwijają się szybciej niż te od nich oddalone, za charakterystykę stopnia rozwoju przyjąć można również odległość od ośrodków centralnych.

Na podsumowanie rozważań uporządkowano wskazane powyżej potencjalne determinanty lokalnych rynków pracy w podziale na grupy. Uwzględniono charakterystyki strony podaźowej i popytowej lokalnego rynku pracy oraz cen a także odległości oraz położenie gmin.

1. Charakterystyki strony podaźowej rynku pracy:

- Stopień zatrudnienia
- Liczba ludności w wieku produkcyjnym
- Udział ludności z wykształceniem wyższym
- Natężenie bezrobocia

2. Charakterystyki strony popytowej lokalnego rynku pracy:

- Natężenie przedsiębiorstw
- Liczba osób pracująca w warunkach zagrożenia
- Nakłady na działalność innowacyjną w przedsiębiorstwach

3. Charakterystyki ceny na rynku pracy:
 - Średnie wynagrodzenie
4. Charakterystyki odległości, na której podejmowane są dojazdy:
 - Odległość w linii prostej mierzona w kilometrach
 - Ilość dróg gminnych i powiatowych
 - Ilość torów kolejowych
 - Liczba miejsc w pojazdach komunikacji miejskiej
5. Charakterystyki położenia gmin:
 - Odległość od stolicy województwa
 - Odległość od najbliższego ośrodka centralnego

Z powyższych rozważań wynika, że informacje mogące posłużyć w dalszej części pracy do oszacowania danych dotyczących dojazdów do pracy dostępne są na różnych poziomach agregacji (gmin, powiatów oraz województw). Ograniczenie badania do poziomu gmin, dla którego przeprowadzone zostaną szacunki spowodowałoby istotną utratę informacji. Jako rozwiązanie tego problemu zaproponowano, w rozdziale czwartym, konstrukcję modelu wielopoziomowego. Dzięki temu podejściu, możliwe będzie nie tylko uwzględnienie charakterystyk rynku pracy na różnych poziomach oraz relacji pomiędzy nimi, ale także uchwycenie specyfiki poszczególnych województw czy powiatów.

Rozdział 2. Metody delimitacji lokalnych rynków pracy

W literaturze znaleźć można różne klasyfikacje metod delimitacji lokalnych rynków pracy (por. Casado-Diaz, Coombes 2011). W niniejszym rozdziale omówione zostaną dwa sposoby klasyfikacji - pierwszy związany z założeniem ośrodków centralnych a drugi z zastosowanym sposobem łączenia. Opisane zostaną także przykładowe metody delimitacji, stosowane w różnych krajach, z uwzględnieniem podanych podziałów. Na podstawie przedstawionego przeglądu literatury wybrane zostaną dwie metody, wykorzystane następnie w celu delimitacji lokalnych rynków pracy w Polsce. Na podstawie analizy obu metod zaproponowano także samodzielnie opracowane podejście. Ze względu na przyjęte rozwiązania metodologiczne określono je mianem taksonomicznego.

2.1. Metody delimitacji ich klasyfikacja

Metody delimitacji lokalnych rynków pracy najogólniej podzielić można na dwie rozłączne grupy: homogeniczne, zakładające jednorodność lokalnych rynków pracy i heterogeniczne, dopuszczające różnorodne lokalne rynki pracy. W przypadku metod homogenicznych zakłada się, że każda jednostka bazowa musi zostać przyporządkowana dokładnie do jednego lokalnego rynku pracy, co nie ma miejsca w przypadku metod heterogenicznych (por. Van der Laan, Schalke, 2001). Część autorów wskazuje, że metody heterogeniczne lepiej oddają złożoność relacji wynikających z dojazdów do pracy (por. Kallioras, Kandyliis, Kromyidakis, Pantazis, 2011). Z drugiej strony tylko metody homogeniczne spełniają omówiony wcześniej postulat delimitacji lokalnych rynków pracy, dotyczący rozłączności jednostek bazowych (por. Eurostat, 1992; Casado-Diaz, Coombes, 2011). W związku z powyższym prezentacja metod ograniczona zostanie do metod homogenicznych.

METODY DEDUKCYJNE I INDUKCYJNE

Wśród metod homogenicznych wyróżnić można podejścia dedukcyjne, określane też jako „z góry na dół” (top-down) oraz indukcyjne, nazywane także „od dołu do góry” (bottom-up) (por. Casado-Diaz, Coombes, 2011; Fusco, Caglioni 2011). Klasyfikacja ta jest pełna i rozłączna, to znaczy, każda metoda homogeniczna określona może być jako dedukcyjna albo indukcyjna.

W przypadku metod dedukcyjnych przyjmuje się z góry pewne założenia dotyczące końcowego kształtu i liczby lokalnych rynków pracy. W szczególności przed rozpoczęciem delimitacji określone są ośrodki centralne, a następnie wokół każdego z nich tworzony jest dokładnie jeden lokalny rynek pracy. Wśród tych metod wyróżnić można podejścia, w których wybór ośrodków centralnych dokonywany jest a priori i nie jest częścią procedury a informacją wejściową (por. Van der Haegen, Pattyn, 1980). W pozostałych przypadkach metod dedukcyjnych stosowane są algorytmy rozpoczynające się od wyodrębnienia zestawu ośrodków centralnych (por. Lagnerö, 2003). Metody dedukcyjne oparte są więc na założeniu, że rozwój każdego lokalnego rynku pracy jest stymulowany przez dokładnie jeden ośrodek centralny. Założenie to może być poparte teorią Johanna Heinricha von Thunena i w przypadku wielu lokalnych rynków pracy wydaje się być słuszne. Biorąc jednak pod uwagę aglomeracje złożone z więcej niż jednego dużego miasta oraz wciąż pogłębiającą się decentralizację zatrudnienia łatwo wskazać sytuacje, w których opisane podejście nie powinno być stosowane.

Rzeczywistą złożoność struktury zatrudnienia pozwalają lepiej uwzględnić metody indukcyjne. W przypadku tej grupy metod, lokalne rynki pracy wyodrębniane są na podstawie analizy dwukierunkowych zależności. Najczęściej, do pomiaru tych relacji wykorzystuje się dojazdy do pracy pomiędzy wszystkimi parami jednostek bazowych. Ponadto, wszystkie jednostki bazowe traktowane są jednakowo.

METODY OPARTE NA ZASADACH I HIERARCHICZNE

Inną klasyfikacją metod delimitacji lokalnych rynków pracy, opartą na kryterium sposobu łączenia, jest podział na podejścia oparte na zasadach (ang. rules-based methods) oraz hierarchiczne (por. Casado-Diaz, Coombes, 2011). W algorytmach hierarchicznych stosuje się jedno kryterium łączenia, aż do momentu kiedy spełniony zostanie warunek końcowy. Z kolei w przypadku metod opartych na zasadach, na różnych etapach procedury, stosuje się różne kryteria łączenia. Przewaga metod należących do którejś z tych grup jest dyskusyjna i zależy od wielu czynników takich jak liczba parametrów czy przyjęte kryterium łączenia. Warto jednak podkreślić, że zaletą metod hierarchicznych jest zazwyczaj zdecydowanie mniejsza liczba parametrów i wartości progowych, koniecznych do przyjęcia a priori niż w przypadku algorytmów opartych na zasadach.

W celu usystematyzowania przedstawionego dalej przeglądu metod delimitacji lokalnych rynków pracy zestawiono je w Tabelicy 2.1 w czterech grupach, powstałych jako kombinacja dwóch opisanych sposobów klasyfikacji. Górne wiersze Tabelicy 2.1 zawierają metody dedukcyjne a dolne – metody indukcyjne. Z kolei w lewej kolumnie tabelicy umieszczono metody oparte na zasadach zaś w prawej – metody hierarchiczne.

Tabela 2.1. Klasyfikacja przykładowych metod delimitacji lokalnych rynków pracy, stosowanych w różnych krajach

Rodzaj metody	oparte na zasadach	hierarchiczne
dedukcyjne	<p>USA:</p> <ul style="list-style-type: none"> ▪ Spotila, 2000 ▪ US Department of Labor, 2003, 2013 <p>Słowenia:</p> <ul style="list-style-type: none"> ▪ Konjar, Lisec, Drobne, 2010 	<p>Szwecja:</p> <ul style="list-style-type: none"> ▪ Carlsson et al., 1993, ▪ SCB, Statistics Sweden, 2010
indukcyjne	<p>Wielka Brytania:</p> <ul style="list-style-type: none"> ▪ Coombes, Green, Oppenshaw, 1986 <p>Włochy:</p> <ul style="list-style-type: none"> ▪ ISTAT, 1997 ▪ ISTAT, 2005 <p>Hiszpania:</p> <ul style="list-style-type: none"> ▪ Casado-Diaz, 2000 <p>Nowa Zelandia:</p> <ul style="list-style-type: none"> ▪ Papps, Newell, 2002 ▪ Newell, Perry, 2005 <p>Australia:</p> <ul style="list-style-type: none"> ▪ Watts, 2004, ▪ Bamber, Walter, 2009 <p>Grecja:</p> <ul style="list-style-type: none"> ▪ Prodromidis, 2008, 2010 	<p>USA:</p> <ul style="list-style-type: none"> ▪ Tolbert, Sizer, 1996 ▪ US Department of Agriculture, 2004 <p>Francja:</p> <ul style="list-style-type: none"> ▪ INSEE, 1998 <p>Dania:</p> <ul style="list-style-type: none"> ▪ Kristensen, 1998

Zródło: Opracowanie własne

Poniżej opisano w skrócie metody delimitacji lokalnych rynków pracy, stosowane przez organy administracji publicznej w wybranych krajach (por. Casado-Diaz, Coombes 2011). W USA kilka urzędów użyteczności publicznej przeprowadza niezależnie delimitacje. I tak Office of Management and Budget wyodrębnia obszary metropolitalne, które traktowane są również jako lokalne rynki pracy (Core based Statistical Areas, CBSAs), (por. Spotila, 2000). Jest to metoda heterogeniczna, ponieważ obszary peryferyjne nie są zazwyczaj włączane do żadnego z obszarów metropolitalnych. Ponadto jest to podejście dedukcyjne; ośrodki centralne wybierane są według kryterium liczby

mieszkańców. Ze względu na kolejne kryterium klasyfikacji jest to algorytm oparty na zasadach. W poszczególnych etapach algorytmu jednostki bazowe dołączane są do ośrodków centralnych na podstawie określonego zestawu kryteriów. Wartości wykorzystywanych w tym celu parametrów obliczane są na podstawie macierzy migracji. Przyłączenie jednostki bazowej do wyznaczanego obszaru metropolitalnego jest możliwe tylko wówczas, jeżeli mają one wspólną granicę.

Wyznaczone w powyższy sposób obszary metropolitalne są z kolei podstawą delimitacji lokalnych rynków pracy (Labor Market Areas, LMAs) przez US Department of Labor (por. US Department of Labor, 2003, 2013). Mianowicie, każdy z wyznaczonych obszarów metropolitalnych uznawany jest za odrębny lokalny rynek pracy, a pozostałe jednostki bazowe łączone ze sobą na podstawie kryterium dojazdów do pracy. Ciągłość wyznaczonych obszarów osiągnąca jest w ten sposób, że jeżeli po zakończeniu procedury któryś z lokalnych rynków pracy nie jest spójnym terenem, traktowany jest, jako dwa (lub więcej) lokalne rynki pracy. Warto zauważyć, że wynikiem jest więc zestaw pokrywających cały badany obszar oraz ciągłych lokalnych rynków pracy. Jeżeli jednak, wyznaczone przez Office of Management and Budget obszary metropolitalne nie są rozłączne, również wyznaczone lokalne rynki pracy nie będą rozłączne. Oznacza to, że nie jest spełniony trzeci postulat delimitacji lokalnych rynków pracy, dotyczący rozłączności.

W USA delimitację lokalnych rynków pracy przeprowadza również US Department of Agriculture (por. Tolbert, Sizer, 1996; US Department of Agriculture, 2004). Wyznaczone w jej wyniku obszary określane są mianem stref dojazdów do pracy (Commuting Zones). W tym przypadku wykorzystywany jest algorytm indukcyjny o charakterze hierarchicznym. Pary gmin są łączone, jeżeli spełnione jest kryterium oparte na dojazdach do pracy. Ponadto, ciągłość wyznaczanych obszarów nie jest osiągnąca przy pomocy dodatkowego warunku. W związku z powyższym, otrzymane lokalne rynki pracy mogą charakteryzować się niewielkimi odstępstwami od zasady ciągłości. Przy pomocy dodatkowego warunku narzucana jest jednak minimalna wielkość lokalnego rynku pracy, mierzona przy pomocy liczby mieszkańców.

Także Urząd Statystyczny Francji wyznacza lokalne rynki pracy, nazywane strefami zatrudnienia (Zones d'emploi). Metodologia i wyniki tej delimitacji zostały

przedstawione w opracowaniu National Institute for Statistics and Economic Studies²² (por. INSEE, 1998). Wykorzystano w niej hierarchiczny algorytm indukcyjny. Procedura grupowania jednostek bazowych przebiega na podstawie kryteriów obliczanych na podstawie macierzy migracji związanych z zatrudnieniem. Ponadto przyjmuje się minimalny rozmiar lokalnego rynku pracy. Ta minimalna wielkość wyznaczana jest na podstawie zasobów siły roboczej danego obszaru. W pewnych sytuacjach dopuszczane są jednak odstępstwa od tej zasady.

Jedną z najpowszechniej stosowanych na Świecie metod delimitacji lokalnych rynków pracy²³ jest indukcyjny algorytm oparty na zasadach. Został on opracowany przez M.G.Coombes'a, A.E.Greena i S.Openshawa w 1986 roku. Procedura ta stosowana jest do dziś, w pierwotnej formie lub po wprowadzeniu pewnych modyfikacji, przez urzędy statystyczne w wielu krajach. Algorytm M.G.Coombes'a jest wykorzystywany między innymi w Wielkiej Brytanii przez Department of Labour (por. Coombes, 1998), we Włoszech przez Italian Institute of Statistics (por. ISTAT, 1997; ISTAT, 2005), w Nowej Zelandii (por. Papps, Newell, 2002; Newell, Perry, 2005), Australii (Watts, 2004; Bamber, Walter, 2009), Hiszpanii (Casado-Diaz, 2000) oraz Grecji (Prodromidis, 2008; 2010). Procedura ta jest także rekomendowana przez Eurostat, jako standardowe podejście do definiowania lokalnych rynków pracy w krajach europejskich (por. Eurostat, 1992). W metodzie M.G. Coombes'a generowane są obszary lokalnych rynków pracy, dla których spełnione są założone warunki. Warunki te dotyczą zarówno samowystarczalności w zakresie podaży pracy (udział procentowy mieszkańców danego obszaru w ogólnej liczby pracujących na tym obszarze) jak i w zakresie popytu na pracę (udział procentowy miejsc pracy na danym obszarze, zajętych przez jego mieszkańców).

Algorytm ten, ze względu na jego powszechność oraz szerokie zastosowanie w wielu krajach, opisany zostanie szczegółowo w podrozdziale 2.3. W dalszej części rozprawy zostanie także podjęta próba jego zaimplementowania w warunkach polskich.

Również w Norwegii, stosowana jest procedura wyodrębniania lokalnych rynków pracy oparta na danych dotyczących dojazdów do pracy. Jednak równocześnie, w roli zmiennych wspomagających wykorzystywane są informacje o wielkości sprzedaży,

²² National Institute for Statistics and Economic Studies jest francuskim odpowiednikiem Głównego Urzędu Statystycznego.

²³ Również w tym algorytmie, lokalne rynki pracy nazywane są strefami dojazdów do pracy (Travel-to-Work Areas).

liczbie mieszkańców, obszarach dystrybucji gazet, migracjach stałych oraz innej wiedzy lokalnej (por. Statistics Norway, 2001). W Szwecji oraz Danii stosowana jest metoda dedukcyjna (por. Carlsson et al., 1993; SCB Statistics Sweden, 2010). Ośrodki centralne wyznaczane są na podstawie informacji o dojazdach do pracy, jeżeli jednak wykazują silną interakcję, mogą zostać połączone. Następnie, pozostałe jednostki bazowe przyporządkowywane są do ośrodków centralnych. Kryterium łączenia określone jest na podstawie natężenia dojazdów do pracy. Także w Słowenii zastosowano metodę dedukcyjną w celu określenia rejonów funkcjonalnych, przy czym do ich oszacowania wykorzystano informacje o codziennych dojazdach do pracy (por. Konjar, Lisec, Drobne, 2010).

Obok metody dedukcyjnej, w Danii opracowano metodę indukcyjną. Zastosowano ją w celu wyodrębnienia rejonów funkcjonalnych, na podstawie rozkładu dojazdów do pracy. Metodę tą zaliczyć można do metod hierarchicznych (por. Kristensen, 1998).

Ponadto w wielu krajach w dziedzinie delimitacji lokalnych rynków pracy prowadzone są badania akademickie nad wykorzystaniem nowych metod obliczeniowych, m.in. algorytmów genetycznych (por. Florez-Revuelta, Casado-Diaz, Martinez-Bernabeu 2008) oraz metody hierarchicznej agregacji Intramax (por. Mitchell, Bill, Watts, 2007; Nel, Krygsman, Jong, 2008; Landre, 2012; Landre, Hakansson, 2013). Jednakże metody te nie uzyskały dotychczas rekomendacji do powszechnego stosowania.

Widoczne jest, że najliczniejszą grupę stanowią metody indukcyjne, oparte na zasadach (por. tab. 2.1). Podejścia indukcyjne mają swoje silne i słabe strony. Umożliwiają uwzględnienie różnych złożonych rozkładów dojazdów, ale nie mają prostej, łatwo rozpoznawalnej postaci, tak jak podejścia dedukcyjne. Jednakże, ze względu na coraz szybsze, ciągłe zmiany rozkładów dojazdów, założenie jako podstawy ośrodków centralnych, jak to ma miejsce w metodach dedukcyjnych, wydaje się być znacznie mniej dogodne.

Dlatego też w celu przeprowadzenia delimitacji lokalnych rynków pracy w Polsce wybrano, jako pierwszy najpowszechniej stosowany algorytm indukcyjny Coombes'a, oparty na zasadach (por. Eurostat, 1992), opisany w podrozdziale 2.3.

Dla umożliwienia porównania z procedurą z grupy metod hierarchicznych wybrano algorytm indukcyjny, opracowany i zastosowany w Danii (por. Kristensen, 1998), opisany w podrozdziale 2.4.

2.2. Stan badań związanych z delimitacją lokalnych rynków pracy w Polsce

Od wielu lat w Polsce nie przeprowadzono kompleksowej delimitacji lokalnych rynków pracy. Głównym powodem braku opracowań w tym zakresie był brak odpowiednich danych. W przeszłości jednak, prowadzono badania, umożliwiające ocenę mobilności pracowniczej w Polsce. Pierwsze badanie przeprowadzone zostało już w 1956 roku (por. Gawryszewski, 2005; Rosik, Stępiak, Wiśniewski, 2010). Źródłem informacji o zatrudnionych w gospodarce uspołecznionej był spis kadrowy. Obejmował on m.in. takie dane jak miejsce zamieszkania, z którego pracownik dojeżdża lub dochodzi do pracy i miejsce faktycznego zatrudnienia. Następne spisy kadrowe odbywały się dość regularnie, w latach: 1956, 1958, 1964, 1968, 1973, 1977 oraz 1983. W roku 1978 dane o dojazdach do pracy zebrano także w trakcie Narodowego Spisu Powszechnego (por. Gawryszewski, 1989). Ponadto w roku 1959 oraz latach następnych przeprowadzono badania ankietowe dojazdów do pracy we wszystkich uspołecznionych zakładach pracy (Lijewski, 1967).

Materiały statystyczne, zawierające dane o pewnego rodzaju macierzy przepływów, pochodzące ze Spisu Kadrowego 1973, przedstawiono w pracy (por. Korcelli, Potrykowska, Bodzak, 1981). Dane te dotyczą kierunków dojazdów do pracy i obejmują dojazdy z 3071 miast i gmin, jako miejsc zamieszkania ludności do 351 miast uznanych za ośrodki zatrudnienia. Dostępna na podstawie Spisu Kadrowego macierz przepływów była macierzą prostokątną o wymiarach 3071x351. Odpowiadała więc ona fragmentowi macierzy przepływów stosowanej w większości metod delimitacji lokalnych rynków pracy. Niestety, nie zawierała ona informacji o przyjeżdżających do pracy do gmin, które nie pełniły roli ośrodków centralnych. Ponieważ na podstawie takiej macierzy nie można było ocenić powiązań pomiędzy dowolną parą gmin, nie umożliwiała ona delimitacji lokalnych rynków pracy rozumianej jak opisano powyżej.

Z drugiej strony, informacje publikowane w Spisach Kadrowych dotyczyły powiązań pomiędzy miastami a wszystkimi gminami. Ujęcie takie stwarzało możliwość delimitacji

obszarów metropolitalnych. Badania w tym zakresie były prowadzone w Polsce w okresie Spisów Kadrowych na szeroką skalę. Pod wieloma względami jest to zagadnienie podobne do delimitacji lokalnych rynków pracy. Jednak celem wyodrębniania obszarów metropolitalnych jest określenie zakresu oddziaływania ośrodka centralnego na obszary go otaczające. Z odmiennej motywacji przeprowadzania obydwóch delimitacji, wynikają główne różnice pomiędzy nimi. Do najważniejszych należą: położenie nacisku na zależności pomiędzy ośrodkami centralnymi i pozostałymi jednostkami terytorialnymi oraz pominięcie powiązań pomiędzy jednostkami peryferyjnymi. Ponadto, wynikiem procesu delimitacji lokalnych rynków pracy jest rozłączny i pokrywający całą zbiorowość podział jednostek na grupy. Natomiast, wyodrębnione obszary metropolitalne nie muszą pokrywać całej przestrzeni, ponadto mogą na siebie nachodzić. Jako podobieństwo obu procesów należy wskazać wykorzystywanie danych o dojazdach do pracy.

W związku z brakiem danych o dojazdach do pracy, badania i delimitacje lokalnych rynków pracy i zasięgów oddziaływania miast podupadły w okresie 1989-2011 (por. Śleszyński, Czapiewski, 2011). Ostatnie badania z tego zakresu prowadzone były w latach dziewięćdziesiątych. I tak, przeprowadzone zostały podziały lokalnych rynków pracy w Polsce przy pomocy modelu potencjału (por. Potrykowski, Taylor, 1982), modelu grawitacji J.D.Nystuena i M.F.Dacey (1961) (por. Chojnicki, 1966) oraz na podstawie odsetka osób dojeżdżających do pracy (por. Korcelli, Gawryszewski, Potrykowska., 1992). Ponadto Główny Urząd Statystyczny dwukrotnie opublikował wyodrębnione obszary metropolitalne, nazywane strefami intensywnych dojazdów do pracy, wokół największych miast polskich (por. GUS, 1973, GUS, 1986). W ramach wielu opracowań nie tylko przeprowadzono delimitację obszarów metropolitalnych, ale również dokonano hierarchicznej klasyfikacji otrzymanych ośrodków dojazdów (por. Gontarski, 1980; Gawryszewski, 1977; Korcelli 1981, Ohme, 1988).

Na podstawie danych ze spisu kadrowego 1973 przeprowadzona została również delimitacja obszarów metropolitalnych przy wykorzystaniu metody grupowania hierarchicznego. Grupowanie przeprowadzono na podstawie relacji największego przepływu (por. Nstuyen, Dacey, 1961), przy uwzględnieniu rangi ośrodka centralnego. W metodzie tej zakłada się występowanie jednostek nadrzędnych i podrzędnych oraz asymetryczne, przechodnie relacje między nimi. Jej wadą jest niemożliwość bezpośredniej oceny stanu zrównowżenia zasobów miejsc pracy i rąk do pracy. W celu weryfikacji istotności uwzględnionych powiązań zastosowano metodę J.D.Nystuena

i M.F.Dacey , uwzględniając drugi, co do wielkości przepływ (por. Korcelli, Potrykowska, Bodzak, 1981). Metoda ta pozwoliła na określenie struktury powiązań w złożonym układzie relacji i pozycji poszczególnych ośrodków centralnych, mierzonej według wielkości dojazdów do pracy.

Z kolei na podstawie danych ze spisu kadrowego, z roku 1983, wydzielone zostały rejony powiązań poszczególnych społeczności z miejscami zatrudnienia (por. Gocał, Rakowski, 1991). W tym celu autorzy posłużyli się autorskim systemem kryteriów wyznaczonych w oparciu o intensywność dojazdów do pracy.

A.Gawryszewski i A.Potrykowska (1980) zastosowali model odległości dojazdów do pracy o funkcji potęgowej i dokonali analizy rozkładów odległości przemieszczeń wahadłowych na przykładzie 49 ówczesnych miast wojewódzkich i wylosowanych dziewięciu innych ośrodków miejskich.

Obszar metropolitalny, w przeciwieństwie do lokalnego rynku pracy, wyznaczany może być indywidualnie dla wybranego miasta. W literaturze polskiej, znaleźć można między innymi, opis delimitacji obszaru metropolitalnego miasta Warszawy, wyznaczonego na podstawie danych PKP. W jego ramach autorzy dokonali analizy dojazdów do pracy do stolicy (por. Cegielski, Kluszewski, 1952). Również wiele innych opracowań z tej dziedziny ogranicza się do analizy strefy wpływów Warszawy (por. Potrykowska, 1983; Gawryszewski, Korcelli, Potrykowska, 1998). Przedmiotem pracy T. Czyż było z kolei badanie zróżnicowania regionalnego Polski z uwzględnieniem regionów rdzeniowych i obszarów peryferyjnych, na podstawie zastosowania modelu potencjału (por. Czyż, 1996).

W okresie braku danych o dojazdach do pracy podejmowane były próby wyznaczenia obszarów metropolitalnych. Rozważania te, oparte były na przykład, na pośredniej identyfikacji oddziaływania miast, mierzonej przy pomocy cech strukturalno-morfologicznych, takich jak gęstość zaludnienia, saldo migracji, nowa zabudowa mieszkaniowa, odsetek przedsiębiorstw usługowych wyższego rzędu, wykształcenie mieszkańców (por. Smętkowski, 2001, 2007; Smętkowski, Jałowiecki, Gorzelak, 2008; Korcelli, Śleszyński, 2006). Innym podejściem było zastąpienie informacji o rzeczywistych dojazdach do pracy, dostępnością czasowo-przestrzenną. Wykorzystano w tym celu oszacowanie czasu potrzebnego na pokonanie dystansu pomiędzy miejscem zamieszkania a potencjalnym miejscem pracy (por. Komornicki, Śleszyński, 2009).

W sytuacji udostępnienia wyników *Badania przepływów ludności związanych z zatrudnieniem*, prace nad wyznaczeniem obszarów metropolitalnych zostały ponownie podjęte. P.Śleszyński i K.Czapiewski (2011) przeprowadzili delimitację regionów miejskich, stosując metodę grupowania hierarchicznego J.D.Nystuena i M.F.Daceya (1961). Wyodrębnienie regionów przeprowadzono na podstawie danych macierzy migracji udostępnionej w ramach *Badania przepływów ludności związanych z zatrudnieniem* (por. Dojazdy do pracy w Polsce, red. K. Kruszka, 2010). Macierz przepływów wykorzystana została także w pracy S. Filas-Przybył, T. Klimanka i J. Kowalewskiego (2012). W pracy tej podjęto próbę zastosowania modelu gravitacji do opisu ciężenia określonych jednostek podziału terytorialnego kraju do ośrodków miejskich, odgrywających istotną rolę w analizie lokalnych i regionalnych rynków pracy. Rozważania przeprowadzono na przykładzie województwa wielkopolskiego.

Ze względu na niedostępność danych, delimitacja lokalnych rynków pracy w Polsce, na podstawie macierzy migracji, nie była możliwa aż do roku 2011. Z wiedzy autorki wynika, że delimitacja taka nie została przeprowadzona również po roku 2011 do chwili obecnej. Należy jednak wspomnieć, że podejmowane były wcześniej pewne badania w tym zakresie. W swojej pracy T.Szumlicz (1993) wskazuje na zasadność delimitacji lokalnych rynków pracy na podstawie informacji o dojazdach do pracy oraz omawia pewne aspekty metodologiczne tego zagadnienia. Jednak, prawdopodobnie ze względu na brak odpowiednich danych, ogranicza się do rozważań teoretycznych. Również J.Stachowski (1993) omawia w swojej pracy możliwe podejścia do delimitacji lokalnych rynków pracy. Wskazuje on między innymi, jak do tego celu zastosować metodę J.D.Nystuena i M.F.Daceya. Omawia także metodę porządkowania wierzchołków sieci (Maik, 1976) oraz jedno z podejść stosowanych w Wielkiej Brytanii (por. Smart, 1974; 1981), ogranicza się jednak, do rozważań teoretycznych. Z kolei J.Runge swoje rozważania nad rozwojem lokalnych rynków pracy, z regionu Zagłębia Dąbrowskiego, rozpoczyna od określenia ich zasięgu terytorialnego przy pomocy metod eksperckich opartych na przesłankach historycznych (por. Runge, 1993).

2.3. Europejski algorytm regionalizacji

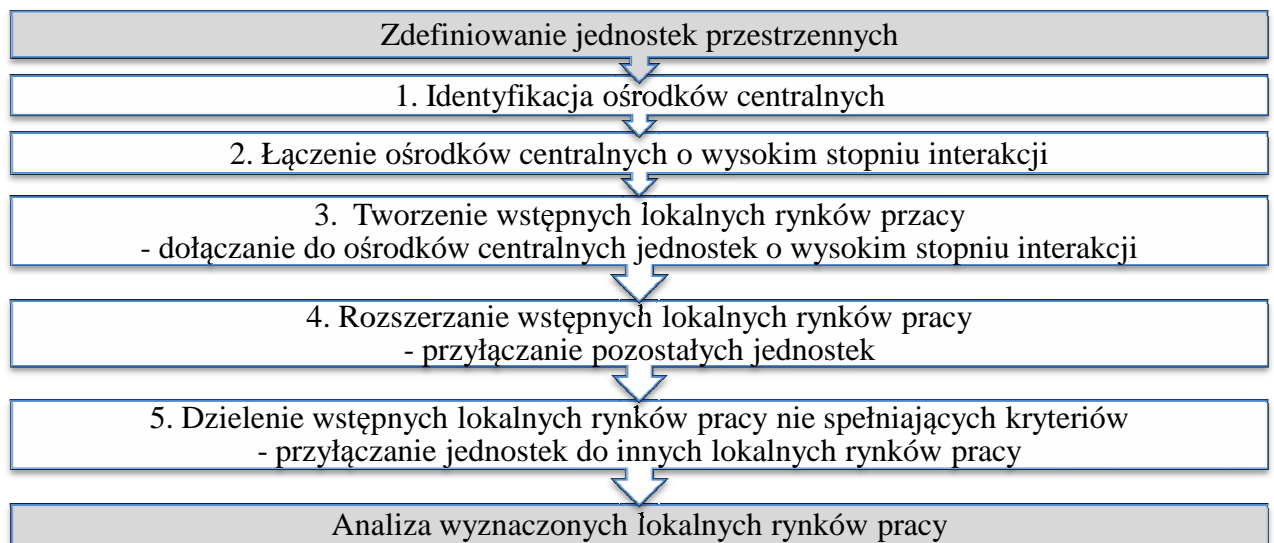
Pierwszym podejściem, wskazanym na podstawie przeglądu literatury, jest algorytm indukcyjny, oparty na zasadach, opracowany przez M.G.Coombes'a, A.E.Greena i S.Openshaw (1986) Algorytm ten, został pierwotnie skonstruowany do delimitacji

lokalnych rynków pracy w Wielkiej Brytanii, a w dalszej części pracy wykorzystany zostanie, w celu delimitacji lokalnych rynków pracy w Polsce.. Zgodnie z podaną klasyfikacją algorytm ten reprezentuje grupę metod indukcyjnych, opartych na zasadach. Omawiana procedura znana jest jako Europejski Algorytm Regionalizacji, i rekomendowana przez Eurostat, jako standardowe podejście do definiowania lokalnych rynków pracy w krajach europejskich (por. Eurostat, 1992). Oprócz Wielkiej Brytanii jest ona stosowana w wielu krajach, między innymi Hiszpanii, Włoszech, Danii, Nowej Zelandii oraz w Australii (por. Casado-Diaz, 2000; Sforzi, 1987; Andersen, 2002; Newell, Papps, 2001; Bamber, Walter, 2009).

Europejski algorytm regionalizacji został szczegółowo opisany zarówno przez jego autorów (por. Coombes, Green, Openshaw, 1986) jak i przez część jego użytkowników (por. Casado-Diaz, 2000; Newell, Papps, 2001). W niniejszej pracy przedstawiona zostanie idea jego działania.

W pierwszej fazie omawianej procedury wytypowane zostają te jednostki bazowe, które pełnią rolę tak zwanych ognisk, czyli ośrodków centralnych wokół których skupiać się będą lokalne rynki pracy. Następnie jednostki bazowe przydzielane są do ognisk, z którymi wykazują odpowiedni stopień interakcji. W trakcie działania algorytmu, niektóre z wcześniej wybranych ośrodków centralnych mogą stracić ten status. Również przynależność jednostki bazowej do danego lokalnego rynku pracy może zostać negatywnie zweryfikowana.

Schemat. 2.1. Etapy Europejskiego Algorytmu Regionalizacji



Źródło: Opracowanie własne na podstawie Papps i Newell 2002

Wszystkie kryteria, zarówno służące do wyboru ognisk, jak i do przydzielania jednostek bazowych do lokalnych rynków pracy, są wyliczane tylko i wyłącznie na podstawie danych zawartych w macierzy przepływów związanych z zatrudnieniem (oznaczanej dalej jako T). Czynności te wykonywane są w pięciu etapach opisanych poniżej (por. schemat 2.1).

Etap 1

W pierwszym etapie algorytmu, w celu zdefiniowania zbioru jednostek bazowych, które mogą zostać wstępnie uznane jako ośrodki centralne, sprawdza się dwa kryteria.

Pierwsze z nich określa stosunek (α_{1i}) liczby mieszkańców rozważanej gminy, którzy pracują w dowolnej gminie (w tym również gminie zamieszkania) do ludności pracującej, której miejsce pracy znajduje się na terenie rozważanej gminy. Stosunek (α_{1i}) w przypadku ośrodków centralnych powinien być większy od zadanej wartości α_1 . Wartości α_1 określana jest przez badacza przed rozpoczęciem procedury i zależy od natężenia dojazdów do pracy w kraju, dla którego prowadzona jest delimitacja:

$$\alpha_{1i} = \frac{\sum_{j=1}^N T_{ji}}{\sum_{j=1}^N T_{ij}} > \alpha_1 \quad (2.1)$$

gdzie:

α_{1i} - stosunek liczby mieszkańców i -tej gminy, którzy pracują w dowolnej gminie (w tym również gminie zamieszkania), do ludności pracującej na terenie tej gminy

α_1 - wartość minimalna stosunku α_{1i} dla gmin będących ośrodkami centralnymi,

T_{ij} – liczba ludności zamieszkałej na terenie i -tego obszaru i pracującej na terenie j -tego obszaru (element macierzy przepływów związanych z zatrudnieniem),

N – liczba jednostek bazowych.

Drugie kryterium dotyczy udziału mieszkańców rozważanej gminy, którzy podejmują zatrudnienie na jej terenie, wśród wszystkich pracujących mieszkańców (niezależnie od miejsca pracy). Udział (α_{2i}), powinien przekraczać zadaną wartość α_2 aby dana gmina mogła zostać rozważona jako ośrodek centralny:

$$\alpha_{2i} = \frac{T_{ii}}{\sum_{j=1}^N T_{ij}} > \alpha_2 \quad (2.2)$$

gdzie:

α_{2i} - udział mieszkańców i -tej gminy, podejmujących zatrudnienie na jej terenie, wśród wszystkich pracujących mieszkańców (niezależnie od miejsca pracy),

α_2 - wartość minimalna stosunku α_{2i} dla gmin będących ośrodkami centralnymi,

Etap 2

W drugim etapie sprawdza się stopień interakcji, między wybranymi w etapie pierwszym „wstępnymi” ośrodkami centralnymi. Jeżeli dwa lub więcej z nich, są ze sobą zbyt silnie powiązane, w dalszych rozważaniach traktowane będą jako jeden ośrodek. Status ośrodka centralnego przyznawany jest jednostce najsilniej oddziaływującej, jednak siła oddziaływania jest sumowana. Wszystkie wybrane w etapie pierwszym ośrodki centralne, szeregowane są malejąco ze względu na liczbę ludności dojeżdżającej do nich do pracy. Następnie dla każdego ośrodka centralnego, wybierany jest mniejszy z udziałów: (i) ludności pracującej na terenie tego ośrodka wśród wszystkich jego pracujących mieszkańców (niezależnie od miejsca pracy) oraz (ii) udziału osób mieszkających na terenie tego ośrodka wśród ludności pracującej na jego terenie (por. wzór (2.3)). Minimum to (α_{3i}) powinno przekraczać zadaną wartość α_3 , co świadczy o odpowiednio dużej liczbie mieszkańców zatrudnionych na terenie badanej jednostki bazowej. Wtedy dana jednostka uznawana jest ostatecznie za ośrodek centralny.

$$\alpha_{3i} = \min \left(\frac{T_{ii}}{\sum_{j=1}^N T_{ij}}; \frac{T_{ii}}{\sum_{j=1}^N T_{ji}} \right) > \alpha_3 \quad (2.3)$$

gdzie:

α_{3i} - minimum z udziału ludności pracującej na terenie i -tego ośrodka centralnego wśród wszystkich jego pracujących mieszkańców oraz udziału osób mieszkających na terenie i -tego ośrodka centralnego wśród ludności pracującej na jego terenie,

α_3 - minimalna wartość kryterium α_{3i} dla gmin, które ostatecznie zostaną uznane za ośrodki centralne.

W przeciwnym przypadku, nie spełnienia warunku określonego wzorem (2.3), spośród pozostałych ośrodków centralnych wytypowanych w etapie pierwszym, wybierane są te, które spełniają dwa kolejne warunki. Kryteria te określają nasilenie interakcji z badanym obszarem. Po pierwsze wybierane są te ośrodki centralne, które mają odpowiednio wysoki udział (α_{4ji}) zatrudnionych na terenie ośrodka (j -tego) nie spełniającego poprzedniego kryterium (por. wzór (2.3)), wśród wszystkich swoich pracujących mieszkańców:

$$\alpha_{4ji} = \frac{T_{ij}}{\sum_{k=1}^N T_{ik}} > \alpha_4 \quad (2.4)$$

gdzie:

α_{4ji} - udział zatrudnionych na terenie j -tego ośrodka centralnego wśród wszystkich pracujących mieszkańców i -tej gminy,

α_4 - minimalna wartość kryterium α_{4ji} , świadcząca o występowaniu istotnej interakcji pomiędzy obydwoma obszarami.

W drugiej kolejności, spośród ośrodków centralnych spełniających warunek (2.4) typowane są te, do których dojeżdża odpowiednio wysoki procent mieszkańców ośrodka j -tego (który nie spełnił kryterium (2.3), jednak spełnił kryterium (2.4)) wśród wszystkich pracujących mieszkańców j -tego ośrodka (2.5):

$$\alpha_{5ji} = \frac{T_{ji}}{\sum_{k=1}^N T_{jk}} > \alpha_5 \quad (2.5)$$

gdzie:

α_{5ji} – udział zatrudnionych na terenie i -tego ośrodka centralnego wśród wszystkich pracujących mieszkańców j -tej gminy,

α_5 - minimalna wartość kryterium α_{5ji} , świadcząca o występowaniu istotnej interakcji pomiędzy obydwoma obszarami.

W ten sposób, jeżeli j -ty ośrodek centralny nie spełniał kryterium określonego wzorem (2.3), wytypowano dla niego zbiór ośrodków centralnych o wysokiej z nim interakcji (spełniającej kryteria określone wzorami (2.4) i (2.5)). Jeżeli zbiór ten jest pusty, j -ty ośrodek centralny zachowuje status samodzielnego ośrodka centralnego. W przeciwnym razie ze zbioru tego wybierany jest ośrodek, który charakteryzuje się najwyższą wartością wskaźnika ważonej interakcji (I_{ji}) z ośrodkiem j -tym:

$$I_{ji} = \frac{T_{ij}^2}{\sum_{k=1}^N T_{ik} * \sum_{k=1}^N T_{kj}} + \frac{T_{ji}^2}{\sum_{k=1}^N T_{jk} * \sum_{k=1}^N T_{ki}} \geq \alpha_{6j} \quad (2.6)$$

gdzie:

I_{ji} - wskaźnik ważonej interakcji pomiędzy i -tym oraz j -tym obszarem,

α_{6j} – minimalna wartość wskaźnika ważonej interakcji I_{ji} dla pary obszarów, pomiędzy którymi występuje istotna interakcja.

Jeżeli dwa (lub więcej) ośrodki mają równe wartości wskaźnika ważonej interakcji z ośrodkiem j -tym, wybierany jest ten z nich, dla którego suma przepływów bezwzględnych pomiędzy nim a ośrodkiem j -tym jest większa. Następnie ośrodek j -ty przestaje pełnić rolę samodzielnego ośrodka centralnego i przyłączony zostaje do lokalnego rynku pracy, którego rozwój stymulowany jest przez ośrodek centralny o najwyższej z nim interakcji. W dalszych rozważaniach oddziaływanie obu tych ośrodków jest sumowane.

Następnie można przystąpić do badania kolejnego ośrodka centralnego (pod względem liczby ludności dojeżdżającej do pracy), aż do momentu, kiedy zostaną rozważone wszystkie wytypowane w pierwszym etapie ośrodki centralne. Pomijane są te ośrodki, które wcześniej zostały przyłączone do innych, wyższych w rankingu.

Etap 3

W etapie trzecim rozpoczyna się przydzielanie jednostek bazowych do lokalnych rynków pracy, których rozwój stymulowany jest przez ośrodki centralne zdefiniowane w etapie drugim. W etapie tym nie wszystkie jednostki bazowe muszą zostać rozdzielone. Procedurę rozpoczyna uszeregowanie dotychczas utworzonych lokalnych rynków pracy rosnąco ze względu na wartość funkcji celu (F_j) (por. wzór (2.7)). Warto podkreślić, że indeks j we wzorze (2.7) dotyczy całego wstępnego lokalnego rynku pracy stymulowanego przez dany ośrodek centralny. Zatem T_{jk} oznacza wszystkich mieszkańców jednostek bazowych wchodzących w skład tego lokalnego rynku pracy zatrudnionych na terenie k -tej jednostki bazowej. Pierwszy składnik iloczynu dotyczy więc, łącznej liczby pracujących mieszkańców badanego lokalnego rynku pracy (bez względu na miejsce zatrudnienia). Drugi ze składników iloczynu, zależy od liczby mieszkańców wstępnego lokalnego rynku pracy zatrudnionych na jego terenie. Wielkość ta porównywana jest z maksimum z liczby ludności pracującej na terenie j -tego lokalnego rynku pracy (bez względu na miejsce zamieszkania) oraz liczby pracujących mieszkańców (zatrudnionych w dowolnym obszarze). Zatem wartość funkcji celu (F_j), dla wstępnych lokalnych rynków pracy, których rozwój stymulowany jest danym ośrodkiem centralnym, będzie się zwiększać w miarę dołączania do niego kolejnych jednostek bazowych.

$$F_j = \min \left(1; \frac{c_1 * \sum_{k=1}^N T_{jk}}{\alpha_7}; \frac{c_2 * \sum_{k=1}^N T_{jk} + c_3}{\alpha_7} \right) * \min \left(1; \frac{\alpha_8 * T_{jj}}{\max(\sum_{k=1}^N T_{kj}; \sum_{k=1}^N T_{jk})} \right) \quad (2.7)$$

gdzie:

F_j - wartość funkcji celu,

α_7 - parametr ustalany ze względu na wielkość populacji,

α_8 - parametr zależny od przeciętnej liczby ludności pracującej w swoim obszarze zamieszkania,

c_1, c_2, c_3 - parametry zależne od nasilenia wyjazdów do pracy z ośrodków centralnych

α_9 - minimalna wartość kryterium zapobiegania tworzeniu lokalnych rynków pracy na terenie, na którym mieszka zbyt mało osób pracujących: $F_j > \alpha_9$

Następnie sprawdzany zostaje warunek, czy wartość funkcji celu lokalnego rynku pracy, którego rozwój stymuluje dany (j -ty) ośrodek centralny (por. wzór (2.7)), przekracza zadaną wartość α_9 . Powyższe kryterium ma na celu zapobieganie tworzeniu lokalnych rynków pracy, na terenie których mieszka zbyt mało osób pracujących lub na terenie których pracuje zbyt mało jego mieszkańców. Jeżeli funkcja celu nie przekracza zadanej wartości, dany rynek nie jest dalej rozważany w tym etapie.

W przeciwnym wypadku spośród pozostałych jednostek bazowych wybierane są takie, które spełniają poniżej określone kryterium. W tym etapie rozważane są zarówno jednostki bazowe nie będące, jak i będące ośrodkami centralnymi. Sprawdza się dla nich, czy udział zatrudnionych na terenie j -tego ośrodka centralnego wśród wszystkich pracujących mieszkańców danej jednostki bazowej (por. wzór (2.4)) przekracza wartość α_4 . Jeżeli zbiór ten jest pusty, do lokalnego rynku pracy, którego rozwój stymulowany jest przez j -ty ośrodek centralny nie zostaje dołączona żadna jednostka bazowa. W przeciwnym razie ze zbioru jednostek bazowych spełniających ten warunek wybierana jest ta, która charakteryzuje się najwyższą wartością wskaźnika ważonej interakcji z ośrodkiem j -tym (por. wzór (2.6)). Jeżeli dwie (lub więcej) jednostek bazowych ma równe wartości wskaźnika ważonej interakcji z ośrodkiem j -tym, wybierana jest ta z nich, dla której suma przepływów bezwzględnych jest większa. Następnie wybrana jednostka bazowa zostaje przyłączona do lokalnego rynku pracy, którego rozwój stymuluje j -ty ośrodek centralny. W dalszych rozważaniach suma ich oddziaływania traktowana jest, jako oddziaływanie lokalnego rynku pracy.

Po dokonaniu każdej modyfikacji lokalnego rynku pracy, ponownie przeprowadzany jest ranking nowo otrzymanych lokalnych rynków pracy ze względu na funkcję celu (por. wzór (2.7)). Następnie cała procedura etapu trzeciego zostaje przeprowadzona od nowa. Dzieje się to do momentu, aż do każdego ośrodka centralnego dołączonych zostanie tyle jednostek bazowych, aby otrzymany lokalny rynek pracy charakteryzował się odpowiednio wysoką wartością funkcji celu (α_9) (por. wzór (2.7)). Niektóre z otrzymanych w tym etapie, lokalnych rynków pracy mogą składać się wyłącznie z ośrodka centralnego, choć w etapie następnym zostaną rozbudowane.

Etap 4

W etapie czwartym wszystkie jednostki bazowe zostaną przydzielone do lokalnych rynków pracy. Wyjątek stanowią jedynie jednostki bazowe o bardzo słabej interakcji z otoczeniem (np. gdyby na terenie takiej jednostki był rezerwat przyrody, gdzie nikt nie mieszka ani nie pracuje). Tak więc, wszystkie nie przyporządkowane do tej pory jednostki bazowe, szeregowane są malejąco ze względu na liczbę pracujących mieszkańców (niezależnie od miejsca pracy). Następnie, kolejno dla każdej z nich, spośród lokalnych rynków pracy wybierane są te, dla których suma przepływów związanych z zatrudnieniem w obie strony jest niezerowa. Oznacza to, że rozważane są tylko te lokalne rynki pracy, do których dojeżdża choć jeden²⁴ mieszkaniec rozważanej jednostki bazowej, lub z którego dojeżdża choć jeden pracownik (w praktyce polskiej 9).

Jeżeli, dla pewnej jednostki bazowej, nie występuje niezerowy przepływ związany z zatrudnieniem do żadnego z wyodrębnionych do tego momentu lokalnych rynków pracy, zostaje ona dołączona do listy rezerwowej w celu rozważenia w późniejszym terminie.

Nie jest to równoznaczne z tym, że pomiędzy badaną jednostką a otoczeniem nie ma żadnego przepływu ludności związanego z zatrudnieniem, może być tak, że przepływy takie następują pomiędzy nią a innymi jednostkami bazowymi nieprzyporządkowanymi do tej pory do żadnego z lokalnych rynków pracy. W przeciwnym razie, ze zbioru tych ośrodków centralnych, pomiędzy których lokalnymi rynkami pracy a rozważaną jednostką bazową występują niezerowe przepływy, wybierany jest ten, który charakteryzuje się najwyższą wartością wskaźnika ważonej interakcji (por. wzór (2.6))

²⁴ Ponieważ w ramach badania przepływów ludności związanych z zatrudnieniem prezentowane są wyłącznie przepływy przekraczające dziewięć osób, w praktyce polskiej rozważane są tylko te obszary, do których dojeżdża przynajmniej dziesięć osób.

z badaną jednostką bazową. Jeżeli dwa (lub więcej) lokalne rynki pracy mają równe wartości wskaźnika ważonej interakcji z badaną jednostką bazową, wybierany jest ten z nich, dla którego suma przepływów bezwzględnych jest większa. Następnie rozważana jednostka bazowa zostaje przyłączona do tego lokalnego rynku pracy. W dalszych rozważaniach suma ich oddziaływania jest traktowana, jako oddziaływanie lokalnego rynku pracy.

Po próbie przyłączenia wszystkich jednostek bazowych do lokalnych rynków pracy, rozpatrywane są jednostki z listy rezerwowej, które w momencie wcześniejszej próby ich przyłączenia nie wykazywały przepływów ludności związanych z zatrudnieniem z żadnym lokalnym rynkiem pracy. Ponieważ teraz lokalne rynki pracy się powiększyły, ponowiona zostaje próba ich przyporządkowania do któregoś z nich. W tym celu, dla każdej jednostki bazowej będącej na liście rezerwowej przeprowadza się ponownie procedurę etapu czwartego. Ponadto, próba ta ponawiana jest do momentu, aż nie uda się przyporządkować ani jednej jednostki bazowej pozostającej do danego przebiegu na liście rezerwowej. Jest to gwarancją, że wszystkie jednostki bazowe zostały przyporządkowane do lokalnych rynków pracy, jeśli było to tylko możliwe. Jednostki bazowe, pozostające w dalszym ciągu na liście rezerwowej, albo nie mają żadnych przepływów związanych z zatrudnieniem z innymi obszarami, albo przepływy te następują tylko między nimi samymi.

Etap 5

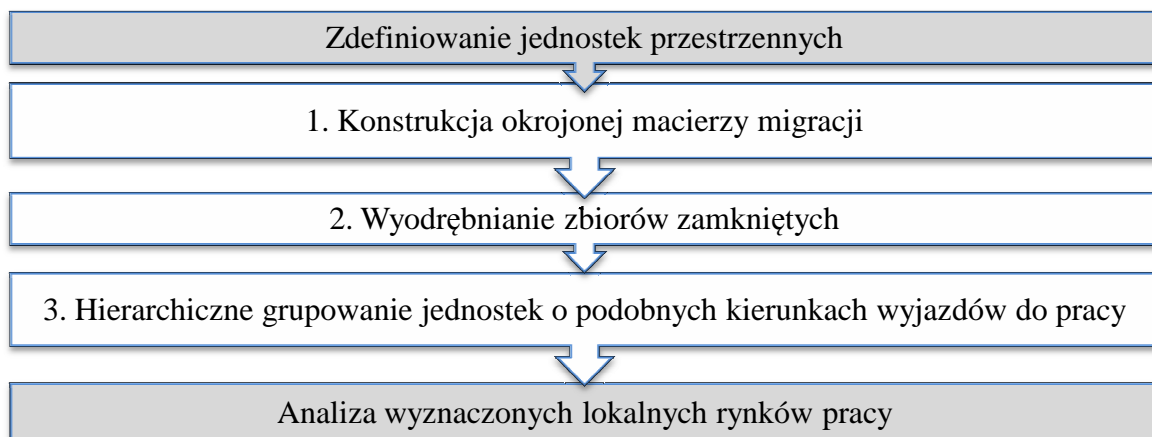
W etapie piątym sprawdzane jest, czy otrzymane lokalne rynki pracy charakteryzują się odpowiednio wysokimi wartościami funkcji celu (por. wzór (2.7)). Podobnie, jak na początku etapu trzeciego, lokalne rynki pracy uszeregowane zostają rosnąco ze względu na tę funkcję. Następnie sprawdza się, czy dla każdego z nich wartość funkcji celu przekracza zadaną wartość graniczną α_9 (w tym wypadku większą niż w etapie 3). Jeżeli któryś z lokalnych ryków pracy nie spełnia tego warunku, zaklasyfikowane do niego jednostki bazowe pozostają się na liście rezerwowej i przydzielone zostaną do pozostałych lokalnych rynków pracy zgodnie z procedurą z etapu czwartego. Następnie powraca się do etapu piątego, aby zweryfikować, czy tak zmodyfikowane lokalne rynki pracy charakteryzują się odpowiednio wysoką wartością funkcji celu. Procedura kończy się wówczas, gdy podczas etapu piątego nie zostaje rozbity żaden lokalny rynek pracy.

Przedstawiony algorytm jest dotychczas jedynym, który znalazł tak szerokie zastosowanie w praktyce. Jest to jego niezaprzeczalną zaletą, gdyż umożliwia wykorzystanie doświadczeń, uzyskanych w różnych innych krajach. Zaletą jest także to, że algorytm europejski umożliwia, zmiany lokalnych rynków pracy. Dopuszcza mianowicie, by obszary połączone na wcześniejszym etapie mogły być w dalszych etapach rozdzielone i dołączone do innych. Za pewną niedogodność można uznać jednak złożoność kolejnych etapów algorytmu i dużą liczbę parametrów wejściowych, potrzebnych dla przeprowadzenia analizy.

2.4. Algorytm Kristensena

W celu porównania przedstawionego algorytmu Coombesa, z metodą hierarchiczną, wybrano podejście zaproponowane przez Kristensena, które jest wykorzystywane do delimitacji lokalnych rynków pracy w Danii (por. Kristensen, 1998). Jest to algorytm indukcyjny, niewykorzystujący ośrodków centralnych, w którym agregacja lokalnych rynków pracy odbywa się w sposób hierarchiczny. Algorytm ten jest znacznie mniej złożony od algorytmu Coombes'a.

Schemat. 2.2. Etapy algorytmu Kristensena



Źródło: Opracowanie własne

Algorytm podzielić można na trzy główne etapy (por. schemat 2.2).

Etap 1

W pierwszym etapie przeprowadzana jest transformacja macierzy przepływów związanych z zatrudnieniem T na macierz binarną TT o takim samym wymiarze. W macierzy TT , istotnym przepływem, w świetle kryterium ilorazu lokacji,

przypisywane są jedynki. Charakteryzują one względną siłę natężenia przepływów ludności związanych z zatrudnieniem. Jedynki przypisywane są również wszystkim elementom nowej macierzy leżącym na diagonalu.

$$TT_{ij} = \begin{cases} 1, & \text{jeżeli } \frac{\frac{T_{ij}}{\sum_{i=1}^N T_{ij}}}{\frac{\sum_{j=1}^N T_{ij}}{\sum_{i=1}^N \sum_{j=1}^N T_{ij}}} > \alpha_1 \quad \forall i = j \\ 0, & \text{w pozostałych przypadkach} \end{cases} \quad (2.8)$$

gdzie:

TT – okrojona macierz przepływów związanych z zatrudnieniem, uwzględniająca wyłącznie istotne przepływy,

T_{ij} – liczba ludności zamieszkałej na terenie i -tej gminy i pracującej na terenie j -tej (element macierzy przepływów związanych z zatrudnieniem),

N – liczba jednostek bazowych,

α_1 - minimalna wartość względnego przepływu pomiędzy gminami dla przepływów uznawanych za istotne.

Etap 2

Okrojona macierz przepływów związanych z zatrudnieniem stanowi podstawę wyodrębnienia tzw. zbiorów zamkniętych. Jako zbiór zamknięty rozumiany jest taki podzbiór jednostek bazowych, z których nie ma istotnych wyjazdów związanych z zatrudnieniem do jednostek spoza obszaru. Zidentyfikowane zbiory zamknięte uznawane zostają za lokalne rynki pracy.

Etap 3

Pozostałe jednostki bazowe łączone są w etapie trzecim w lokalne rynki pracy według kryterium funkcji podobieństwa P_{ij} (por wzór (2.9)). Proces ich łączenia przebiega iteracyjnie. W pierwszej kolejności wyszukiwana jest para jednostek bazowych o największej wartości funkcji podobieństwa (P_{ij}), które nie należą do żadnego lokalnego rynku pracy. Jeżeli wartość funkcji podobieństwa przekracza wartość podobieństwa „pierwszego rzędu” jednostki są ze sobą łączone, a następnie dołączane są do nich wszystkie jednostki bazowe nienależące jeszcze do lokalnych rynków pracy.

Podobieństwo pierwszego rzędu obliczane jest jako $1-\alpha_2$, gdzie α_2 jest wartością z przedziału (0;1) (raczej zbliżoną do 0) określaną przed rozpoczęciem procedury przez badacza, o którą stopniowo obniża się podobieństwo wymagane przy łączeniu jednostek bazowych. Wszystkie połączone jednostki tworzą nowy lokalny rynek pracy.

$$P_{ij} = \frac{\overline{W_i \cap W_j}}{\overline{W_i \cup W_j}} \quad (2.9)$$

gdzie:

P_{ij} – funkcja podobieństwa,

$W_i = \{k; TT_{ik} = 1\}$ – zbiór jednostek bazowych, do których następują istotne wyjazdy związane z zatrudnieniem z jednostki i -tej,

\bar{A} – moc zbioru A , czyli liczba elementów zbioru A .

Jeżeli wśród jednostek bazowych nienależących do lokalnych rynków pracy nie ma żadnej pary charakteryzującej się odpowiednio wysoką wartością funkcji podobieństwa, kryterium podobieństwa zostaje obniżone o α_2 i cała procedura rozpoczyna się od nowa. Postępowanie to powtarzane jest do momentu, kiedy wszystkie jednostki bazowe pogrupowane są w lokalne rynki pracy, lub nie ma już pary jednostek bazowych nienależących do lokalnych rynków pracy, dla których wartość funkcji podobieństwa przekracza α_2 .

Opisany algorytm można zaliczyć do metod hierarchicznych. W algorytmie tym, trakcie całego procesu łączenia obszarów wykorzystywana jest jedna i ta sama zasada. Wśród zalet tych metod wskazać można mniejszą złożoność algorytmu oraz mniejszą liczbę progowych parametrów, które należy ustalić przed rozpoczęciem analizy. Jednakże poważną wadą może być to, że procedury hierarchiczne wykazują tendencję do dostarczania niezupełnie optymalnych wyników przy analizie w większej skali (Casado-Diaz, Coombes, 2011). Wynika to między innymi stąd, że połączenia obszarów, dokonane na wczesnych etapach analizy w istotny sposób ograniczają opcje dostępne w etapach późniejszych.

2.5. Podejście taksonomiczne

Wśród cech obu podejść opisanych powyżej wskazać można zarówno wady jak i zalety. Jako alternatywę zaproponowano więc podejście taksonomiczne, łączące zalety obu metod (por. tab. 2.2). Zaproponowany algorytm można zaliczyć do grupy metod

indukcyjnych, hierarchicznych. Podobnie jak w opisanych algorytmach zaczerpniętych z literatury, jako dane wyjściowe wymagana jest macierz przepływów związanych z zatrudnieniem, zaś rezultatem jest podział jednostek bazowych na rozłączne i pokrywające całą zbiorowość lokalne rynki pracy.

Obydwa podejścia różnią się jednak między innymi pod względem liczby parametrów wejściowych, określanych a priori, które wpływają na wynik delimitacji. Wydaje się, że zbyt duża liczba parametrów nie jest pożądaną cechą tego typu algorytmów, ponieważ ich różne zdefiniowanie może prowadzić do sprzecznych wyników, a także stwarza pole do manipulacji wynikami. Konieczność zdefiniowania aż dwunastu parametrów wejściowych w europejskim algorytmie regionalizacji autorka uznaje za główną wadę tego podejścia. W metodzie zaproponowanej przez Kristensen'a wystarczy określić dwie wartości. W zaproponowanym autorskim podejściu, przyjęć trzeba tylko jeden parametr, który określić można na dwa sposoby. W zależności od tego wyboru, stosowany jest poniżej opisany algorytm z mocnym lub słabym kryterium grupowania.

Kolejną istotną różnicą jest określanie, w europejskim algorytmie regionalizacji, dla każdego lokalnego rynku pracy dokładnie jednego ośrodka centralnego stymulującego jego rozwój. Podejście takie wydaje się być dyskusyjne, stąd brak wyodrębniania ośrodków centralnych można uznać za kolejną przewagę algorytmu Kristensena. W zaproponowanym podejściu taksonomicznym również nie wyodrębnia się ośrodków centralnych.

Z kolei przewagą europejskiego algorytmu regionalizacji, jest niewątpliwie uwzględnianie powiązań, zarówno pomiędzy parami jednostek bazowych oraz pomiędzy jednostkami bazowymi a wstępnymi lokalnymi rynkami pracy. W procedurze zaproponowanej przez Kristensena uwzględniane są wyłącznie związki pomiędzy parami jednostek bazowych, co należy traktować, jako istotną wadę tego podejścia. W proponowanej metodzie taksonomicznej, uwzględniane są zależności pomiędzy pojedynczą jednostką bazową a innymi jednostkami i grupami jednostek oraz pomiędzy dwoma wstępnie utworzonymi lokalnymi rynkami pracy.

Kolejną zaletą europejskiego algorytmu regionalizacji, jest dynamiczny charakter budowania lokalnych rynków pracy, pozwalający w każdym momencie na dołączenie do już istniejącego lokalnego rynku pracy kolejnej jednostki bazowej, a także na rozbijanie w etapie piątym wcześniej zbudowanych lokalnych rynków pracy.

Poważnym brakiem algorytmu Kristensen’a jest zaś, dość kontrowersyjna kolejność budowania lokalnych rynków pracy. Mianowicie, na określonym poziomie podobieństwa może nie być możliwości dołączenia kolejnej jednostki bazowej do już istniejącego lokalnego rynku pracy podczas, kiedy jednocześnie na tym samym poziomie podobieństwa istnieje możliwość zbudowania nowego lokalnego rynku pracy. W autorskim podejściu taksonomicznym, w każdym momencie trwania procedury możliwe jest poszerzenie każdego ze wstępnie wyznaczonych lokalnych rynków pracy. Nie ma, co prawda możliwości rozbicia raz połączonych jednostek bazowych, jednak możliwe jest połączenie dwóch, wcześniej traktowanych oddzielnie, wstępnych lokalnych rynków pracy.

Tabela 2.2. Porównanie europejskiego algorytmu regionalizacji, algorytmu Kristensena oraz podejścia taksonomicznego

Lp	Cecha:	Europejski algorytm regionalizacji	Algorytm Kristensena	Podejście taksonomiczne
1	Dane wejściowe	Macierz przepływów związanych z zatrudnieniem		
2	Liczba definiowanych parametrów	12	2	1
3	Ośrodki centralne	Wyłaniane w trakcie działania	brak	brak
4	Uwzględniane rodzaje powiązań	Pomiędzy pojedynczą jednostką bazową a innymi jednostkami oraz grupami jednostek	Wyłącznie pomiędzy jednostkami bazowymi	Pomiędzy pojedynczą jednostką bazową a innymi jednostkami i grupami jednostek oraz pomiędzy dwoma wstępnie utworzonymi lokalnymi rynkami pracy
5	Procedura delimitacji	Równoległe budowanie lokalnych rynków pracy	Hierarchiczne budowanie lokalnych rynków pracy	Równoległe budowanie lokalnych rynków pracy
6	Rezultat	Podział jednostek bazowych na rozłączne i pokrywające całą zbiorowość lokalne rynki pracy		

Źródło: Opracowanie własne

Taksonomia jest dziedziną statystycznej analizy wieloczynnikowej, która zajmuje się teoretycznymi zasadami i regułami klasyfikacji obiektów wielocechowych. (por. Sokołowski, 1992; Strahl D., red, 1998). W analizie taksonomicznej występują dwa główne zadania badawcze: zadanie grupowania operacyjnych jednostek taksonomicznych (ang. cluster analysis) oraz zadanie liniowego porządkowania obiektów (por. Sokołowski, 1992).

W procesie delimitacji zbiorów jednostek bazowych dzielony jest na rozłączne i pokrywające całość lokalne rynki pracy. Łatwo zauważyć więc, że jeżeli zbiór jednostek bazowych potraktujemy jako zbiór operacyjnych jednostek taksonomicznych, zaś możliwe wyniki delimitacji jako niepustą rodzinę jego podzbiorów spełnione są obydwie warunki klasyfikacji taksonomicznej (por. Sokołowski 1992). Zgodnie z tą obserwacją opisaną w tym podrozdziale metodę delimitacji określić można podejściem taksonomicznym. Należy przy tym jednak zauważyć, że w zagadnieniach taksonomicznych postuluje się, aby przestrzeń klasyfikacji była przestrzenią metryczną, zatem w szczególności kryterium według którego łączone są jednostki taksonomiczne powinno charakteryzować się własnościami metryki (por. Sokołowski, 1992). Stosowane w tym celu w zaproponowanym podejściu kryterium natężenia dojazdów do pracy nie jest jednak metryką.

W ramach podejścia taksonomicznego zaproponowano dwie wersje algorytmu. Pierwsza z nich charakteryzuje się mocnym kryterium (2.11), które musi być spełnione, aby dwa wstępne lokalne rynki pracy zostały połączone (por. Gruchociak, 2012d). W drugim wariantcie, warunek ten jest mniej restrykcyjny i ma postać (2.12). Można zauważyć, że w wyniku zastosowania słabszego kryterium łączenia, uzyskana zostanie mniejsza liczba lokalnych rynków pracy.

W obu wersjach algorytm opiera się na łączeniu jednostek bazowych, pomiędzy którymi występuje silny przepływ ludności związany z zatrudnieniem. Rozważane są więc, pary jednostek bazowych, uszeregowane malejąco według względnej liczby wyjazdów WT_{ij} z pierwszej jednostki do drugiej (por. wzór (2.10)).

$$WT_{ij} = \frac{T_{ij}}{\sum_{j=1}^N T_{ij}} \quad (2.10)$$

gdzie:

WT_{ij} – udział zatrudnionych na terenie j -tej jednostki bazowej wśród wszystkich pracujących mieszkańców i -tej jednostki bazowej (bez względu na miejsce zatrudnienia),

T – macierz przepływów związanych z zatrudnieniem.

Dla każdej, kolejno rozważanej pary jednostek bazowych możliwe są cztery sytuacje:

1. Obie jednostki bazowe nie zostały wcześniej włączone do żadnego ze wstępnych lokalnych rynków pracy. Wówczas zostają one połączone tworząc nowy lokalny rynek pracy.
2. Jednostka, z której następuje wyjazd, nie została jeszcze przyłączona do żadnego wstępnego lokalnego rynku pracy. A jednostka, do której wyjazd jest skierowany należy już do wstępnego lokalnego rynku pracy. Wówczas jednostka bazowa, z której następuje wyjazd jest przyłączana do lokalnego rynku pracy.
3. Jednostka wyjazdu należy już do wstępnego lokalnego rynku pracy a jednostka, do której wyjazd następuje jeszcze nie. W takim wypadku jednostki nie są ze sobą łączone.
4. Obie jednostki bazowe należą już do wstępnych lokalnych rynków pracy. W takiej sytuacji lokalne rynki pracy są ze sobą łączone, jeżeli spełnione są dwa poniżej określone warunki. W zależności od zastosowania, w pierwszym z warunków mocnego lub słabego kryterium łączenia, algorytm będzie się charakteryzował mniejszą lub większą skłonnością do łączenia lokalnych rynków pracy.

Przy pomocy pierwszego warunku sprawdza się, czy istnieją przesłanki, aby lokalny rynek pracy, z którego następuje wyjazd, przyłączyć do innego lokalnego rynku pracy. Sytuacja taka występuje, jeżeli wstępne lokalne rynki pracy nie charakteryzują się dostatecznie wysokim procentem mieszkańców zatrudnionych na ich terenie. W przypadku zastosowania mocnego kryterium łączenia sprawdza się, czy więcej mieszkańców lokalnego rynku pracy, z którego następuje wyjazd pracuje na terenie lokalnego rynku pracy, do którego jest on skierowany niż na jego terenie (por. wzór (2.11)).

$$\sum_{i=1}^{\bar{A}} \sum_{j=1}^{\bar{B}} T_{A_i B_j} > \sum_{i=1}^{\bar{A}} \sum_{j=1}^{\bar{A}} T_{A_i A_j} \quad (2.11)$$

gdzie:

A – zbiór jednostek bazowych zaliczonych do tego wstępnego lokalnego rynku pracy,
co jednostka, z której następuje wyjazd,

B - zbiór jednostek bazowych zaliczonych do tego wstępnego lokalnego rynku pracy,
co jednostka, do której skierowany jest wyjazd,

A_i – i -ty element zbioru A .

W przypadku zastosowania słabego kryterium łączenia wystarczy, aby liczba mieszkańców lokalnego rynku pracy, z którego następuje wyjazd, pracujących poza jego granicami, przewyższała liczbę mieszkańców zatrudnionych na jego terenie (por. wzór (2.12)).

$$\sum_{i=1}^{\bar{A}} \sum_{k=1}^l \sum_{\substack{j=1 \\ A^k \neq A}}^{\bar{A}^k} T_{A_i A_j^k} > \sum_{i=1}^{\bar{A}} \sum_{j=1}^{\bar{A}} T_{A_i A_j} \quad (2.12)$$

gdzie:

l - liczba lokalnych rynków pracy w danym momencie działania algorytmu,

A^k – zbiór jednostek bazowych zaliczonych do k -tego wstępnego lokalnego rynku pracy, $k=1, \dots, l$.

Następnie sprawdza się, czy oddziaływanie lokalnego rynku pracy, do którego następuje wyjazd, na lokalny rynek pracy z którego on pochodzi jest możliwie największe. W tym celu porównuje się liczbę osób wyjeżdżających do pracy z lokalnego rynku pracy wyjazdu do pozostałych lokalnych rynków pracy (por. wzór (2.13)).

$$\frac{\sum_{i=1}^{\bar{A}} \sum_{j=1}^{\bar{B}} T_{A_i B_j}}{\bar{B}} = \max_k \frac{\sum_{i=1}^{\bar{A}} \sum_{j=1}^{\bar{A}^k} T_{A_i A_j^k}}{\bar{A}^k} \quad (2.13)$$

Powyżej opisana procedura kontynuowana jest do momentu, aż rozważone zostaną wszystkie pary jednostek bazowych, pomiędzy którymi odnotowano niezerowy przepływ ludności związany z zatrudnieniem.

Zaproponowany algorytm łączy zalety procedury Coombes'a, takie jak uwzględnianie powiązań zarówno pomiędzy jednostkami centralnymi jak i ich grupami oraz równoległe budowanie lokalnych rynków pracy, z zaletami metody Kristensena, za jakie można uznać małą liczbę parametrów wejściowych, brak konieczności uwzględniania ośrodków centralnych i mniejszą złożoność algorytmu.

Wyniki zastosowania dwóch powyżej opisanych algorytmów, zaczerpniętych z literatury oraz autorskiego podejścia taksonomicznego w warunkach polskich przedstawione zostaną w rozdziale piątym. Do przeprowadzenia delimitacji lokalnych rynków pracy według każdej z tych metod potrzebne są informacje o dojazdach do pracy. Dane takie w Polsce dostępne są na chwilę obecną wyłącznie dla roku 2006, otrzymane wyniki miałyby być zaś aktualne także dla roku 2011. W związku z tym dwa kolejne rozdziały poświęcone zostaną oszacowaniu danych dotyczących dojazdów do pracy, przy czym pierwszy z nich dotyczyć będzie zastosowanej w tym celu metodologii, zaś kolejny jej aplikacji oraz uzyskanych wyników.

Rozdział 3. Modelowanie wielopoziomowe

3.1. Idea, korzyści i założenia modelowania wielopoziomowego

HISTORIA

Idea modelowania wielopoziomowego zrodziła się na początku lat 70-tych, kiedy to zwrócono uwagę na zróżnicowanie poziomu nauczania w klasach oraz szkołach. D. Lindley i A. Smith opracowali wówczas ogólne ramy dla badania zagnieżdżonych danych o złożonych strukturach błędów losowych (por. Lindley, Smith, 1972). Z kolei opracowanie A. Dempstera, N. Lairda i D. Rubina (1977) stanowiło istotny przełom w estymacji parametrów. Niestety, badania te zostały na pewien czas odłożone, ze względu na zbyt dużą złożoność obliczeniową. Następne opracowania dotyczące modelowania wielopoziomowego powstały dopiero w latach 80-tych.

NAZEWNICTWO

Omawiana metoda modelowania wielopoziomowego w literaturze przedmiotu funkcjonuje pod wieloma nazwami. Tak więc, opisywana jest pod nazwą liniowego modelowania wielopoziomowego (multilevel linear models) (por. Goldstein 2003), modeli efektów mieszanych bądź losowych (mixed-effects models, random-effects models) (por. Laird, Ware, 1982, Singer, 1998, Littell, Milliken, Stroup, Wolfinger, 1996), modeli losowego współczynnika regresji (random-coefficient regression models) (por. Longford, 1993; de Leeuw & Kreft, 1986) oraz modeli elementów kowariancji (covariance components models) (por. Dempster, Rubin, Tsutakawa 1981; Longford, 1987). Również modelowanie hierarchiczne (hierarchical linear models) (por. Lindley, Smith, 1972; Raudenbush, Bryk, 2002), nazywane także hierarchicznymi metodami Baysowskimi (hierarchical Bayes Method) (por. Rao, 2003), zalicza się do metod modelowania wielopoziomowego, jednak w zakresie ograniczonym do poziomów zagnieżdżanych, nie obejmując zaś modeli wielopoziomowych z dwoma kryteriami grupowania. Z kolei model wielopoziomowy z dwoma kryteriami grupowania (por. Biecek, 2011) funkcjonuje w literaturze pod nazwą modelu krzyżowego (cross model) (por. Bates, 2013a; Hox, 2010). Taki rodzaj modelu omówiony zostanie w podrozdziale 3.2. Ponadto wśród modeli krzyżowych wyróżnia się modele częściowo i całkowicie krzyżowe (partially/completely crossed). Modele częściowo krzyżowe, to takie, w których przekrój jednostek drugiego poziomu, otrzymanych ze względu na różne

kryteria grupowania, może zawierać dowolną liczbę jednostek pierwszego poziomu. Natomiast modelami całkowicie krzyżowymi określane są takie, w których przekrój jednostek drugiego poziomu zawierać może co najwyżej jedną jednostkę pierwszego poziomu (por. Bates, 2013a). W podrozdziale 3.3. omówiona zostanie teoria modelu całkowicie krzyżowego.

Warto zauważyć, że w literaturze, zwłaszcza dotyczącej modeli dwupoziomowych, jako poziomy (levels) rozumiane są czasem warianty kryterium grupowania. W prezentowanej pracy warianty kryterium grupowania określane są, jako jednostki drugiego oraz wyższych poziomów. Jako kolejne poziomy rozumie się zaś, populacje grup jednostek badania określonych przez kryterium grupowania odpowiadające temu poziomowi (czyli zbiorowości jednostek danego poziomu).

IDEA I KORZYŚCI

Metodologia modelowania wielopoziomowego pozwala na uwzględnienie podobieństwa pomiędzy jednostkami badania z pierwszego poziomu, które należą do tych samych grup utworzonych poprzez zmienną grupującą na drugim lub wyższych poziomach. W przeciwieństwie do klasycznej regresji liniowej nie zakłada się, że wszystkie obserwacje są niezależne; uwzględniona zostaje bowiem zależność pomiędzy jednostkami pierwszego poziomu, należącymi do tej samej jednostki na poziomie wyższym. W przypadku nie uwzględnienia takiej zależności standardowe błędy szacunku są niedoszacowane (por. Klimanek, 2003; Hox, 2002). Tak więc, w modelach wielopoziomowych obserwowane są dwa rodzaje zmienności: zróżnicowanie pomiędzy jednostkami pierwszego poziomu, należącymi do tych samych jednostek poziomów wyższych oraz zróżnicowanie pomiędzy jednostkami wyższych poziomów (por. Frątczak, Mianowska, 2012). Uwzględnienie zależności pomiędzy jednostkami poziomu pierwszego należącymi do tych samych jednostek na poziomach wyższych pozwala w znacznym stopniu poprawić precyzję szacunku w stosunku do klasycznej regresji liniowej, pod warunkiem, że szacowana zmienna charakteryzuje się strukturą wielopoziomową. Jest tak zarówno wówczas, gdy w klasycznym modelu regresji podział jednostek pierwszego poziomu nie zostaje uwzględniony w żaden sposób, ale również, jeżeli obserwacje należące do tych samych grup są traktowane jako całość (i oszacowania prowadzone są dla całych grup) (por. Twisk, 2010, rozdział 3).

Ponadto, dzięki podzieleniu badanej populacji na grupy, możliwe jest wyjaśnienie części zmienności szacowanej cechy przy pomocy charakterystyk z wyższych poziomów, nie popełniając przy tym błędu nazywanego paradoksem Simpsona (Simpson's paradox). Błąd ten polega na łączeniu ze sobą danych z niejednorodnych populacji i traktowaniu ich tak, jakby odnosiły się do jednej zbiorowości (por. Klimanek, 2003, Hox, 1995). Potrzebę agregacji informacji dostępnych na różnych poziomach podkreślają w swojej pracy również T. Bołt, K. Krauze i T. Kulawczuk (1985).

We wcześniejszych pracach (por. Gruchociak 2012a, 2012b) autorka podjęła próbę oceny przydatności modelowania wielopoziomowego w szacowaniu charakterystyk społeczno-gospodarczych w przekroju terytorialnym. W pracach tych zastosowano modele dwupoziomowe w celu oszacowania wskaźnika zatrudnienia w przekroju powiatów. Dzięki zastosowaniu metodologii modelowania dwupoziomowego udało się uwzględnić zróżnicowanie poziomu badanej cechy pomiędzy województwami, co nie byłoby możliwe przy zastosowaniu klasycznej funkcji regresji liniowej. Poprawiło to znacznie precyzję szacunku.

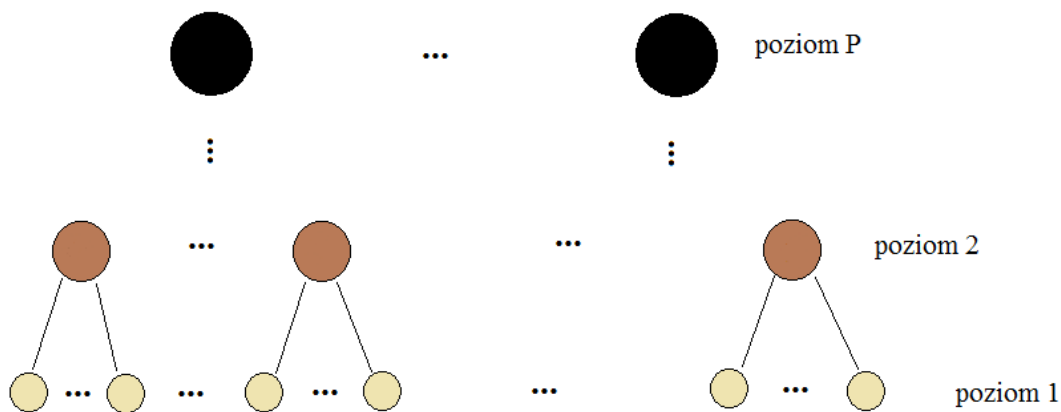
ZAŁOŻENIA

Należy podkreślić, że stosowanie metodologii modelowania wielopoziomowego jest uzasadnione tylko dla specyficznych zbiorowości oraz zmiennych. Zarówno populacja jak i zmienna winny charakteryzować się wielopoziomową strukturą. W odniesieniu do populacji oraz klasycznego modelu wielopoziomowego oznacza to, że populację podzielić można na skończoną liczbę rozłącznych i pokrywających wszystkie jednostki pierwszego poziomu grup (inaczej jednostek drugiego poziomu).

W przypadku modelu więcej niż dwupoziomowego, jednostki drugiego poziomu mogą również zostać podzielone na rozłączne i pokrywające całą zbiorowość grupy (nazywane dalej jednostkami trzeciego poziomu) itd. (por. schemat 3.1).

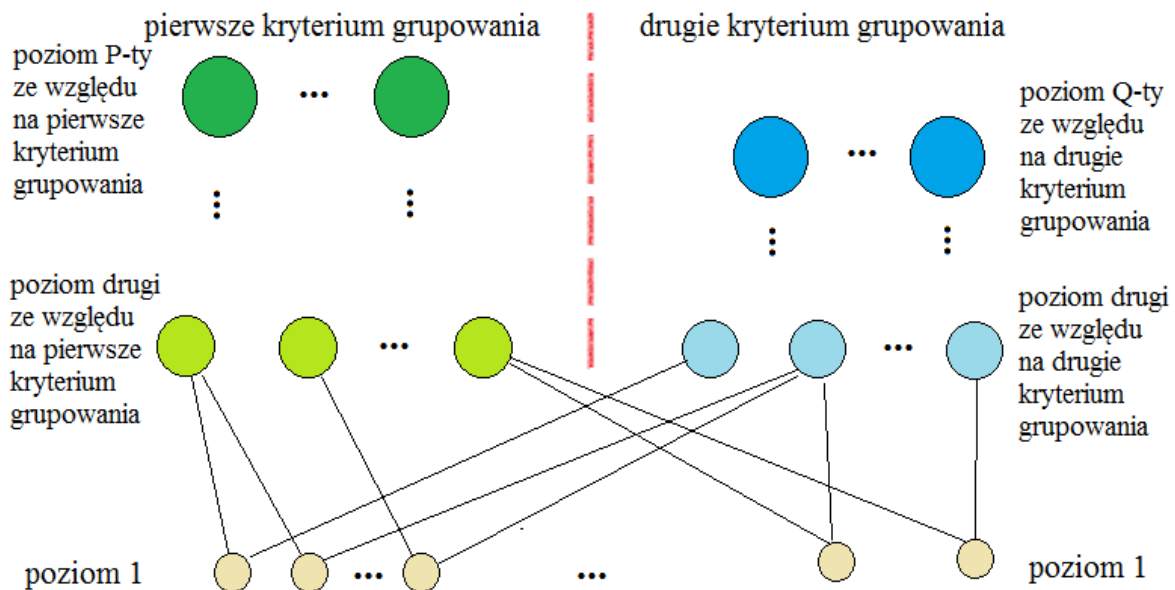
W przypadku modelu z dwoma kryteriami grupowania (tzw. cross model) jednostki drugiego poziomu zostają zdefiniowane dwukrotnie, przy czym każda jednostka pierwszego poziomu należy do dokładnie jednej jednostki poziomu drugiego określonej ze względu na każde z dwóch kryteriów podziału. Dla jednostek drugiego poziomu, zdefiniowanych przez oba kryteria, określone mogą zostać również jednostki wyższych poziomów (por. schemat 3.2).

Schemat 3.1. Schemat jednostek klasycznej populacji o strukturze P-poziomowej



Źródło: Opracowanie własne

Schemat 3.2. Schemat jednostek populacji o strukturze P oraz Q poziomowej z dwoma kryteriami grupowania



Źródło: Opracowanie własne

Należy podkreślić, że modele wielopoziomowe z dwoma kryteriami grupowania zaliczane są do modeli wielopoziomowych. Jednak modele wielopoziomowe wywodzą się od modeli dla danych o strukturze hierarchicznej i często w ten sposób są rozumiane (por. Raudenbush, Bryk, 2002), zaś modele dla danych krzyżowych traktowane są raczej jako ich rozszerzenie. Stąd w prezentowanej pracy, ze względu na przejrzystość zapisu, modele wielopoziomowe dla danych o strukturze hierarchicznej nazywane

są „klasycznymi”, co może być dyskusyjne, ponieważ można je traktować po prostu jako modele z jednym kryterium grupowania (por. Bates, 2010; Biecek, 2011).

Stosowanie modelu wielopoziomowego jest uzasadnione tylko wówczas, jeżeli także szacowana zmienna charakteryzuje się wielopoziomową strukturą. Oznacza to, że jej wartość powinna różnić się istotnie pomiędzy grupami, czyli jednostkami na każdym z wyższych poziomów. Zróżnicowanie to wynikać może z bezpośredniej zależności pomiędzy badaną zmienną a przynależnością jednostki pierwszego poziomu do jednostek wyższych poziomów. Klasycznym, podawanym w literaturze przykładem takiej sytuacji jest zróżnicowanie poziomu nauczania wynikające zarówno z indywidualnych zdolności i predyspozycji ucznia (czynników na pierwszym – jednostkowym poziomie) oraz kwalifikacji nauczyciela i stosowanej metody nauczania (czynniki na poziomie drugim – grupowym) (por. Hox, 2002; Cronbach, Webb, 1975; Goldstein, 2003).

Jako inną przyczynę zróżnicowania poziomu badanej zmiennej pomiędzy grupami, wskazać można zależność tej zmiennej oraz podziału na grupy z pewną ukrytą, często niemierzalną, zmienną. Jako przykład można podać relację między aktywnością zawodową i stopniem rozwoju regionów. Przestrzenne zróżnicowanie czynników określających popyt na pracę, lokalizacja zasobów naturalnych, zakładów produkcyjnych, rozwój infrastruktury technicznej, komunikacyjnej, edukacyjnej są istotnymi determinantami określającymi aktywność ekonomiczną ludności, obok czynników charakteryzujących przedsiębiorczość poszczególnych osób.

Jeżeli struktura zmiennej objaśnianej spełnia obydwie opisane wyżej warunki konieczne; występowanie wielopoziomowej struktury populacji oraz przesłanek aby podejrzewać, że poziom badanej zmiennej jest zróżnicowany pomiędzy grupami; należy zweryfikować, czy występujące pomiędzy grupami zróżnicowanie jest istotne statystycznie na obranym poziomie istotności. W tym celu zastosować można np. test analizy wariancji.

Kolejne założenie dotyczy normalności rozkładu szacowanej zmiennej, przy czym wartość oczekiwana różni się pomiędzy jednostkami wyższych poziomów, co jest konsekwencją wielopoziomowej struktury zmiennej, jednak wariancja powinna być stała w całej populacji. Jak podaje w swojej pracy J.Hox, przy dużych próbach szacowanie parametrów modelu, metodą największej wiarygodności, daje stabilne wyniki nawet

jeżeli rozkład szacowanej zmiennej nieco odbiega od normalnego (por. Hox, 2010, rozdział 3).

Należy także podkreślić, że w rozprawie pod pojęciem modelu wielopoziomowego rozumiany jest liniowy model wielopoziomowy. Tak więc przyjmuje się, że występuje liniowa zależność pomiędzy zmienną objaśnianą a zmiennymi objaśniającymi (która może mieć jednak zróżnicowany charakter dla jednostek niższego poziomu należących do różnych jednostek na poziomie wyższym). W literaturze funkcjonują także opisy modeli wielopoziomowych o charakterze nieliniowym, jednak nie są one przedmiotem tej rozprawy (por. Rydlewski, 2009).

W przypadku estymowania zmiennych w populacji o wielopoziomowej strukturze, zastosowanie odpowiednio dobranego modelu w istotnym stopniu poprawia precyzję szacunku, pod warunkiem, że szacowana zmienna charakteryzuje się również wielopoziomową strukturą (por. Goldstein, 2003; Hox, 2010; Raudenbush i Bryk, 2002; Twisk, 2010).

PROBLEMY

Należy zauważyć, że modele więcej niż dwu poziomowe charakteryzują się tendencją do szybkiego komplikowania się. Nie chodzi tu tylko o trudność w interpretacji, ale przede wszystkim trudność oszacowania parametrów, nawet przy wykorzystaniu nowoczesnego oprogramowania (por. Hox, 2010; DiPrete, Forristal, 1994, s. 349). Niemniej jednak, w analizie wielopoziomowej występują modele o trzech i więcej poziomach. W przypadku takich modeli, należy być szczególnie ostrożnym przy wprowadzaniu kolejnych komplikacji. Dotyczy to zwłaszcza uwzględniania interakcji pomiędzy zmiennymi objaśniającymi z różnych poziomów, czyli inaczej wyjaśnianiu części zmienności współczynników kierunkowych przy zmiennych objaśniających z poziomów niższych, przy pomocy charakterystyk wyższych poziomów. Interakcje takie należy wprowadzać wyłącznie, jeżeli istnieją silne przesłanki merytoryczne, aby twierdzić, że wartość zmiennej z poziomu wyższego faktycznie wyjaśnia charakter zależności pomiędzy zmienną objaśnianą a objaśniającą z niższego poziomu lub, jeżeli oszacowanie odpowiedniego współczynnika kierunkowego charakteryzuje się szczególnie dużą wariancją (por. Hox, 2010).

CENTROWANIE ZMIENNYCH

Jak wskazuje się w literaturze przedmiotu, przed rozpoczęciem oszacowań należy przeprowadzić centrowanie zmiennych objaśniających (ang. centering explanatory variables) odnoszących się do wszystkich poziomów (por. Hox, 2010; Bates, 2010). Umożliwia to interpretację wyrazu wolnego oraz współczynników kierunkowych przy interakcjach pomiędzy zmiennymi z różnych poziomów. Przede wszystkim jednak, pozwala zminimalizować wariancję części modelu, w przypadku modelu z losowym współczynnikiem kierunkowym. Zwróćmy uwagę, że w przypadku modelu wielopoziomowego z losowymi współczynnikami kierunkowymi, w przeciwieństwie do klasycznej regresji liniowej, składnik losowy może zależeć od niektórych zmiennych objaśniających. Kolejną przesłanką do stosowania zmiennych scentrowanych jest dodatni wpływ na zbieżność oszacowań (por. Hox, 2010).

W literaturze opisywane są dwa główne podejścia do centrowania zmiennych objaśnianych. W pierwszym z nich, od wartości zmiennych objaśnianych odejmuje się średnią globalną. W drugim natomiast, średnią dla jednostki wyższego poziomu. Stosowanie drugiego podejścia znajduje uzasadnienie w teorii nazywanej „efektem stawu”. Mianowicie, żaba przeciętnej wielkości może wydawać się duża, jeżeli żyje w stawie z małymi żabami, lub mała, jeżeli żyje w stawie, w którym pozostałe żaby są od niej większe. Takie uwzględnienie lokalnego punktu odniesienia może znaleźć uzasadnienie w przypadku wielu badań, np. związanych z edukacją (por. Hox, 2010). Z drugiej jednak strony odjęcie średnich grupowych powoduje zniwelowanie zróżnicowania poziomu zmiennej pomiędzy jednostkami należącymi do różnych grup na poziomie wyższym, co w oczywisty sposób wpływa negatywnie na własności modelu (por. Kreft, de Leeuw, Aiken, 1995; Enders, Tofighi, 2007; Hofmann, Gavin, 1998; Paccagnella, 2006). Utraconą informację można „odzyskać” dodając do modelu średnie dla tych jednostek wyższego poziomu, jednak w znacznym stopniu zwiększa to złożoność modelu (por. Hox, 2010; Enders, Tofighi, 2007). Kolejnym problemem przy centrowaniu metodą wykorzystującą średnią grupową, w przypadku modelu więcej niż dwupoziomowego jest określenie, ze względu na przynależność do jednostek którego z wyższych poziomów należy te średnie wyznaczyć.

W prezentowanej pracy zastosowano podejście pierwsze, czyli od zmiennych objaśniających odjęto ich średnie globalne.

SPOSOBY KONSTRUOWANIA MODELU

W literaturze prezentowane są dwa podejścia do konstrukcji modeli wielopoziomowych (por. Frątczak, Mianowska, 2012, str. 246; West, Welch, Gałęcki, 2007). W pierwszym z nich, nazywanym „góra-dół” (ang. top-bottom), najpierw dodawane są zmienne objaśniające, a dopiero później efekty losowe zależne od przynależności jednostek pierwszego poziomu do jednostek wyższych poziomów. W następnej kolejności z modelu usuwane są te zmienne, które po uwzględnieniu efektów losowych związanych z przynależnością do jednostek wyższych poziomów nie poprawiają, w istotnym stopniu, jakości dopasowania modelu (por. Verbeke, Molenberghs, 2000). Drugie podejście, nazywane „krok wyżej” (ang. step-up), zakłada z kolei dodawanie do modelu najpierw efektów losowych, a dopiero następnie zmiennych objaśniających (por. Raudenbush, Bryk, 2002, Snijders, Bosker, 1999, Biecek, 2011, Bates, 2010). Różne możliwości stopniowego komplikowania modelu w przypadku struktury dwupoziomowej w swojej pracy szczegółowo omawiają S. Raudenbush i A. Bryk (por. Raudenbush, Bryk, 2002). Podobnie jak w podejściu „góra-dół”, przy każdej modyfikacji sprawdza się, czy poprawia ona, w istotnym stopniu, jakość dopasowania modelu. W pracy omówione zostały etapy komplikowania modelu wielopoziomowego zgodne z podejściem szczegółowo omówionym dla struktury dwupoziomowej w pracach J.Hox’a oraz P.Bliese (por. Hox, 2010; Bliese, 2012).

Konstruując modele wielopoziomowe w rozdziale czwartym, zastosowano podejście „krok wyżej” W trakcie konstrukcji modelu przeprowadzono także weryfikację przewagi podejścia wielopoziomowego nad klasyczną regresją liniową. Tak więc, punktem wyjścia będzie klasyczny model regresji liniowej nie uwzględniający wielopoziomowej struktury danych. Konstrukcję modelu wielopoziomowego rozpocznie ujęcie, w którym zmienność cechy objaśnianej wyjaśniana jest tylko i wyłącznie przez przynależność jednostek poziomu pierwszego do jednostek poziomów wyższych. W etapie końcowym zaprezentowany zostanie model uwzględniający wpływ losowego wyrazu wolnego oraz losowych współczynników, zależnych od zmiennych objaśniających z wyższych poziomów. Warto w tym miejscu zaznaczyć, że końcowa postać modelu dwupoziomowego, w przypadku konstrukcji „krok wyżej”, jest zazwyczaj taka sama dla dostępnych w literaturze opracowań (por. Raudenbush, Bryk, 2002; Węziak, 2007; Bliese, 2012). Jednak zaobserwować można różne podejścia w stopniowym komplikowaniu modelu. Według jednej z propozycji, zmienne objaśniające z wyższych

poziomów zostają dołączone do modelu przed zmiennymi objaśniającymi z poziomu pierwszego (por. Klimanek, 2003; Raudenbush, Bryk, 2002). W niniejszym opracowaniu omówiono natomiast podejście, w którym najpierw dodaje się zmienne objaśniające z poziomu pierwszego (por. Bliese, 2012; Twisk, 2010). Taki sposób postępowania uznano za bardziej intuicyjny i lepiej odpowiadający sytuacjom rzeczywistym.

Stopniowe komplikowanie modelu ma na celu uniknięcie uwzględniania losowego charakteru parametrów modelu, jeżeli nie poprawia to, w istotny sposób, dopasowania modelu do konkretnych danych rzeczywistych. Dlatego też, omówiono kilka wybranych kryteriów, pozwalających ocenić poprawę jakości oszacowania modelu, w porównaniu z modelem prostszym (z poprzedniego etapu). W zastosowaniach praktycznych należy po każdej komplikacji modelu zweryfikować, czy wprowadzona zmiana w sposób istotny poprawiła jakość modelu. Tak więc, jeżeli stwierdzone zostanie, że realizacja któregoś z etapów nie poprawia, w świetle określonych kryteriów, jakości modelu, należy ten etap pominąć i przejść do następnego (np. z etapu 1 do 3). Najczęściej, wprowadzenie komplikacji dla określonego poziomu poprawia precyzję szacunku tylko w przypadku niektórych wprowadzanych zmiennych objaśniających. W takiej sytuacji należy uwzględnić tylko te zmiany, dzięki którym uzyskano poprawę precyzji szacunku. Weryfikacja zasadności wprowadzania kolejnych komplikacji, zgodnych z przedstawianą metodologią modelowania wielopoziomowego jest niezwykle istotna, ponieważ niektóre wprowadzane do modelu zmienne, mogą w rzeczywistości nie mieć struktury wielopoziomowej.

3.2. Klasyczny model wielopoziomowy

W podrozdziale 3.2 omówione zostaną etapy konstrukcji modelu opartego na danych o klasycznej wielopoziomowej strukturze (por. schemat. 3.1). Jak wspomniano w poprzednim podrozdziale zastosowane zostanie podejście omówione w literaturze dla modelu dwupoziomowego (por. Hox, 2010; Bliese, 2012), jednak zostanie ono uogólnione na model o dowolnej liczbie poziomów. Konstrukcja modelu odbywa się w pięciu etapach, które w przypadku modelu więcej niż dwupoziomowego, dzielą się na podetapy.

- I. W etapie pierwszym zróżnicowanie zmiennej objaśnianej tłumaczone jest przez przynależność jednostek pierwszego poziomu do jednostek poziomów wyższych.

Etap pierwszy składa się z tylu podetapów, ile jest wyższych niż pierwszy poziomów.

- II. W etapie drugim wprowadzane są zmienne objaśniające z pierwszego poziomu. Etap ten nie dzieli się na podetapy.
- III. W etapie trzecim wprowadzane są zmienne objaśniające z poziomów wyższych. Etap ten, podobnie jak etap pierwszy, dzieli się na tyle podetapów, ile jest wyższych niż pierwszy poziomów.
- IV. W etapie czwartym uwzględniane jest zróżnicowanie zależności pomiędzy zmienną objaśnianą a zmiennymi objaśniającymi z poziomów niższych.
- V. W etapie piątym część zróżnicowania wytłumaczona może zostać przy pomocy zmiennych objaśniających z wyższych poziomów. Zarówno etap czwarty jak i piąty składają się z podetapów odpowiadających parom poziomów: niższego, z którego pochodzi zmienna, przy której stoi wyjaśniany współczynnik kierunkowy oraz wyższego, w ramach jednostek którego, współczynnik ten jest zróżnicowany.

Tak więc liczba podetapów, zarówno w etapie czwartym jak i piątym, odpowiada liczbie wariacji dwuelementowych bez powtórzeń zbioru złożonego ze wszystkich poziomów.

Konstrukcję modelu poprzedzać winna weryfikacja założeń omówionych w podrozdziale 3.1. Następnie, celem porównania w późniejszych etapach, wyznaczyć należy dwie funkcje regresji liniowej. W pierwszej z nich nie uwzględnia się żadnych zmiennych objaśniających. W drugiej funkcji regresji, jako zmienne objaśniające przyjąć należy zmienne z pierwszego poziomu.

W opisie konstrukcji modelu P-poziomowego ($P \geq 2$) przyjęto następujące oznaczenia wzorowane na oznaczeniach modelu dwupoziomowego stosowanych w pracy S.W.Raudenbush'a i A.S. Bryka (por. Raudenbush, Bryk, 2002). Podobne oznaczenia stosowane były w wielu pracach na ten temat (por. Snijders, Bosker, 1999; Klimanek, 2003). Zamiast jednak oznaczać wyrazy wolne i współczynniki kierunkowe z różnych poziomów różnymi literami greckimi, dla wszystkich poziomów zastosowano β z indeksem odpowiadającym poziomowi i zmiennej (szczegółowy opis poniżej).

Pozwoliło to dokonać uogólnienia na przypadek z dowolną liczbą poziomów (por. Hox, 2010).

W tym momencie rozważań podkreślić należy, że najczęściej stosowanym sposobem zapisu modelu wielopoziomowego, zwłaszcza w przypadku występowania więcej niż dwóch poziomów, jest zapis macierzowy (por. Hox, 2010; Bates, 2010; Biecek, 2011; Frątczak, Mianowska, 2012). Pozwala on opisać nawet bardzo skomplikowane modele przy pomocy krótkiego równania, ponadto zapis taki wykorzystywany jest podczas szacowania parametrów modelu. Mianowicie każdy model wielopoziomowy może zostać zapisany przy pomocy równia:

$$Y = X\beta + Ze + r \quad , \quad (3.1)$$

gdzie:

Y – wektor zmiennej objaśnianej,

X – macierz zmiennych objaśniających dla efektów stałych modelu,

β – wektor współczynników kierunkowych,

Z – macierz zmiennych objaśniających dla efektów losowych modelu,

e – wektor efektów losowych, $e \sim N(\mathbf{0}; \Sigma)$,

r – składnik resztowy modelu, $r \sim N(\mathbf{0}; \sigma^2 \mathbf{I}_d)$.

Jak zaznaczono w wyjaśnieniach przyjmuje się, że wektor efektów losowych oraz składnik resztowy charakteryzują się wielowymiarowymi rozkładami normalnymi. Jednocześnie dopuszcza się występowanie zależności pomiędzy niektórymi efektami. Z założenia, jednak przyjmuje się, że poszczególne składniki resztowe są od siebie niezależne. Ponadto zakłada się, że nie występuje korelacja pomiędzy efektami losowymi a składnikiem resztowym. Rozkład warunkowy zmiennej objaśnianej, przy określonym wektorze efektów losowych, można zapisać więc jako $(Y|e = \hat{e}) \sim N(X\beta + Z\hat{e}; \sigma^2 \mathbf{I}_d)$.

Jako wadę powyższego zapisu, uznać można jednak jego małą przejrzystość, pod względem zestawu zmiennych objaśniających, ich interakcji, czy efektów losowych wchodzących w skład modelu. W związku z tym, powyższy zapis nie wystarcza do opisu etapów konstrukcji modelu, który wprowadzony zostanie w dalszej części podrozdziału. Pomimo tego, na zakończenie opisu każdego z pięciu etapów konstrukcji modelu przedstawiono budowę macierzy wchodzących w skład równania (3.1). Dzięki temu

zaobserwować można, jak ich zawartość oraz wymiar zmieniają się w zależności od stopnia skomplikowania modelu.

Poniżej zestawiono przyjęte w dalszym ciągu oznaczenia.

n – liczba jednostek na pierwszym poziomie, czyli liczebność próby,

P – liczba poziomów ($P > 1$),

l_p - liczebność próby na p -tym poziomie, dla $p = 1, \dots, P$ ($l_1 = n$),

$n_{p,i}$ - liczba jednostek pierwszego poziomu, należących do i -tej jednostki na p -tym poziomie, dla $p = 1, \dots, P$, $i = 1, \dots, l_p$, $\sum_{i=1}^{l_p} n_{p,i} = n$.

Q_p – liczba zmiennych objaśniających z p -tego poziomu,

Y_i - wartość zmiennej objaśnianej dla i -tej jednostki pierwszego poziomu, $i = 1, \dots, n$,

X_{pqj} - wartość q -tej zmiennej objaśniającej (scentrowanej) z p -tego poziomu dla

i -tej jednostki tego poziomu, $p = 1, \dots, P$, $q = 1, \dots, Q_p$, $j = 1, \dots, l_p$,

$X_{p0i} = 1$ – przyjęto, że „zerowa” zmienna objaśniająca z każdego poziomu to jedynka dla $i = 1, \dots, l_p$. Ułatwi to zapis, gdyż „zerowa” zmienna pełni rolę stałej, stojącej przy wyrazie wolnym,

$\beta_{0\dots 0q0\dots 0}^{kr}$ - współczynnik kierunkowy; pierwsza wartość indeksu górnego mówi, na którym etapie ($k=1, \dots, 5$) współczynnik ten został zdefiniowany, druga określa podetap ($r=1, \dots, P$). Podetapy numerowane są ze względu na to, którego poziomu dotyczyło przeprowadzone w nich rozwinięcie.

Indeks dolny składa się z p elementów, z których prawie wszystkie są zerami. Jeżeli wszystkie elementy w indeksie dolnym są zerami, to jest to wyraz wolny. Jeżeli istnieje dokładnie jedna niezerowa wartość, to jest to współczynnik kierunkowy, przy czym pozycja tej niezerowej wartości wskazuje, przy zmiennej z którego poziomu on stoi, zaś jej wartość określa która to z kolei zmienna z tego poziomu. Jeżeli w indeksie dolnym znajduje się więcej niż jedna niezerowa wartość, jest to współczynnik kierunkowy przy interakcji między zmiennymi z różnych poziomów, przy czym pozycje niezerowych elementów wskazują z których poziomów są te zmienne, zaś ich wartości, które to są zmienne z tych poziomów. Jeżeli któraś z wartości w indeksie dolnym jest w nawiasie ($\beta_{0\dots 0q0\dots(j)0\dots 0}^{kr}$), oznacza to, że dany

współczynnik może przyjmować różne wartości dla jednostek pierwszego poziomu należących do różnych jednostek na poziomie, któremu odpowiada pozycja indeksu w nawiasie. Sytuacja taka nie jest jednak możliwa przy zapisie jednym równaniem.

$r_{ij\dots k}^{kr}$ - składnik resztowy z pierwszego poziomu, indeks górny tak samo jak w przypadku współczynników kierunkowych, $i = 1, \dots, n$, $j = 1, \dots, l_2, \dots$, $k = 1, \dots, l_p$,

$$r_{ij\dots k}^{\text{kr}} \sim N(0; \sigma_{\text{kr}}^2),$$

$e_{\underbrace{0\dots 0}_{1\dots r} \underbrace{0\dots 0}_{r+1\dots p} (k) \underbrace{0\dots 0}_{p+1\dots P}}$ - składnik losowy związany z poziomem p -tym, stojący przy q -tej zmiennej objaśniającej z r -tego poziomu, $p = 2, \dots, P$, $r = 1, \dots, p-1$, $q = 1, \dots, Q_r$,

$\tau_{p,0,0,0,0}^{\text{kr}}$ - wariancja składnika losowego z p -tego poziomu w wyrazie wolnym, $p=2, \dots, P$,

$\tau_{p,r,q,r,q}^{\text{kr}}$ - wariancja składnika losowego z p -tego poziomu we współczynniku kierunkowym q -tej zmiennej z r -tego poziomu, $p=2, \dots, P$, $r = 1, \dots, p-1$, $q=1, \dots, Q_r$,

$\tau_{p,r_1,q_1,r_2,q_2}^{\text{kr}}$ - kowariancja składnika losowego z p -tego poziomu we współczynniku kierunkowym q_1 -tej zmiennej z r_1 -tego poziomu oraz składnika losowego z p -tego poziomu we współczynniku kierunkowym q_2 -tej zmiennej z r_2 -tego poziomu, $p= 2, \dots, P$. Jeżeli $r_1, q_1 = 0,0$ lub $r_2, q_2 = 0,0$, chodzi o składnik losowy wyrazu wolnego, $r_1, r_2 = 0, \dots, p-1$, $q_1 = 0, \dots, Q_{r_1}$, $q_2 = 1, \dots, Q_{r_2}$,

e_p^{kr} - wektor zawierający wszystkie efekty losowe wprowadzone w r tym podetapie etapu k tego oraz w podetapach wcześniejszych związane z poziomem p -tym, $p=1, \dots, P$; zakłada się, że ma on wielowymiarowy rozkład normalny,

ze średnią równą zero oraz macierzą kowariancji τ_p^{kr} ; $e_p^{\text{kr}} \sim N(0; \tau_p^{\text{kr}})$,

$r_{ij\dots k}^{\text{kr}}$ - składnik resztowy; zakłada się, że ma on rozkład normalny z wartością oczekiwaną 0 oraz wariancją σ_{kr}^2 ; $r_{ij\dots k}^{\text{kr}} \sim N(0; \sigma_{\text{kr}}^2)$.

ETAP 1 Model pusty

W tym etapie zmienna Y_i objaśniana jest wyłącznie przez przynależność jednostek z pierwszego poziomu do jednostek poziomów wyższych, bez użycia zmiennych objaśniających. Model rozszerzany jest w $P-1$ podetapach, począwszy od modelu z efektem losowym wyłącznie z drugiego poziomu, do modelu z efektami losowymi z wszystkich P poziomów.

Podetap 1.1.

Model z efektem losowym z drugiego poziomu można zapisać w wersji P -równaniowej (dla każdego poziomu oddzielnie):

Na poziomie pierwszym model zapisać można przy pomocy równania:

$$Y_i = \beta_{\underbrace{0(j)0\dots 0}_P}^{12} + r_{ij\dots k}^{\text{kr}} \quad (3.2a)$$

Na poziomie drugim:

$$\beta_{\underbrace{0(j)0\dots 0}_P}^{12} = \beta_{\underbrace{00(1)0\dots 0}_P}^{12} + e_{\underbrace{0(j)0\dots 0}_P}^{12} \quad (3.2b)$$

Na poziomie trzecim:

$$\beta_{\underbrace{00(1)0\dots 0}_P}^{12} = \beta_{\underbrace{000(m)0\dots 0}_P}^{12} \quad (3.2c)$$

...

Na poziomie P -tym:

$$\beta_{\underbrace{0\dots 0(k)}_P}^{12} = \beta_{\underbrace{0\dots 0}_P}^{12} \quad (3.2d)$$

Łącznie:

$$Y_i = \beta_{\underbrace{0\dots 0}_P}^{12} + e_{\underbrace{0(j)0\dots 0}_P}^{12} + r_{\underbrace{ij\dots k}_P}^{12}, \quad (3.3)$$

gdzie:

$$r_{\underbrace{ij\dots k}_P}^{12} \sim N(0; \sigma_{kr}^2) \quad , \quad e_{\underbrace{0(j)0\dots 0}_P}^{12} \sim N(0; \tau_{2,0,0,0}^{12})$$

Porównanie wyników z liniową funkcją regresji bez zmiennych objaśniających pozwoli ocenić, czy samo uwzględnienie struktury dwupoziomowej, wprowadzone w podetapie 1.1 poprawia precyzję szacunku.

Podetap 1.(P-1)

Model z efektami losowymi ze wszystkich P poziomów można zapisać w wersji P -równaniowej (dla każdego poziomu oddzielnie):

Na poziomie pierwszym model zapisać można przy pomocy równania:

$$Y_i = \beta_{\underbrace{0(j)0\dots 0}_P}^{1P} + r_{\underbrace{ij\dots k}_P}^{1P} \quad (3.4a)$$

Na poziomie drugim:

$$\beta_{\underbrace{0(j)0\dots 0}_P}^{1P} = \beta_{\underbrace{00(1)0\dots 0}_P}^{1P} + e_{\underbrace{0(j)0\dots 0}_P}^{1P} \quad (3.4b)$$

...

Na poziomie P -tym:

$$\beta_{\underbrace{0\dots 0(k)}_P}^{1P} = \beta_{\underbrace{0\dots 0}_P}^{1P} \quad (3.4c)$$

Łącznie:

$$Y_i = \beta_{\underbrace{0\dots 0}_P}^{1P} + \sum_{p=2}^P e_{\underbrace{0\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{1P} + r_{\underbrace{ij\dots k}_P}^{1P}, \quad (3.5)$$

gdzie:

$$r_{ij\dots k}^{1P} \sim N(0; \sigma_{1P}^2) \quad , \quad e_{0\dots 0(k)0\dots 0}^{1P} \sim N(0; \tau_{p,0,0,0,0}^{1P}) \quad , \quad \text{dla } p=2,\dots,P.$$

Jeżeli uwzględnienie losowego składnika, związanego z którymś poziomem dla wyrazu wolnego, nie poprawia w istotnym stopniu jakości dopasowania modelu, nie należy takiego składnika dodawać. W takim przypadku model miałby o jeden poziom mniej aniżeli sugeruje wielopoziomowa struktura badanej populacji.

Jeżeli powyższe równanie łączne modelu z etapu pierwszego byłoby przedstawione w zapisie macierzowym (por. wzór (3.1)), macierz \mathbf{X} przyjęłaby postać n wymiarowego wektora jedynek, wektor $\boldsymbol{\beta}$ byłby jedną liczbą, $\beta_{0\dots 0}^{1P}$.

Macierz \mathbf{Z} byłaby wymiaru $n \times (\sum_{p=2}^P l_p)$, przy czym można by ją traktować, jako macierz blokową, złożoną z $P-1$ macierzy

zero-jedynkowych wymiaru $n \times l_p$, których kolumny odpowiadałyby jednostkom poziomu p -tego, a jedynki pojawiałyby się na przecięciach z wierszami odpowiadającymi tym jednostkom z poziomu pierwszego, które należą do danej jednostki poziomu p -tego, dla $p=2,\dots,P$. Wektor \mathbf{e} byłby wymiaru $\sum_{p=2}^P l_p$, i zawierałby kolejno błędy losowe odpowiadające wszystkim jednostkom każdego poziomu wyższego niż drugi.

ETAP 2 Model z losowym wyrazem wolnym

W następnym etapie do modelu wprowadzone zostają zmienne objaśniające z pierwszego poziomu. Nachylenie krzywej regresji traktujemy nadal jako stałe, bez względu na przynależność jednostek pierwszego poziomu do jednostek na wyższych poziomach. Tak więc, w składzie współczynników kierunkowych, przy dodanych zmiennych objaśniających z pierwszego poziomu, występują jedynie stałe, równe dla wszystkich jednostek pierwszego poziomu. Zmienne należy oczywiście dodawać kolejno, tak aby uniknąć wprowadzania charakterystyk, których uwzględnienie nie poprawia jakości dopasowania modelu. Ponadto również oszacowane współczynniki kierunkowe powinny być istotne statystycznie. W celu dokonania wyboru zmiennych wprowadzanych do modelu można posłużyć się na przykład procedurą regresji krokowej.

Model dla jednostek pierwszego poziomu jest następujący:

$$Y_i = \sum_{q=0}^{Q_1} \beta_{q(j)0\dots 0}^{21} X_{1,q,i} + r_{ij\dots k}^{21} \quad (3.6a)$$

Na poziomie drugim:

$$\begin{cases} \beta_{0(j)0\dots 0}^{21} = \beta_{00(1)\dots 0}^{21} + e_{0j0\dots 0}^{21} \\ \beta_{q(j)0\dots 0}^{21} = \beta_{q0(1)\dots 0}^{21}, \text{ dla } q = 1, \dots, Q_1 \end{cases} \quad (3.6b)$$

...

Na poziomie P -tym:

$$\begin{cases} \beta_{0\dots 0(k)}^{21} = \beta_{0\dots 0}^{21} + e_{0\dots k}^{21} \\ \beta_{q(j)0\dots 0}^{21} = \beta_{q0\dots 0}^{21}, \text{ dla } q = 1, \dots, Q_1 \end{cases} \quad (3.6c)$$

Łącznie:

$$Y_i = \sum_{q=0}^{Q_1} \beta_{q00\dots 0}^{21} X_{1,q,i} + \sum_{p=2}^P e_{0\dots 0(k)0\dots 0}^{21} + r_{ij\dots k}^{21} \quad , \quad (3.7)$$

gdzie:

$$r_{ij\dots k}^{21} \sim N(0; \sigma_{21}^2) \quad , \quad e_{0\dots 0(k)0\dots 0}^{21} \sim N(0; \tau_{p,0,0,0,0}^{21}) \quad , \quad \text{dla } p=2, \dots, P.$$

Tak wyznaczony model może być traktowany jako bezpośrednie rozwinięcie modelu opisanego w etapie 1, przez dodanie zmiennych objaśniających z pierwszego poziomu. Jednak może on być traktowany również jako rozwinięcie modelu klasycznej regresji liniowej ze zmiennymi objaśniającymi z pierwszego poziomu. W związku z powyższym należy wykazać jego wyższość nad oboma z nich.

Jeżeli powyższe równanie łączne modelu z etapu pierwszego zostałyby przedstawione w zapisie macierzowym (por. wzór (3.1)), macierz \mathbf{X} byłaby wymiaru $n \times (Q_1 + 1)$ przy czym jej pierwsza kolumna składałaby się z samych jedynek, kolejne zaś zawierałyby wartości zmiennych objaśniających z pierwszego poziomu. Wektor $\boldsymbol{\beta}$ zawierałby zaś wyraz wolny oraz współczynniki kierunkowe odpowiadające tym zmiennym z pierwszego poziomu. Macierz \mathbf{Z} oraz wektor \mathbf{e} nie zmieniłyby się w stosunku do poprzedniego etapu.

ETAP 3 Model z losowym wyrazem wolnym zależnym od zmiennych z wyższych poziomów

W porównaniu z modelem z etapu drugiego, model wzbogacony zostanie o zmienne objaśniające z wyższych poziomów. Oznacza to, że wyraz wolny będzie nie tylko od przynależności do jednostek na wyższych poziomach, ale objaśniany będzie także

przy pomocy zmiennych z tych poziomów. Rozszerzenie to następuje w P-1 podetapach, przy czym w pierwszym z nich, dołączane są zmienne z drugiego poziomu. W ostatnim podetapie dołączane są zmienne z P-tego poziomu. Podobnie jak w etapie drugim, zmienne

z każdego poziomu (w ramach podetapów) należy wprowadzać kolejno, tak aby uniknąć wprowadzania charakterystyk, których uwzględnienie nie poprawia jakości dopasowania modelu lub których współczynniki kierunkowe nie są statystycznie różne od zera. W większości podetapów wprowadzona zostanie część wstępnych zmiennych objaśniających, jednak wystąpić może sytuacja, że w modelu na tym etapie nie zostanie uwzględniona żadna zmienna objaśniająca z niektórych poziomów. Nie oznacza to jednak jeszcze, że odrzucone w tym etapie zmienne nie pojawią się w końcowej wersji modelu. Zdarzyć może się bowiem, że część z nich będzie zmiennymi objaśniającymi współczynniki kierunkowe przy charakterystykach z poziomów niższych (etap 5) lub wartość współczynnika kierunkowego, która nie różni się istotnie od zera w przekroju całej populacji, różnić będzie się istotnie w przekroju jednostek któregoś z wyższych poziomów (etap 4).

Podetap 3.1

Model dla jednostek pierwszego poziomu jest następujący:

$$Y_i = \sum_{q=0}^{Q_1} \beta_{\underbrace{q(j)0\dots0}_P}^{32} X_{1,q,i} + r_{\underbrace{ij\dots k}_P}^{32} \quad (3.8a)$$

Na poziomie drugim:

$$\begin{cases} \beta_{\underbrace{0(j)0\dots0}_P}^{32} = \sum_{r=0}^{Q_2} \beta_{\underbrace{0r(1)0\dots0}_P}^{32} X_{2,r,i} + e_{\underbrace{0(j)0\dots0}_P}^{32} \\ \beta_{\underbrace{q(j)0\dots0}_P}^{32} = \beta_{\underbrace{q0(1)0\dots0}_P}^{32}, \text{ dla } q = 1, \dots, Q_1 \end{cases}, \quad (3.8b)$$

Na poziomie trzecim:

$$\begin{cases} \beta_{\underbrace{00(1)0\dots0}_P}^{32} = \beta_{\underbrace{000(m)0\dots0}_P}^{32} + e_{\underbrace{00(1)0\dots0}_P}^{32} \\ \beta_{\underbrace{q0(1)0\dots0}_P}^{32} = \beta_{\underbrace{q00(m)0\dots0}_P}^{32}, \text{ dla } q = 1, \dots, Q_1, \\ \beta_{\underbrace{0r(1)0\dots0}_P}^{32} = \beta_{\underbrace{0r0(m)0\dots0}_P}^{32}, \text{ dla } r = 1, \dots, Q_2 \end{cases} \quad (3.8c)$$

...

Na poziomie P -tym:

$$\begin{cases} \beta_{\underline{0\dots 0(k)}\underline{P}}^{32} = \beta_{\underline{0\dots 0}\underline{P}}^{32} + e_{\underline{0\dots k}\underline{P}}^{32} \\ \beta_{\underline{q0\dots 0(k)}\underline{P}}^{32} = \beta_{\underline{q0\dots 0}\underline{P}}^{32}, \text{ dla } q = 1, \dots, Q_1, \\ \beta_{\underline{0r0\dots 0(k)}\underline{P}}^{32} = \beta_{\underline{0r0\dots 0}\underline{P}}^{32}, \text{ dla } r = 1, \dots, Q_2 \end{cases} \quad (3.8d)$$

Łącznie:

$$Y_i = \beta_{\underline{0\dots 0}\underline{P}}^{32} + \sum_{q=0}^{Q_2} \beta_{\underline{q0(1)0\dots 0}\underline{P}}^{32} X_{2,q,i} + \sum_{q=1}^{Q_1} \beta_{\underline{q00\dots 0}\underline{P}}^{32} X_{1,q,i} + \sum_{p=2}^P e_{\underline{0\dots 0(p)0\dots 0}\underline{P}}^{32} + r_{\underline{ij\dots k}\underline{P}}^{32} \quad (3.9)$$

gdzie:

$$r_{\underline{ij\dots k}\underline{P}}^{32} \sim N(0; \sigma_{32}^2), \quad e_{\underline{0\dots 0(k)0\dots 0}\underline{P}}^{32} \sim N(0; \tau_{\underline{p,0,0,0}}^{32}), \text{ dla } p=2, \dots, P.$$

Otrzymany model jest rozwinięciem modelu z losowym wyrazem wolnym i zmiennymi objaśniającymi z pierwszego poziomu (etap 2), przez dodanie do niego zmiennych objaśniających z drugiego poziomu.

Podetap 3.(P-1)

Model dla jednostek pierwszego poziomu jest następujący:

$$Y_i = \sum_{q=0}^{Q_1} \beta_{\underline{q(j)0\dots 0}\underline{P}}^{3P} X_{1,q,i} + r_{\underline{ij\dots k}\underline{P}}^{3P} \quad (3.10a)$$

Na poziomie drugim:

$$\begin{cases} \beta_{\underline{0(j)0\dots 0}\underline{P}}^{3P} = \sum_{r=0}^{Q_2} \beta_{\underline{0r(1)0\dots 0}\underline{P}}^{3P} X_{2,r,i} + e_{\underline{0(j)0\dots 0}\underline{P}}^{3P} \\ \beta_{\underline{q(j)0\dots 0}\underline{P}}^{3P} = \beta_{\underline{q0(1)\dots 0}\underline{P}}^{3P}, \text{ dla } q = 1, \dots, Q_1 \end{cases}, \quad (3.10b)$$

Na poziomie trzecim:

$$\begin{cases} \beta_{\underline{00(1)0\dots 0}\underline{P}}^{3P} = \sum_{s=0}^{Q_3} \beta_{\underline{00s(m)0\dots 0}\underline{P}}^{3P} X_{3,s,i} + e_{\underline{00(1)0\dots 0}\underline{P}}^{3P} \\ \beta_{\underline{q0(1)0\dots 0}\underline{P}}^{3P} = \beta_{\underline{q00(m)0\dots 0}\underline{P}}^{3P}, \text{ dla } q = 1, \dots, Q_1 \\ \beta_{\underline{0r(1)0\dots 0}\underline{P}}^{3P} = \beta_{\underline{0r0(m)0\dots 0}\underline{P}}^{3P}, \text{ dla } r = 1, \dots, Q_2 \end{cases}, \quad (3.10c)$$

...

Na poziomie P -tym:

$$\left\{ \begin{array}{l} \beta_{\underbrace{0\dots 0(k)}_P}^{3P} = \sum_{t=0}^{Q_P} \beta_{\underbrace{00\dots t}_P}^{3P} X_{P,t,i} + e_{\underbrace{0\dots k}_P}^{3P} \\ \beta_{\underbrace{0\dots 0q0\dots 0(k)}_{1\dots p\dots P}}^{3P} = \beta_{\underbrace{0\dots 0q0\dots 0}_{1\dots p\dots P}}^{3P}, \text{ dla } p = 1, \dots, P-1, q = 1, \dots, Q_p \end{array} \right. \quad (3.10d)$$

Łącznie:

$$Y_i = \sum_{p=1}^P \sum_{q=0}^{Q_P} \beta_{\underbrace{0\dots 0q0\dots 0}_{1\dots p\dots P}}^{3P} X_{p,q,i} + \sum_{p=2}^P e_{\underbrace{0\dots 0(p)0\dots 0}_P}^{3P} + r_{\underbrace{ij\dots k}_P}^{3P} \quad (3.11)$$

gdzie:

$$r_{\underbrace{ij\dots k}_P}^{3P} \sim N(0; \sigma_{3P}^2) \quad , \quad e_{\underbrace{0\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{3P} \sim N(0; \tau_{p,0,0,0,0}^{3P}) \quad , \quad \text{dla } p = 2, \dots, P.$$

Modele opisane w etapach 2 i 3 często nazywane są modelami z losowym wyrazem wolnym ponieważ zmienność wyrazu wolnego jest w nich wyjaśniana przez przynależność

do jednostek wyższych poziomów. Tak więc zakłada się, że wyraz wolny różni się w przekrojach jednostek wyższych poziomów, ale jednocześnie współczynniki kierunkowe przy zmiennych objaśniających są stałe w całej populacji.

Tak wyznaczony model może być traktowany jako bezpośrednie rozwinięcie modelu opisanego w etapie 2, przez dodanie zmiennych objaśniających z pierwszego poziomu. Jednak może być traktowany również jako rozwinięcie modelu klasycznej regresji liniowej

ze zmiennymi objaśniającymi ze wszystkich poziomów. Zatem konieczne jest wykazanie jego przewagi nad oboma z nich.

Jeżeli powyższe równanie łączne modelu z etapu pierwszego byłoby przedstawione w zapisie macierzowym (por. wzór (3.1)), macierz \mathbf{X} byłaby wymiaru $n \times (\sum_{p=1}^P Q_p + 1)$ przy czym jej pierwsza kolumna składałaby się z samych jedynek, kolejne zaś zawierałyby wartości zmiennych objaśniających z kolejnych poziomów, Wektor $\boldsymbol{\beta}$ zawierałby wyraz wolny oraz współczynniki kierunkowe odpowiadające tym zmiennym objaśniającym. Macierz \mathbf{Z} oraz wektor \mathbf{e} nie uległyby zmianie w stosunku do poprzednich etapów.

ETAP 4 Model z losowymi współczynnikami regresji

W tym etapie dopuszczona zostanie możliwość zróżnicowania jednostek pierwszego poziomu w ramach jednostek na wyższych poziomach, nie tylko ze względu na poziom szacowanej zmiennej, ale również ze względu na kształt zależności pomiędzy tą zmienną a zmiennymi objaśniającymi z niższych poziomów. Uwzględniamy więc, możliwość zróżnicowania wartości współczynników kierunkowych przy zmiennych z poziomów niższych, pomiędzy jednostkami poziomów wyższych. Przeprowadzamy to w $P \times (P-1)/2$ podetapach, w pierwszym z nich sprawdzając, czy uwzględnienie losowego charakteru współczynników przy zmiennych z pierwszego poziomu, ze względu na poziom drugi, poprawi jakość dopasowania modelu. Następnie sprawdza się zasadność uwzględniania losowego charakteru związanego z trzecim oraz kolejnymi poziomami współczynników kierunkowych przy zmiennych z pierwszego poziomu. Oczywiście poziom, ze względu na który uwzględniamy losowy charakter współczynnika kierunkowego musi być wyższy, niż poziom na którym określona jest zmienna, przy której on stoi. W ostatnim, $(P \times (P-1)/2)$ -tym podetapie, możliwe jest uwzględnienie zróżnicowania współczynników kierunkowych przy zmiennych z $P-1$ poziomu pomiędzy jednostkami P -tego poziomu.

Zmienne, które zostały pominięte w etapach drugim i trzecim, mogą zostać przeanalizowane ponownie. Całkiem możliwe jest bowiem, że potencjalne zmienne objaśniające, których współczynniki kierunkowe oszacowane dla całej populacji nie były istotnie różne od zera, różnią się jednak istotnie pomiędzy jednostkami wyższych poziomów.

Podetap 4.1

Model dla jednostek pierwszego poziomu jest następujący:

$$Y_i = \sum_{q=0}^{Q_1} \underbrace{\beta_{q(j)0\dots 0}^{4(1/2)}}_P X_{1,q,i} + \underbrace{r_{ij\dots k}^{4(1/2)}}_P \quad (3.12a)$$

Na poziomie drugim:

$$\begin{cases} \underbrace{\beta_{0(j)0\dots 0}^{4(1/2)}}_P = \sum_{r=0}^{Q_2} \underbrace{\beta_{0r(1)0\dots 0}^{4(1/2)}}_P X_{2,r,i} + \underbrace{e_{0(j)0\dots 0}^{4(1/2)}}_P \\ \underbrace{\beta_{q(j)0\dots 0}^{4(1/2)}}_P = \underbrace{\beta_{q0(1)\dots 0}^{4(1/2)}}_P + \underbrace{e_{q(j)0\dots 0}^{4(1/2)}}_P, \text{ dla } q = 1, \dots, Q_1 \end{cases} \quad (3.12b)$$

Na poziomie trzecim:

$$\begin{cases} \beta_{\underbrace{00(1)0\dots0}_P}^{4(1/2)} = \sum_{s=0}^{Q_3} \beta_{\underbrace{00s(m)0\dots0}_P}^{4(1/2)} X_{3,s,i} + e_{\underbrace{00(1)0\dots0}_P}^{4(1/2)} \\ \beta_{\underbrace{q0(1)0\dots0}_P}^{4(1/2)} = \beta_{\underbrace{q00(m)0\dots0}_P}^{4(1/2)}, \text{ dla } q = 1, \dots, Q_1 \\ \beta_{\underbrace{0r(1)0\dots0}_P}^{4(1/2)} = \beta_{\underbrace{0r0(m)0\dots0}_P}^{4(1/2)}, \text{ dla } r = 1, \dots, Q_2 \end{cases}, \quad (3.12c)$$

...

Na poziomie P -tym:

$$\begin{cases} \beta_{\underbrace{0\dots0(k)}_P}^{4(1/2)} = \sum_{t=0}^{Q_P} \beta_{\underbrace{00\dots t}_P}^{4(1/2)} X_{P,t,i} + e_{\underbrace{0\dots k}_P}^{4(1/2)} \\ \beta_{\underbrace{0\dots0q0\dots0(k)}_{1\dots p\dots P}}^{4(1/2)} = \beta_{\underbrace{0\dots0q0\dots0}_{1\dots p\dots P}}^{4(1/2)}, \text{ dla } p = 1, \dots, P-1, q = 1, \dots, Q_p \end{cases}, \quad (3.12d)$$

Łącznie:

$$Y_i = \sum_{p=1}^P \sum_{q=0}^{Q_p} \beta_{\underbrace{0\dots0q0\dots0(k)}_{1\dots p\dots P}}^{4(1/2)} X_{p,q,i} + \sum_{q=1}^{Q_1} e_{\underbrace{q(j)0\dots0}_P}^{4(1/2)} X_{1,q,i} + \sum_{p=2}^P e_{\underbrace{0\dots0(p)0\dots0}_P}^{4(1/2)} + r_{\underbrace{ij\dots k}_P}^{4(1/2)} \quad (3.13)$$

gdzie:

$$r_{\underbrace{ij\dots k}_P}^{4(1/2)} \sim N(0; \sigma_{4(1/2)}^2), \quad e_{\underbrace{0\dots0(k)0\dots0}_{1\dots p\dots P}}^{4(1/2)} \sim N(0; \tau_{p,0,0,0}^{4(1/2)}), \quad \text{dla } p=3, \dots, P,$$

$$\mathbf{e}_2^{4(1/2)} = \begin{bmatrix} e_{\underbrace{0(k)0\dots0}_P}^{4(1/2)} \\ e_{\underbrace{1(k)0\dots0}_{1\dots p\dots P}}^{4(1/2)} \\ \vdots \\ e_{\underbrace{Q_1(k)0\dots0}_{1\dots p\dots P}}^{4(1/2)} \end{bmatrix}, \quad \mathbf{e}_2^{4(1/2)} \sim N(\mathbf{0}; \boldsymbol{\tau}_2^{4(1/2)}).$$

W tym miejscu należy wyjaśnić, że $e_{\underbrace{0\dots0q0\dots0(k)0\dots0}_{1\dots r\dots p\dots P}}^{4(p_1/p_2)}$ dla $r=0$ rozumiane jest jako

$e_{\underbrace{0\dots0(k)0\dots0}_{1\dots p\dots P}}^{4(p_1/p_2)}$. Jest to pewnego rodzaju nieścisłość notacji, ponieważ indeks r powinien

przyjmować wartości całkowite pomiędzy 1 a P . Przyjęcie takiego uproszczenia zapisu powodowane było skróceniem zapisu równania, a tym samym równań opisujących kolejne etapy.

Po zdecydowaniu, które ze współczynników kierunkowych, przy zmiennych z pierwszego poziomu różnią się istotnie pomiędzy jednostkami drugiego poziomu,

należy dodać w odpowiednich współczynnikach kierunkowych efekty losowe związane z przynależnością do jednostek drugiego poziomu. Następnie weryfikowane jest, czy tak skonstruowany model charakteryzuje się lepszym dopasowaniem niż model oszacowany w etapie 3. Ponieważ zmiany wprowadzone zostały tylko w losowej części modelu, porównanie to może zostać przeprowadzone przy pomocy testu ilorazu wiarygodności, nawet jeżeli parametry szacowane są przy pomocy resztowej metody największej wiarygodności (por. Hox, 2010).

Podetap 4.Px(P-1)/2

Model dla jednostek pierwszego poziomu jest następujący:

$$Y_i = \sum_{q=0}^{Q_1} \underbrace{\beta_{q(j)0\dots 0}^{4(P-1/P)}}_P X_{1,q,i} + \underbrace{r_{ij\dots k}^{4(P-1/P)}}_P \quad (3.14a)$$

Na poziomie drugim:

$$\begin{cases} \underbrace{\beta_{0(j)0\dots 0}^{4(P-1/P)}}_P = \sum_{r=0}^{Q_2} \underbrace{\beta_{0r(l)0\dots 0}^{4(P-1/P)}}_P X_{2,r,i} + \underbrace{e_{0(j)0\dots 0}^{4(P-1/P)}}_P \\ \underbrace{\beta_{q(j)0\dots 0}^{4(P-1/P)}}_P = \underbrace{\beta_{q0(l)0\dots 0}^{4(P-1/P)}}_P + \underbrace{e_{q(j)0\dots 0}^{4(P-1/P)}}_P, \text{ dla } q = 1, \dots, Q_1 \end{cases}, \quad (3.14b)$$

Na poziomie trzecim:

$$\begin{cases} \underbrace{\beta_{00(l)0\dots 0}^{4(P-1/P)}}_P = \sum_{s=0}^{Q_3} \underbrace{\beta_{00s(m)0\dots 0}^{4(P-1/P)}}_P X_{3,s,i} + \underbrace{e_{00(l)0\dots 0}^{4(P-1/P)}}_P \\ \underbrace{\beta_{q0(l)0\dots 0}^{4(P-1/P)}}_P = \underbrace{\beta_{q00(m)0\dots 0}^{4(P-1/P)}}_P + \underbrace{e_{q0(l)0\dots 0}^{4(P-1/P)}}_P, \text{ dla } q = 1, \dots, Q_1 \\ \underbrace{\beta_{0r(l)0\dots 0}^{4(P-1/P)}}_P = \underbrace{\beta_{0r0(m)0\dots 0}^{4(P-1/P)}}_P + \underbrace{e_{0r0(m)0\dots 0}^{4(P-1/P)}}_P, \text{ dla } r = 1, \dots, Q_2 \end{cases}, \quad (3.14c)$$

...

Na poziomie (P-1):

$$\begin{cases} \underbrace{\beta_{0\dots 0(w)0}^{4(P-1/P)}}_P = \sum_{t=0}^{Q_P} \underbrace{\beta_{00\dots w(k)}^{4(P-1/P)}}_P X_{P,t,i} + \underbrace{e_{0\dots k}^{4(P-1/P)}}_P \\ \underbrace{\beta_{0\dots 0q0\dots 0(w)0}^{4(P-1/P)}}_{1\dots p\dots P} = \underbrace{\beta_{0\dots 0q0\dots 0(k)}^{4(P-1/P)}}_{1\dots p\dots P} + \underbrace{e_{0\dots 0q0\dots 0(k)}^{4(P-1/P)}}_{1\dots p\dots P}, \text{ dla } p = 1, \dots, P-1, q = 1, \dots, Q_p \end{cases}, \quad (3.14d)$$

Na poziomie P -tym:

$$\begin{cases} \beta_{\underbrace{0\dots 0(k)}_P}^{4(P-1/P)} = \sum_{t=0}^{Q_P} \beta_{\underbrace{00\dots t}_P}^{4(P-1/P)} X_{P,t,i} + e_{\underbrace{0\dots k}_P}^{4(P-1/P)} \\ \beta_{\underbrace{0\dots 0q0\dots 0(k)}_{1\dots p\dots P}}^{4(P-1/P)} = \beta_{\underbrace{0\dots 0q0\dots 0}_{1\dots p\dots P}}^{4(P-1/P)}, \text{ dla } p = 1, \dots, P-1, q = 1, \dots, Q_p \end{cases}, \quad (3.14e)$$

Łącznie:

$$Y_i = \sum_{p=1}^P \sum_{q=0}^{Q_P} \beta_{\underbrace{0\dots 0q0\dots 0}_{1\dots p\dots P}}^{4(P-1/P)} X_{p,q,i} + \sum_{p=2}^P \sum_{r=0}^{p-1} \sum_{q=1}^{Q_p} e_{\underbrace{0\dots 0q0\dots 0(k)0\dots 0}_{1\dots r\dots p\dots P}}^{4(P-1/P)} X_{p,q,i} + \Gamma_{\underbrace{ij\dots k}_P}^{4(P-1/P)}, \quad (3.15)$$

gdzie: $\Gamma_{\underbrace{ij\dots k}_P}^{4(P-1/P)} \sim N(0; \sigma_{4(P-1/P)}^2)$, $e_p^{4(P-1/P)} = \begin{bmatrix} e_{\underbrace{00\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{4(P-1/P)} \\ e_{\underbrace{10\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{4(P-1/P)} \\ \vdots \\ e_{\underbrace{Q_1 0\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{4(P-1/P)} \\ e_{\underbrace{01\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{4(P-1/P)} \\ \vdots \\ e_{\underbrace{0\dots 01(k)0\dots 0}_{1\dots p\dots P}}^{4(P-1/P)} \\ \vdots \\ e_{\underbrace{0\dots 0Q_{p-1}(k)0\dots 0}_{1\dots p\dots P}}^{4(P-1/P)} \end{bmatrix}$,

$$e_p^{4(P-1/P)} \sim N(0; \tau_p^{4(P-1/P)}), \text{ dla } p = 2, \dots, P.$$

Należy pamiętać, że w przypadku dodania każdego kolejnego efektu losowego związanego z danym poziomem, konieczne jest oszacowanie nie tylko jego wariancji, ale również kowariancji ze wszystkimi wcześniej dodanymi efektami losowymi z tego poziomu. Łatwo zauważyć, że w takiej sytuacji liczba szacowanych parametrów szybko rośnie,

co w znacznym stopniu komplikuje model. W związku z tym należy być szczególnie ostrożnym w przypadku dodawania kolejnych efektów losowych we współczynnikach kierunkowych (por. Hox, 2010). Należy rozważać wprowadzenie efektów losowych tylko w tych współczynnikach kierunkowych, w przypadku których istnieją przesłanki merytoryczne, aby podejrzewać, że różnią się one pomiędzy jednostkami wyższego poziomu.

Jeżeli powyższe równanie łączne modelu z etapu pierwszego zostałyby przedstawione w zapisie macierzowym (por. wzór (3.1)), macierz \mathbf{X} wektor $\boldsymbol{\beta}$ pozostały by takie same

jak w etapie poprzednim. Macierz Z byłaby wymiaru $n \times (\sum_{p=2}^P (l_p * (1 + \sum_{r=1}^{p-1} Q_r)))$, przy czym można by ją traktować, jako macierz blokową, złożoną z $P-1$ macierzy zero-jedynkowych wymiaru $n \times l_p$, w których kolumny odpowiadałyby jednostkom poziomu p -tego i jedynki pojawiałyby się na przecięciach z wierszami odpowiadającymi tym jednostkom z poziomu pierwszego, które należą do danej jednostki poziomu p -tego oraz z $\sum_{r=1}^P Q_r$ macierzy o takim samym wymiarze, jednak złożonych z wartości zmiennych objaśniających z niższych niż p -ty poziomów, dla $p = 2, \dots, P$. Wektor e byłby wymiaru $\sum_{p=2}^P l_p * (1 + \sum_{r=1}^P Q_r)$ i zawierałby kolejno błędy losowe wchodzące w skład wyrazu wolnego oraz współczynników kierunkowych przy zmiennych objaśniających z niższych poziomów, odpowiadające wszystkim jednostkom każdego poziomu wyższego niż drugi.

ETAP 5 Model z losowymi współczynnikami regresji zależnym od zmiennych z wyższych poziomów

W modelu regresji liniowej dla losowego wyrazu wolnego oraz losowych współczynników kierunkowych, objaśnianych przy użyciu zmiennych z wyższych poziomów uwzględnia się możliwość wpływu przynależności do jednostki wyższego poziomu tak na poziom jak i kształt zależności pomiędzy szacowaną zmienną a zmiennymi objaśniającymi z niższych poziomów. Dodatkowo objaśnia się zmienność szacowanej cechy na niższym poziomie przy pomocy zmiennych określonych na wyższym poziomie. Tak więc przyjmuje się, że współczynniki kierunkowe dla zmiennych objaśniających z niższych poziomów nie tylko różnią się zależnie od przynależności do jednostek z wyższych poziomów, ale mogą być objaśniane przez zmienne z tych poziomów.

Innymi słowy, część zróżnicowania współczynników kierunkowych przy zmiennych objaśniających związana z podziałem na jednostki poziomu wyższego niż ten, dla którego określona jest zmienna przy której stoi odpowiedni współczynnik, wyjaśniona może zostać przez zmienne objaśniające z tego samego, wyższego poziomu z którym związane jest to zróżnicowanie. W ten sposób w równaniu opisującym model pojawiają się iloczyny zmiennych objaśniających z różnych poziomów. Rozszerzanie modelu w tym etapie, podobnie

jak w etapie 4, następuje w $P \times (P-1)/2$ podetapach. W pierwszym z nich współczynniki kierunkowe przy zmiennych objaśniających z poziomu pierwszego wyjaśniane zostają

przez zmienne z poziomu drugiego. W kolejnych podetapach współczynniki przy zmiennych

z poziomu pierwszego wyjaśniane zostają również przez zmienne objaśniające z poziomu trzeciego itd. W ostatnim podetapie współczynniki kierunkowe przy zmiennych z przedostatniego poziomu wyjaśniane zostają przez zmienne z poziomu ostatniego.

Podetap 5.1.

Model dla jednostek pierwszego poziomu jest następujący:

$$Y_i = \sum_{q=0}^{Q_1} \underbrace{\beta_{q(j)0\dots0}^{5(1/2)}}_P X_{1,q,i} + \underbrace{r_{ij\dots k}^{5(1/2)}}_P \quad (3.16a)$$

Na poziomie drugim:

$$\begin{cases} \underbrace{\beta_{0(j)0\dots0}^{5(1/2)}}_P = \sum_{r=0}^{Q_2} \underbrace{\beta_{0r(1)0\dots0}^{5(1/2)}}_P X_{2,r,i} + \underbrace{e_{0(j)0\dots0}^{5(1/2)}}_P \\ \underbrace{\beta_{q(j)0\dots0}^{5(1/2)}}_P = \sum_{s=0}^{Q_2} \underbrace{\beta_{qs(1)\dots0}^{5(1/2)}}_P X_{2,s,i} + \underbrace{e_{q(j)0\dots0}^{5(1/2)}}_P, \text{ dla } q = 1, \dots, Q_1 \end{cases} \quad (3.16b)$$

Na poziomie trzecim:

$$\begin{cases} \underbrace{\beta_{00(1)0\dots0}^{5(1/2)}}_P = \sum_{s=0}^{Q_3} \underbrace{\beta_{00s(m)0\dots0}^{5(1/2)}}_P X_{3,s,i} + \underbrace{e_{00(1)0\dots0}^{5(1/2)}}_P \\ \underbrace{\beta_{q0(1)0\dots0}^{5(1/2)}}_P = \underbrace{\beta_{q00(m)0\dots0}^{5(1/2)}}_P + \underbrace{e_{q0(1)0\dots0}^{5(1/2)}}_P, \text{ dla } q = 1, \dots, Q_1 \\ \underbrace{\beta_{0r(1)0\dots0}^{5(1/2)}}_P = \underbrace{\beta_{0r0(m)0\dots0}^{5(1/2)}}_P + \underbrace{e_{0r0(m)0\dots0}^{5(1/2)}}_P, \text{ dla } r = 1, \dots, Q_2 \\ \underbrace{\beta_{qs(1)\dots0}^{5(1/2)}}_P = \underbrace{\beta_{qs0(m)\dots0}^{5(1/2)}}_P, \text{ dla } q = 1, \dots, Q_1, s = 1, \dots, Q_2 \end{cases} \quad (3.16c)$$

...

Na poziomie (P-1):

$$\begin{cases} \underbrace{\beta_{0\dots0(w)0}^{5(1/2)}}_P = \sum_{t=0}^{Q_P} \underbrace{\beta_{00\dots w(k)}^{5(1/2)}}_P X_{P,t,i} + \underbrace{e_{0\dots k}^{5(1/2)}}_P \\ \underbrace{\beta_{0\dots0q0\dots0(w)0}^{5(1/2)}}_{\substack{1\dots p\dots\dots P \\ P}} = \underbrace{\beta_{0\dots0q0\dots0(k)}^{5(1/2)}}_{\substack{1\dots p\dots\dots P \\ P}} + \underbrace{e_{0\dots0q0\dots0(k)}^{5(1/2)}}_{\substack{1\dots p\dots\dots P \\ P}}, \text{ dla } p = 1, \dots, P-2, q = 1, \dots, Q_p \\ \underbrace{\beta_{qs0\dots0(w)0}^{5(1/2)}}_P = \underbrace{\beta_{qs0\dots0(k)}^{5(1/2)}}_P, \text{ dla } q = 1, \dots, Q_1, s = 1, \dots, Q_2 \end{cases} \quad (3.16d)$$

Na poziomie P -tym:

$$\begin{cases} \beta_{\underbrace{0\dots 0(k)}_P}^{5(1/2)} = \sum_{t=0}^{Q_P} \beta_{\underbrace{00\dots t}_P}^{5(1/2)} X_{P,t,i} + e_{\underbrace{0\dots k}_P}^{5(1/2)} \\ \beta_{\underbrace{0\dots 0q0\dots 0(k)}_{1\dots p\dots P}}^{5(1/2)} = \beta_{\underbrace{0\dots 0q0\dots 0}_{1\dots p\dots P}}^{5(1/2)}, \text{ dla } p = 1, \dots, P-1, q = 1, \dots, Q_p, \\ \beta_{\underbrace{qs0\dots 0(k)}_P}^{5(1/2)} = \beta_{\underbrace{qs0\dots 00}_{P}}^{5(1/2)}, \text{ dla } q = 1, \dots, Q_1, s = 1, \dots, Q_2 \end{cases} \quad (3.16e)$$

Łącznie:

$$Y_i =$$

$$\begin{aligned} & \sum_{p=1}^P \sum_{q=0}^{Q_p} \beta_{\underbrace{0\dots 0q0\dots 0}_{1\dots p\dots P}}^{5(1/2)} X_{p,q,i} + \sum_{q=1}^{Q_1} \sum_{s=0}^{Q_2} \beta_{\underbrace{qs0\dots 0}_{P}}^{5(1/2)} X_{1,q,i} X_{2,s,i} + \\ & \sum_{p=2}^P \sum_{r=0}^{p-1} \sum_{q=1}^{Q_p} e_{\underbrace{0\dots 0q0\dots 0(k)0\dots 0}_{1\dots r\dots p\dots P}}^{5(1/2)} X_{p,q,i} + r_{\underbrace{ij\dots k}_P}^{5(1/2)} \end{aligned} \quad (3.17)$$

$$\text{gdzie: } r_{\underbrace{ij\dots k}_P}^{5(1/2)} \sim N(0; \sigma_{5(1/2)}^2), \quad e_p^{5(1/2)} = \begin{bmatrix} e_{\underbrace{00\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{5(1/2)} \\ e_{\underbrace{10\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{5(1/2)} \\ \vdots \\ e_{\underbrace{Q_1 0\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{5(1/2)} \\ e_{\underbrace{01\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{5(1/2)} \\ \vdots \\ e_{\underbrace{0\dots 01(k)0\dots 0}_{1\dots p\dots P}}^{5(1/2)} \\ \vdots \\ e_{\underbrace{0\dots 0Q_{p-1}(k)0\dots 0}_{1\dots p\dots P}}^{5(1/2)} \end{bmatrix}, \quad e_p^{5(1/2)} \sim N(\mathbf{0}; \boldsymbol{\tau}_p^{5(1/2)}), \quad \text{dla}$$

$p=2, \dots, P$.

W kolejnych podetapach należy pamiętać, że nie każda zmienna z wyższego poziomu nadaje się do objaśniania współczynnika kierunkowego przy każdej ze zmiennych z poziomu niższego. Sytuacja taka uwarunkowana może być tak względami merytorycznymi,

jak również brakiem poprawy jakości modelu. W takim wypadku nie należy oczywiście wprowadzać do modelu interakcji pomiędzy takimi zmiennymi. J. Hox w swojej pracy podkreśla, że wprowadzanie zbyt wielu interakcji zwłaszcza między więcej niż dwiema zmiennymi objaśniającymi z różnych poziomów prowadzi do komplikacji przy interpretacji wyników oraz kłopotów ze zbieżnością algorytmów (por. Hox, 2010).

Podetap 5.Px(P-1)/2.

Model dla jednostek pierwszego poziomu jest następujący:

$$Y_i = \sum_{q=0}^{Q_1} \beta_{\underline{q(j)0\dots0}}^{\frac{5(P-1/P)}{P}} X_{1,q,i} + r_{\underline{ij\dots k}}^{\frac{5(P-1/P)}{P}} \quad (3.18a)$$

Na poziomie drugim:

$$\left\{ \begin{array}{l} \beta_{\underline{0(j)0\dots0}}^{\frac{5(P-1/P)}{P}} = \sum_{r=0}^{Q_2} \beta_{\underline{0r(l)0\dots0}}^{\frac{5(P-1/P)}{P}} X_{2,r,i} + e_{\underline{0(j)0\dots0}}^{\frac{5(P-1/P)}{P}} \\ \beta_{\underline{q(j)0\dots0}}^{\frac{5(P-1/P)}{P}} = \sum_{s=0}^{Q_2} \beta_{\underline{qs(l)0\dots0}}^{\frac{5(P-1/P)}{P}} X_{2,s,i} + e_{\underline{q(j)0\dots0}}^{\frac{5(P-1/P)}{P}}, \text{ dla } q = 1, \dots, Q_1 \end{array} \right. , \quad (3.18b)$$

Na poziomie trzecim:

$$\left\{ \begin{array}{l} \beta_{\underline{00(l)0\dots0}}^{\frac{5(P-1/P)}{P}} = \sum_{s=0}^{Q_3} \beta_{\underline{00s(m)0\dots0}}^{\frac{5(P-1/P)}{P}} X_{3,s,i} + e_{\underline{00(l)0\dots0}}^{\frac{5(P-1/P)}{P}} \\ \beta_{\underline{q0(l)0\dots0}}^{\frac{5(P-1/P)}{P}} = \sum_{t=0}^{Q_3} \beta_{\underline{q0t(m)0\dots0}}^{\frac{5(P-1/P)}{P}} X_{3,t,i} + e_{\underline{q0(l)0\dots0}}^{\frac{5(P-1/P)}{P}}, \text{ dla } q = 1, \dots, Q_1 \\ \beta_{\underline{0r(l)0\dots0}}^{\frac{5(P-1/P)}{P}} = \sum_{t=0}^{Q_3} \beta_{\underline{0rt(m)0\dots0}}^{\frac{5(P-1/P)}{P}} X_{3,t,i} + e_{\underline{0r0(m)0\dots0}}^{\frac{5(P-1/P)}{P}}, \text{ dla } r = 1, \dots, Q_2 \\ \beta_{\underline{qs(l)0\dots0}}^{\frac{5(P-1/P)}{P}} = \beta_{\underline{qs0(m)0\dots0}}^{\frac{5(P-1/P)}{P}}, \text{ dla } q = 1, \dots, Q_1, s = 1, \dots, Q_2 \end{array} \right. , \quad (3.18c)$$

...

Na poziomie (P-1):

$$\left\{ \begin{array}{l} \beta_{\underline{0\dots0(w)0}}^{\frac{5(P-1/P)}{P}} = \sum_{t=0}^{Q_P} \beta_{\underline{00\dots w(k)}}^{\frac{5(P-1/P)}{P}} X_{P,t,i} + e_{\underline{0\dots k}}^{\frac{5(P-1/P)}{P}} \\ \beta_{\underline{0\dots0q_0\dots0(w)0}}^{\frac{5(P-1/P)}{1\dots p\dots p\dots P}} = \sum_{q_{P-1}=0}^{Q_{P-1}} \beta_{\underline{0\dots0q_0\dots0q_{P-1}(k)}}^{\frac{5(P-1/P)}{1\dots p\dots p\dots P}} X_{P-1,q_{P-1},i} + e_{\underline{0\dots0q_0\dots0(k)}}^{\frac{5(P-1/P)}{1\dots p\dots p\dots P}}, \\ \text{dla } p = 1, \dots, P-2, q = 1, \dots, Q_p \\ \beta_{\underline{q_1\dots q_{P-2}(w)0}}^{\frac{5(P-1/P)}{P}} = \sum_{q_{P-1}=0}^{Q_{P-1}} \beta_{\underline{q_1\dots q_{P-2}q_{P-1}(k)}}^{\frac{5(P-1/P)}{P}} X_{P-1,q_{P-1},i}, \\ \text{dla } p = 1, \dots, P-2, q_p = 0, 1, \dots, Q_p, \text{ jeżeli po lewej stronie} \\ \text{równania dla więcej niż jednego } p = 1, \dots, P-2, q_p \neq 0 \end{array} \right. \quad (3.18d)$$

Na poziomie P -tym:

$$\begin{cases} \beta_{\underbrace{0\dots 0(k)}_P}^{5(P-1/P)} = \sum_{t=0}^{Q_P} \beta_{\underbrace{00\dots t}_P}^{5(P-1/P)} X_{P,t,i} + e_{\underbrace{0\dots k}_P}^{5(P-1/P)} \\ \beta_{\underbrace{q_1\dots q_{P-2}q_{P-1}(k)}_P}^{5(P-1/P)} = \sum_{q_P=0}^{Q_P} \beta_{\underbrace{q_1\dots q_{P-2}q_{P-1}q_P}_P}^{5(P-1/P)} X_{P,q_P,i} \end{cases}, \text{ dla } p = 1, \dots, P-1, q_p = 0, 1, \dots, Q_p,$$

(3.18e)

Łącznie:

$$Y_i = \sum_{A \subset Z} \beta_{\underbrace{a_1 a_2 \dots a_P}_P}^{5(P-1/P)} \prod_{j=1}^P X_{j,a_j,i} + \sum_{p=2}^P \sum_{r=0}^{p-1} \sum_{q=1}^{Q_p} e_{\underbrace{0\dots 0q0\dots 0(k)0\dots 0}_{1\dots r\dots p\dots P}}^{5(P-1/P)} X_{p,q,i} + r_{\underbrace{ij\dots k}_P}^{5(P-1/P)},$$

(3.19)

gdzie:

$$r_{\underbrace{ij\dots k}_P}^{5(P-1/P)} \sim N(0; \sigma_{5(P-1/P)}^2), \quad e_p^{5(P-1/P)} = \begin{bmatrix} e_{\underbrace{00\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{5(P-1/P)} \\ e_{\underbrace{10\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{5(P-1/P)} \\ \vdots \\ e_{\underbrace{Q_1 0\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{5(P-1/P)} \\ e_{\underbrace{01\dots 0(k)0\dots 0}_{1\dots p\dots P}}^{5(P-1/P)} \\ \vdots \\ e_{\underbrace{0\dots 01(k)0\dots 0}_{1\dots p\dots P}}^{5(P-1/P)} \\ \vdots \\ e_{\underbrace{0\dots 0Q_{p-1}(k)0\dots 0}_{1\dots p\dots P}}^{5(P-1/P)} \end{bmatrix},$$

$$e_p^{5(P-1/P)} \sim N(0; \tau_p^{5(P-1/P)}), \text{ dla } p=2, \dots, P.$$

We wzorze (3.19) zastosowano sumowanie po zbiorze Z , którego elementy są wszystkimi możliwymi zestawami zmiennych po jednej z każdego poziomu. Pod uwagę dla każdego poziomu brane są również „zmiennie zerowe”, czyli zmiennie przyjmujące dla całej populacji stałą wartość 1. Zatem $\bar{Z} = \prod_{p=1}^P (Q_p + 1)$ oraz zbiór Z składa się z $\prod_{p=1}^P (Q_p + 1)$ zbiorów, które z kolei zawierają po P elementów każdy, odpowiadających numerom zmiennych objaśniających z kolejnych poziomów. Można to także zapisać następująco: jeżeli $A \subset Z$, to $\bar{A} = P$, $A = \{a_1, a_2, \dots, a_p\}$, przy czym dla $p=1, \dots, P$, $a_p \in \{0, 1, \dots, Q_p\}$.

Jeżeli powyższe równanie łączne modelu z etapu pierwszego, zostałyby przedstawione w zapisie macierzowym (por. wzór (3.1)), macierz \mathbf{X} byłaby wymiaru $n \times \prod_{p=1}^P (Q_p + 1)$, przy czym jej kolumny odpowiadałyby wszystkim możliwym iloczynom zmiennych

należących do zestawów ze zbioru Z . Wektor β byłby $\prod_{p=1}^P(Q_p + 1)$ wymiarowy i zawierałby współczynniki przy wszystkich możliwych iloczynach zmiennych z różnych poziomów (wliczając w to wyraz wolny oraz zmienne pojedyncze). Z kolei część losowa układu nie uległaby zmianie w stosunku do etapu poprzedniego.

Dalsze komplikowanie modelu jest teoretycznie możliwe (w przypadku więcej niż dwupoziomowym). Można mianowicie wprowadzić elementy losowe we współczynnikach przy interakcjach pomiędzy zmiennymi. Elementy losowe musiałyby być oczywiście związane z poziomem wyższym, niż każda ze zmiennych wchodzących w skład iloczynu. Takie rozszerzenie modelu prowadziłoby jednak do jego ogromnej komplikacji a w konsekwencji do kompletnego braku czytelności znaczenia szacowanych parametrów. Ponadto, ze względu na nadmierną liczbę parametrów można spodziewać się braku zbieżności w praktycznej próbie szacowania parametrów. Z drugiej strony wydaje się, że uwzględnienie losowego charakteru współczynników przy interakcjach między zmiennymi nie powinno w istotnym stopniu poprawić własności predykcyjnych modelu, ponieważ uwzględnione już zostały efekty losowe przy każdej ze zmiennych wchodzących w skład iloczynu z osobna. Zatem w rozdziale 4.1 prezentowanej pracy skonstruowany zostanie model opisany przy pomocy pięciu powyższych etapów.

3.3. Model wielopoziomowy z dwoma kryteriami grupowania

W podrozdziale 3.3 zaproponowane zostaną etapy konstrukcji modelu opartego na danych o wielopoziomowej strukturze z dwoma kryteriami grupowania (ang. cross model), (por. schemat 3.2). Rozszerzanie modelu przebiega w etapach analogicznych do klasycznego modelu wielopoziomowego. Ponieważ etapy konstrukcji klasycznego modelu wielopoziomowego zostały opisane szczegółowo w poprzednim podrozdziale, w przypadku modelu z dwoma kryteriami grupowania (nazywanego także modelem krzyżowym) wskazane zostaną tylko różnice wynikające z uwzględniania bardziej skomplikowanej struktury jednostek badania.

W związku z uwzględnieniem krzyżowej struktury populacji konieczne było przyjęcie odzwierciedlającej ją notacji. Poniżej zestawiono przyjęte w dalszym ciągu oznaczenia.

n – liczba jednostek na pierwszym poziomie, czyli liczebność próby,

P_a – ilość poziomów ze względu na pierwsze kryterium grupowania ($P_a > 1$),
 P_b – ilość poziomów ze względu na drugie kryterium grupowania ($P_b > 1$),
 $l_{a,p}$ - liczebność próby na p -tym poziomie ze względu na pierwsze kryterium grupowania, dla $p = 1, \dots, P_a$ ($l_{a,1} = n$),
 $l_{b,p}$ - liczebność próby na p -tym poziomie ze względu na drugie kryterium grupowania, dla $p = 1, \dots, P_b$ ($l_{b,1} = n$),
 $n_{a,p,i}$ - liczebność jednostek pierwszego poziomu, należących do i -tej jednostki na p -tym poziomie ze względu na pierwsze kryterium grupowania, dla $p = 1, \dots, P_a$, $i = 1, \dots, l_{a,p}$, $\sum_{i=1}^{l_{a,p}} n_{a,p,i} = n$,
 $n_{b,p,i}$ - liczebność jednostek pierwszego poziomu, należących do i -tej jednostki na p -tym poziomie ze względu na drugie kryterium grupowania, dla $p = 1, \dots, P_b$, $i = 1, \dots, l_{b,p}$, $\sum_{i=1}^{l_{b,p}} n_{b,p,i} = n$,
 Q_1 – liczba zmiennych objaśniających z pierwszego poziomu,
 $Q_{a,p}$ – liczba zmiennych objaśniających z p -tego poziomu ze względu na pierwsze kryterium grupowania,
 $Q_{b,p}$ – liczba zmiennych objaśniających z p -tego poziomu ze względu na drugie kryterium grupowania,
 Y_i - wartość zmiennej objaśnianej dla i -tej jednostki pierwszego poziomu, $i = 1, \dots, n$,
 X_{1qi} - wartość q -tej zmiennej objaśniającej (scentrowanej) z pierwszego poziomu dla i -tej jednostki pierwszego poziomu, $q = 1, \dots, Q_1$, $i = 1, \dots, n$,
 X_{apqi} - wartość q -tej zmiennej objaśniającej (scentrowanej) z p -tego poziomu ze względu na pierwsze kryterium grupowania dla i -tej jednostki tego poziomu, $p = 1, \dots, P_a$, $q = 1, \dots, Q_{a,p}$, $j = 1, \dots, l_{a,p}$,
 X_{bpqi} - wartość q -tej zmiennej objaśniającej (scentrowanej) z p -tego poziomu ze względu na drugie kryterium grupowania dla i -tej jednostki tego poziomu, $p = 1, \dots, P_b$, $q = 1, \dots, Q_{b,p}$, $j = 1, \dots, l_{b,p}$,
 $X_{ap0i} = X_{ap0i} = 1$ – przyjęto, że „zerowa” zmienna objaśniająca z każdego poziomu to jedynka. Oznaczenie takie uprości zapis, poprzez dołączenie wyrazu wolnego do współczynników kierunkowych,
 β_q^{kcr} - współczynnik kierunkowy przy q -tej zmiennej objaśniającej z pierwszego poziomu; pierwsza wartość indeksu górnego mówi, na którym etapie ($k=1, \dots, 5$) współczynnik ten został zdefiniowany, druga i trzecia określają podetap; $c \in \{a, b\}$, $r=1, \dots, P_c$. Podetapy numerowane są ze względu na poziom, którego dotyczą i kryterium grupowania, którego dotyczyło przeprowadzone w nich rozwinięcie. Jeżeli $q=0$, to jest to wyraz wolny,

$\beta_{a0\dots0q0\dots0}^{kcr}$ - współczynnik kierunkowy; pierwsza wartość indeksu górnego mówi, na którym etapie ($k=1,\dots,5$) współczynnik ten został zdefiniowany, druga i trzecia określają podetap; $c \in \{a, b\}$, $r=1,\dots, P_c$. Podetapy numerowane są ze względu na poziom, którego dotyczą i kryterium grupowania, którego dotyczyło przeprowadzone w nich rozwinięcie. Indeks dolny składa się z literki określającej, kryterium grupowania oraz z p elementów, z których prawie wszystkie są zerami. Jeżeli wszystkie elementy w indeksie dolnym są zerami, to jest to wyraz wolny. Jeżeli istnieje dokładnie jedna niezerowa wartość, to jest to współczynnik kierunkowy, przy czym pozycja tej niezerowej wartości wskazuje, przy zmiennej z którego poziomu on stoi, zaś jej wartość określa która to z kolei zmienna z tego poziomu. Jeżeli w indeksie dolnym znajduje się więcej niż jedna niezerowa wartość, jest to współczynnik kierunkowy interakcji między zmiennymi z różnych poziomów, przy czym pozycje niezerowych elementów wskazują z których poziomów są to zmienne, zaś ich wartości, które to zmienne z tych poziomów. Jeżeli któraś z wartości w indeksie dolnym jest w nawiasie ($\beta_{0\dots0q0\dots(j)0\dots0}^{kcr}$), oznacza to, że dany współczynnik może przyjmować różne wartości dla jednostek pierwszego poziomu należących do różnych jednostek na poziomie, któremu odpowiada pozycja indeksu w nawiasie. Sytuacja taka nie jest jednak możliwa przy zapisie jednym równaniem.

r_i^{kcr} - składnik resztowy z pierwszego poziomu, indeks górny tak samo jak w przypadku współczynników kierunkowych; $r_i^{kcr} \sim N(0; \sigma_{kcr}^2)$,

$e_{a0\dots0q0\dots(k)0\dots0}^{kcr}$ - składnik losowy związany z poziomem p -tym ze względu na pierwsze kryterium grupowania, stojący przy q -tej zmiennej objaśniającej z r -tego poziomu, $p = 2, \dots, P_a$, $r = 1, \dots, p-1$, $q = 1, \dots, Q_{ar}$,

$e_{b0\dots0q0\dots(k)0\dots0}^{kcr}$ - składnik losowy związany z poziomem p -tym ze względu na drugie kryterium grupowania, stojący przy q -tej zmiennej objaśniającej z r -tego poziomu, $p = 2, \dots, P_b$, $r = 1, \dots, p-1$, $q = 1, \dots, Q_{br}$,

$\tau_{a,p,0,0,0,0}^{kcr}$ - wariancja składnika losowego z p -tego poziomu ze względu na pierwsze kryterium grupowania w wyrazie wolnym, $p=2,\dots, P_a$,

$\tau_{b,p,0,0,0,0}^{kcr}$ - wariancja składnika losowego z p -tego poziomu ze względu na drugie kryterium grupowania w wyrazie wolnym, $p=2,\dots, P_b$,

$\tau_{a,p,r,q,r,q}^{kcr}$ - wariancja składnika losowego z p -tego poziomu ze względu na pierwsze kryterium grupowania we współczynniku kierunkowym q -tej zmiennej z r -tego poziomu, $p=2,\dots, P_a$, $r = 1, \dots, p-1$, $q=1,\dots, Q_{ar}$,

$\tau_{b,p,r,q,r,q}^{kcr}$ - wariancja składnika losowego z p -tego poziomu ze względu na drugie kryterium grupowania we współczynniku kierunkowym q -tej zmiennej z r -tego poziomu, $p=2,\dots, P_b$, $r = 1, \dots, p-1$, $q=1,\dots, Q_{br}$,

$\tau_{a,p,r_1,q_1,r_2,q_2}^{kcr}$ – kowariancja składnika losowego z p -tego poziomu ze względu na pierwsze kryterium grupowania we współczynniku kierunkowym q_1 -tej zmiennej z r_1 -tego poziomu oraz składnika losowego z p -tego poziomu ze względu na pierwsze kryterium grupowania we współczynniku kierunkowym q_2 -tej zmiennej z r_2 -tego poziomu, $p=2, \dots, P_a$. Jeżeli $r_1, q_1 = 0, 0$

lub $r_2, q_2 = 0, 0$, chodzi o składnik losowy wyrazu wolnego, $r_1, r_2 = 0, \dots, p-1$, $q_1 = 0, \dots, Q_{a,r_1}$, $q_2 = 1, \dots, Q_{a,r_2}$,

$\tau_{b,p,r_1,q_1,r_2,q_2}^{kcr}$ – kowariancja składnika losowego z p -tego poziomu ze względu na drugie kryterium grupowania we współczynniku kierunkowym q_1 -tej zmiennej z r_1 -tego poziomu oraz składnika losowego z p -tego poziomu ze względu

na drugie kryterium grupowania we współczynniku kierunkowym q_2 -tej zmiennej z r_2 -tego poziomu, $p=2, \dots, P_b$. Jeżeli $r_1, q_1 = 0, 0$ lub $r_2, q_2 = 0, 0$, chodzi

o składnik losowy wyrazu wolnego, $r_1, r_2 = 0, \dots, p-1$, $q_1 = 0, \dots, Q_{b,r_1}$, $q_2 = 1, \dots, Q_{b,r_2}$,

$\mathbf{e}_{a,p}^{kcr}$ – wektor zawierający wszystkie efekty losowe związane z poziomem p -tym ze względu na pierwsze kryterium grupowania, $p=1, \dots, P_a$. Zakłada się, że ma on wielowymiarowy rozkład normalny, ze średnią równą zero oraz macierzą kowariancji $\boldsymbol{\tau}_{a,p}^{kcr}$; $\mathbf{e}_{a,p}^{kcr} \sim N(\mathbf{0}; \boldsymbol{\tau}_{a,p}^{kcr})$,

$\mathbf{e}_{b,p}^{kcr}$ – wektor zawierający wszystkie efekty losowe związane z poziomem p -tym ze względu na drugie kryterium grupowania, $p=1, \dots, P_b$. Zakłada się, że ma on wielowymiarowy rozkład normalny, ze średnią równą zero oraz macierzą kowariancji $\boldsymbol{\tau}_{b,p}^{kcr}$; $\mathbf{e}_{b,p}^{kcr} \sim N(\mathbf{0}; \boldsymbol{\tau}_{b,p}^{kcr})$,

ETAP 1 Model pusty

W tym etapie zmienna \mathbf{Y}_i objaśniana jest przez przynależność jednostek z pierwszego poziomu do jednostek poziomów wyższych, ze względu na obydwa kryteria grupowania, bez użycia zmiennych objaśniających. Model rozszerzany jest w $P_a + P_b - 2$ podetapach.

W pierwszej kolejności uwzględniony zostaje efekt losowy z drugiego poziomu, ze względu na kryterium grupowania, z którym związana jest większa liczba poziomów. Bez straty ogólności przyjąć można, że jest to pierwsze kryterium grupowania. W drugim podetapie uwzględniany jest efekt losowy z drugiego poziomu ze względu na drugie kryterium grupowania itd.

W ostatnim etapie tak skonstruowany model można zapisać:

$$Y_i = \underbrace{\beta_{0 \dots 0}^{1P}}_P + \sum_{p=2}^{P_a} e^{\underbrace{1aP_a}_{1 \dots p \dots P_a} \underbrace{0 \dots 0(k)0 \dots 0}_{1 \dots p \dots P_a}} + \sum_{p=2}^{P_b} e^{\underbrace{1aP_a}_{1 \dots p \dots P_b} \underbrace{b0 \dots 0(k)0 \dots 0}_{1 \dots p \dots P_b}} + r_i^{1aP_a}, \quad (3.20)$$

gdzie:

$$r_i^{1aP_a} \sim N(0; \sigma_{1aP_a}^2) \quad , \quad e_{a0\dots0(k)0\dots0}^{1aP_a} \sim N(0; \tau_{a,p,0,0,0,0}^{1aP_a}) \quad , \quad \text{dla } p=2, \dots, P_a \quad ,$$

$$e_{b0\dots0(k)0\dots0}^{1aP_a} \sim N(0; \tau_{b,p,0,0,0,0}^{1aP_a}) \quad , \quad \text{dla } p=2, \dots, P_b \quad .$$

Jeżeli liczba poziomów ze względu na obydwie kryteria grupowania byłaby taka sama, badacz sam musi zdecydować, dla którego kryterium grupowania w pierwszej kolejności uwzględnić w modelu efekt losowy związany z drugim poziomem. W takiej sytuacji w ostatnim podetapie uwzględnia się efekt losowy ostatniego poziomu ze względu na to drugie kryterium grupowania.

Jeżeli powyższe równanie łączne modelu z etapu pierwszego zostałyby przedstawione w zapisie macierzowym (por. wzór (3.1)), macierz \mathbf{X} przyjęłaby postać n wymiarowego wektora jedynek, wektor $\boldsymbol{\beta}$ jednej liczby, $\beta_{0\dots0}^{1P}$. Macierz \mathbf{Z} byłaby wymiaru $n \times (\sum_{p=2}^{P_a} l_{a,p} + \sum_{p=2}^{P_b} l_{b,p})$, przy czym można ją traktować, jako macierz blokową, złożoną z macierzy zero-jedynkowych wymiaru, w którym kolumny odpowiadałyby jednostkom danego poziomu i jedynki pojawiałyby się na przecięciach z wierszami odpowiadającym tym jednostkom z poziomu pierwszego, które należą do danej jednostki tego poziomu. Wektor \mathbf{e} byłby wymiaru $\sum_{p=2}^{P_a} l_{a,p} + \sum_{p=2}^{P_b} l_{b,p}$ i zawierałby kolejno błędy losowe odpowiadające wszystkim jednostkom każdego poziomu wyższego niż drugi ze względu na obydwie kryteria podziału.

ETAP 2 Model z losowym wyrazem wolnym

W następnym etapie do modelu wprowadzone zostają zmienne objaśniające z pierwszego poziomu. Zakładamy jednak, że nachylenie krzywej regresji pozostanie stałe bez względu na przynależność na wyższych poziomach. Tak więc, w składzie współczynników kierunkowych przy dodanych zmiennych objaśniających z pierwszego poziomu występują jedynie stałe, równe dla wszystkich jednostek pierwszego poziomu. Zmienne należy, tak jak w przypadku klasycznego modelu wielopoziomowego, dodawać kolejno, aby uniknąć wprowadzania charakterystyk, których uwzględnienie nie poprawia jakości dopasowania modelu. Ponadto również oszacowane współczynniki kierunkowe powinny być istotne statystycznie. W celu wyboru zmiennych wprowadzonych do modelu można posłużyć się na przykład procedurą regresji krokowej.

$$Y_i = \sum_{q=0}^{Q_1} \underbrace{\beta_{q00\dots0}^{21}}_P X_{1,q,i} + \sum_{p=2}^{P_a} \underbrace{e_{a0\dots0(k)0\dots0}^{21}}_{1\dots p\dots P_a} + \sum_{p=2}^{P_b} \underbrace{e_{b0\dots0(k)0\dots0}^{21}}_{1\dots p\dots P_b} + r_i^{21} \quad , \quad (3.21)$$

gdzie:

$$r_i^{21} \sim N(0; \sigma_{21}^2) \quad , \quad \underbrace{e_{a0\dots0(k)0\dots0}^{21}}_{1\dots p\dots P} \sim N(0; \tau_{a,p,0,0,0,0}^{21}) \quad , \quad \text{dla } p=2, \dots, P_a \quad ,$$

$$\underbrace{e_{b0\dots0(k)0\dots0}^{21}}_{1\dots p\dots P} \sim N(0; \tau_{b,p,0,0,0,0}^{21}) \quad , \quad \text{dla } p=2, \dots, P_b.$$

Jeżeli powyższe równanie łączne modelu z etapu pierwszego zostałyby przedstawione w zapisie macierzowym (por. wzór (3.1)), macierz \mathbf{X} byłaby wymiaru $n \times (Q_1 + 1)$ przy czym jej pierwsza kolumna składałaby się z samych jedynek, kolejne zaś zawierałyby wartości zmiennych objaśniających z pierwszego poziomu, wektor $\boldsymbol{\beta}$ zawierałby wyraz wolny

oraz współczynniki kierunkowe odpowiadające tym zmiennym z pierwszego poziomu. Macierz \mathbf{Z} oraz wektor \mathbf{e} nie zmieniałyby się w stosunku do poprzedniego etapu.

ETAP 3 Model z losowym wyrazem wolnym zależnym od zmiennych z wyższych poziomów

W porównaniu z modelem z etapu drugiego, model wzbogacony zostanie o zmienne objaśniające z wyższych poziomów ze względu na oba kryteria grupowania. Rozszerzenie to następuje w $P_a + P_b - 2$ podetapach, w pierwszym z nich dołączane są zmienne z drugiego poziomu ze względu na pierwsze kryterium grupowania. W ostatnim podetapie dołączone są zmienne ze wszystkich poziomów ze względu na obydwa kryteria grupowania.

Łącznie:

$$Y_i = \sum_{q=0}^{Q_1} \underbrace{\beta_{q00\dots0}^{3aP_a}}_P X_{1,q,i} + \sum_{p=2}^{P_a} \sum_{q=1}^{Q_{a,p}} \underbrace{\beta_{a0\dots0q0\dots0}^{3aP_a}}_{1\dots p\dots P} X_{a,p,q,i} + \sum_{p=2}^{P_b} \sum_{q=1}^{Q_{b,p}} \underbrace{\beta_{b0\dots0q0\dots0}^{3aP_a}}_{1\dots p\dots P} X_{b,p,q,i} + \sum_{p=2}^{P_a} \underbrace{e_{a0\dots0(k)0\dots0}^{3aP_a}}_{1\dots p\dots P_a} + \sum_{p=2}^{P_b} \underbrace{e_{b0\dots0(k)0\dots0}^{3aP_a}}_{1\dots p\dots P_b} + r_i^{3aP_a} \quad (3.22)$$

gdzie:

$$r_i^{3aP_a} \sim N(0; \sigma_{3aP_a}^2) \quad , \quad \underbrace{e_{a0\dots0(k)0\dots0}^{3aP_a}}_{1\dots p\dots P} \sim N(0; \tau_{a,p,0,0,0,0}^{21}) \quad , \quad \text{dla } p=2, \dots, P_a \quad ,$$

$$\underbrace{e_{b0\dots0(k)0\dots0}^{3aP_a}}_{1\dots p\dots P} \sim N(0; \tau_{b,p,0,0,0,0}^{3aP_a}) \quad , \quad \text{dla } p=2, \dots, P_b.$$

Jeżeli powyższe równanie łączne modelu z etapu pierwszego zostałyby przedstawione w zapisie macierzowym (por. wzór (3.1)), macierz \mathbf{X} byłaby wymiaru $n \times (Q_1 + \sum_{p=2}^{P_a} Q_{a,p} + \sum_{p=2}^{P_b} Q_{b,p} + 1)$ przy czym jej pierwsza kolumna składałaby się z samych jedynek, kolejne zaś zawierałyby wartości zmiennych objaśniających z kolejnych poziomów ze względu na oba kryteria grupowania, wektor $\boldsymbol{\beta}$ zawierałby wyraz wolny oraz współczynniki kierunkowe odpowiadające tym zmiennym objaśniającym. Macierz \mathbf{Z} oraz wektor \mathbf{e} nie zmieniłyby się w stosunku do poprzednich etapów.

ETAP 4 Model z losowymi współczynnikami regresji

W tym etapie dopuszczona zostanie możliwość zróżnicowania jednostek pierwszego poziomu w ramach jednostek na wyższych poziomach ze względu na oba kryteria grupowania,

nie tylko ze względu na poziom szacowanej zmiennej, ale również ze względu na kształt zależności pomiędzy tą zmienną a zmiennymi objaśniającymi z niższych poziomów, ze względu na to samo kryterium grupowania. Uwzględniamy więc możliwość zróżnicowania wartości współczynników kierunkowych przy zmiennych z poziomów niższych pomiędzy jednostkami poziomów wyższych ze względu na to samo kryterium grupowania. Przeprowadzamy to w $[P_a \times (P_a - 1) + P_b \times (P_b - 1)]/2$ podetapach, w pierwszym z nich sprawdzając, dla których współczynników przy zmiennych z pierwszego poziomu ze względu na poziom drugi ze względu na to samo kryterium grupowania, uwzględnienie losowego charakteru poprawi jakość dopasowania modelu. Następnie sprawdza się zasadność uwzględniania losowego charakteru związanego z drugim poziomem ze względu na drugie kryterium grupowania we współczynnikach kierunkowych przy zmiennych z pierwszego poziomu. Czynniki losowe z wyższych poziomów we współczynnikach kierunkowych, uwzględniane są naprzemiennie ze względu na oba kryteria grupowania. W pierwszej kolejności, uwzględniane jest zróżnicowanie współczynników kierunkowych przy zmiennych objaśniających z pierwszego poziomu, następnie kolejno rozważane są współczynniki przy zmiennych z poziomów wyższych ze względu na oba kryteria grupowania. Należy pamiętać, że współczynniki kierunkowe przy zmiennych z drugiego oraz wyższych poziomów, ze względu na dane kryterium grupowania, wyjaśniane mogą być tylko przez składniki losowe związane z poziomami co najmniej o jeden wyższymi, ze względu na to samo kryterium grupowania.

Łącznie:

$Y_i =$

$$\begin{aligned}
 & \sum_{q=0}^{Q_1} \beta_{\underbrace{q00\dots0}_P}^{4a(P_a-1/P_a)} X_{1,q,i} + \sum_{p=2}^{P_a} \sum_{q=1}^{Q_{a,p}} \beta_{\underbrace{a0\dots0q0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} X_{a,p,q,i} + \\
 & \sum_{p=2}^{P_b} \sum_{q=1}^{Q_{b,p}} \beta_{\underbrace{b0\dots0q0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} X_{b,p,q,i} + \sum_{p=2}^{P_a} \sum_{q=0}^{Q_1} e_{\underbrace{q0\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} X_{1,q,i} + \\
 & \sum_{p=2}^{P_a} \sum_{r=2}^{p-1} \sum_{q=1}^{Q_{a,r}} e_{\underbrace{a0\dots0q0\dots0(k)0\dots0}_{1\dots r\dots p\dots P_a}}^{4a(P_a-1/P_a)} X_{a,r,q,i} + \sum_{p=2}^{P_b} \sum_{r=2}^{p-1} \sum_{q=1}^{Q_{b,r}} e_{\underbrace{a0\dots0q0\dots0(k)0\dots0}_{1\dots r\dots p\dots P_b}}^{4a(P_a-1/P_a)} X_{b,r,q,i} + \\
 & r_i^{4a(P_a-1/P_a)}, \tag{3.23}
 \end{aligned}$$

gdzie:

$$\begin{aligned}
 r_i^{4a(P_a-1/P_a)} \sim N(0; \sigma_{4a(P_a-1/P_a)}^2), \quad e_{a,p}^{4a(P_a-1/P_a)} = & \begin{bmatrix} e_{\underbrace{a00\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ e_{\underbrace{a10\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{\underbrace{aQ_1 0\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ e_{\underbrace{a01\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{\underbrace{a0\dots01(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{\underbrace{a0\dots0Q_{p-1}(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \end{bmatrix} \\
 , \quad e_{b,p}^{4a(P_a-1/P_a)} \sim N(0; \tau_{b,p}^{4a(P_a-1/P_a)}), \quad \text{dla } p=2, \dots, P_a, \quad e_{b,p}^{4a(P_a-1/P_a)} = & \begin{bmatrix} e_{\underbrace{b00\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ e_{\underbrace{b10\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{\underbrace{bQ_1 0\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ e_{\underbrace{b01\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{\underbrace{b0\dots01(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{\underbrace{b0\dots0Q_{p-1}(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \end{bmatrix}, \\
 e_{b,p}^{4a(P_a-1/P_a)} \sim N(0; \tau_{b,p}^{4a(P_a-1/P_a)}), \quad \text{dla } p=2, \dots, P_b. &
 \end{aligned}$$

Jeżeli powyższe równanie łączne modelu z etapu pierwszego zostałyby przedstawione w zapisie macierzowym (por. wzór (3.1)), macierz \mathbf{X} oraz wektor $\boldsymbol{\beta}$ pozostałyby takie same jak w etapie poprzednim. Macierz \mathbf{Z} byłaby wymiaru: $n \times ((\sum_{p=2}^{P_a} l_{a,p} * (1 + \sum_{r=1}^{p-1} Q_{a,r})) + (\sum_{p=2}^{P_b} l_{b,p} * (1 + \sum_{r=1}^{p-1} Q_{b,r})))$, przy czym można ją traktować, jako macierz blokową, złożoną z macierzy zero-jedynkowych, w których kolumny odpowiadałyby jednostkom poziomu p -tego i jedynki pojawiałyby się na przecięciach z wierszami odpowiadającymi tym jednostkom z poziomu pierwszego, które należą do danej jednostki p -tego poziomu ze względu na odpowiednie kryterium grupowania oraz macierzy o takim samym wymiarze, jednak złożonych z wartości zmiennych objaśniających z niższych niż p -ty poziomów dla tego samego kryterium grupowania. Wektor \mathbf{e} byłby wymiaru $(\sum_{p=2}^{P_a} l_{a,p} * (1 + \sum_{r=1}^{p-1} Q_{a,r})) + (\sum_{p=2}^{P_b} l_{b,p} * (1 + \sum_{r=1}^{p-1} Q_{b,r}))$

i zawierałby kolejno błędy losowe wchodzące w skład wyrazu wolnego oraz współczynników kierunkowych przy zmiennych objaśniających z niższych poziomów, odpowiadające wszystkim jednostkom każdego poziomu wyższego niż drugi ze względu na oba kryteria grupowania.

ETAP 5 Model z losowymi współczynnikami regresji zależnym od zmiennych z wyższych poziomów

W etapie piątym część zróżnicowania współczynników kierunkowych przy zmiennych objaśniających związana z podziałem na jednostki poziomu wyższego niż ten, dla którego określona jest zmienna przy której stoi odpowiedni współczynnik, wyjaśniona może zostać przez zmienne objaśniające z tego samego, bądź wyższego poziomu z którym związane jest to zróżnicowanie. W etapie poprzednim współczynniki kierunkowe wyjaśniane były wyłącznie przez efekty losowe związane z wyższymi poziomami ze względu na to samo kryterium grupowania. Podobnie w tym etapie, zmienność współczynnika kierunkowego przy zmiennej określonej dla poziomu wyższego niż pierwszy ze względu na dane kryterium grupowania może być wyjaśniana wyłącznie przez zmienne określone dla wyższych poziomów tego samego kryterium grupowania. W ten sposób, w równaniu opisującym model pojawią się iloczyny zmiennych objaśniających z różnych poziomów, ale tylko w ramach kryteriów grupowania.

Rozszerzanie modelu w tym etapie, podobnie jak w etapie 4, następuje w $[P_a \times (P_a - 1) + P_b \times (P_b - 1)]/2$ podetapach. W pierwszym z nich współczynniki kierunkowe przy zmiennych objaśniających z poziomu pierwszego wyjaśniane zostają przez zmienne z poziomu drugiego ze względu na pierwsze kryterium grupowania. W kolejnych podetapach współczynniki przy zmiennych z poziomu pierwszego wyjaśniane zostają również przez zmienne objaśniające z drugiego poziomu ze względu na drugie kryterium grupowania, trzeciego poziomu ze względu na pierwsze kryterium grupowania itd. W ostatnim podetapie współczynniki kierunkowe przy zmiennych z przedostatniego poziomu wyjaśniane zostają przez zmienne z poziomu ostatniego ze względu na kryterium grupowania, dla którego określono więcej poziomów.

Łącznie:

$$Y_i =$$

$$\begin{aligned} & \sum_{q=0}^{Q_1} \beta_{\underbrace{q00\dots0}_P}^{4a(P_a-1/P_a)} X_{1,q,i} + \sum_{A<Z_a} \beta_{\underbrace{aa_1a_2\dots a_p}_P}^{5a(P_a-1/P_a)} \prod_{j=1}^{P_a} X_{a,j,a_j,i} + \\ & \sum_{B<Z_b} \beta_{\underbrace{bb_1b_2\dots b_p}_P}^{5a(P_a-1/P_a)} \prod_{j=1}^{P_b} X_{b,j,b_j,i} + \sum_{p=2}^{P_a} \sum_{q=0}^{Q_1} e_{\underbrace{q0\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} X_{1,q,i} + \\ & \sum_{p=2}^{P_a} \sum_{r=2}^{p-1} \sum_{q=1}^{Q_{a,r}} e_{\underbrace{a0\dots0q0\dots0(k)0\dots0}_{1\dots r\dots p\dots P_a}}^{4a(P_a-1/P_a)} X_{a,r,q,i} + \sum_{p=2}^{P_b} \sum_{r=2}^{p-1} \sum_{q=1}^{Q_{b,r}} e_{\underbrace{a0\dots0q0\dots0(k)0\dots0}_{1\dots r\dots p\dots P_b}}^{4a(P_a-1/P_a)} X_{b,r,q,i} + \\ & r_i^{4a(P_a-1/P_a)}, \end{aligned} \tag{3.24}$$

$$r_i^{4a(P_a-1/P_a)} \sim N(0; \sigma_{4a(P_a-1/P_a)}^2), e_{a,p}^{4a(P_a-1/P_a)} = \begin{bmatrix} e_{\underbrace{a00\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ e_{\underbrace{a10\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{\underbrace{aQ_10\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ e_{\underbrace{a01\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{\underbrace{a0\dots01(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{\underbrace{a0\dots0Q_{p-1}(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \end{bmatrix},$$

$$\mathbf{e}_{a,p}^{4a(P_a-1/P_a)} \sim N\left(\mathbf{0}; \boldsymbol{\tau}_{a,p}^{4a(P_a-1/P_a)}\right), \text{ dla } p=2, \dots, P_a, \quad \mathbf{e}_{b,p}^{4a(P_a-1/P_a)} = \begin{bmatrix} e_{b \underbrace{00\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ e_{b \underbrace{10\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{b \underbrace{Q_1 0\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ e_{b \underbrace{01\dots0(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{b \underbrace{0\dots01(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \\ \vdots \\ e_{b \underbrace{0\dots0Q_{p-1}(k)0\dots0}_{1\dots p\dots P}}^{4a(P_a-1/P_a)} \end{bmatrix},$$

$$\mathbf{e}_{b,p}^{4a(P_a-1/P_a)} \sim N\left(\mathbf{0}; \boldsymbol{\tau}_{b,p}^{4a(P_a-1/P_a)}\right), \text{ dla } p=2, \dots, P_b.$$

gdzie:

Z_a – zbiór wszystkich możliwych zestawów zmiennych po jednej z każdego poziomu ze względu na pierwsze kryterium grupowania. Pod uwagę dla każdego poziomu brane są również „zmiennne zerowe”, czyli zmiennne przyjmujące w całej populacji stałą wartość 1. Ze zbioru tego wyłączone są tylko takie zestawy zmiennych, w których nie ma zmiennych z poziomów wyższych niż pierwszy. Zatem $\bar{Z} = \prod_{p=1}^{P_a} (Q_p + 1) - (Q_1 + 1)$, ponadto elementami zbioru Z_a są zbiory, które z kolei zawierają po P_a elementów każdy, odpowiadającym numerom zmiennych objaśniających z kolejnych poziomów ze względu na pierwsze kryterium grupowania. Inaczej można to zapisać w następującej postaci: jeżeli $A \subset Z$, to $\bar{A} = P$, $A = \{a_1, a_2, \dots, a_p\}$, przy czym dla $p=1, \dots, P_a$, $a_p \in \{0, 1, \dots, Q_{a,p}\}$,

Z_b – zbiór analogiczny do Z_a , ale ze względu na drugie kryterium grupowania.

Jeżeli powyższe równanie łączne modelu z etapu pierwszego zostałyby przedstawione w zapisie macierzowym (por. wzór (3.1)), macierz \mathbf{X} byłaby wymiaru $n \times (\prod_{p=1}^{P_a} (Q_{a,p} + 1) + \prod_{p=1}^{P_b} (Q_{b,p} + 1) - (Q_1 + 1))$, przy czym jej kolumny odpowiadałyby wyrazowi wolnemu, zmiennym z pierwszego poziomu oraz wszystkim możliwym iloczynom zmiennych należących do zestawów ze zbiorów Z_a i Z_b . Wektor $\boldsymbol{\beta}$ byłby $(\prod_{p=1}^{P_a} (Q_{a,p} + 1) + \prod_{p=1}^{P_b} (Q_{b,p} + 1) - (Q_1 + 1))$ wymiarowy i zawierałby współczynniki przy wszystkich możliwych iloczynach zmiennych z różnych poziomów (wliczając w to wyraz wolny oraz zmiennne pojedyncze). Z kolei część losowa układu nie uległaby zmianie w stosunku do etapu poprzedniego.

Dalsze komplikowanie modelu, podobnie jak w przypadku klasycznego modelu wielopoziomowego, jest teoretycznie możliwe (w przypadku więcej niż dwupoziomowym ze względu na co najmniej jedno kryterium grupowania). Można mianowicie wprowadzić elementy losowe we współczynnikach przy interakcjach pomiędzy zmiennymi. Jednak przy wielopoziomowych modelach krzyżowych powodowałoby to jeszcze większe problemy interpretacyjne oraz obliczeniowe. Zatem w rozdziale 4.2 prezentowanej pracy skonstruowany zostanie model opisany przy pomocy pięciu powyższych etapów.

3.4. Estymacja parametrów modelu

Jak wspomniano w podrozdziale 3.2, do szacowania parametrów modelu stosuje się zapis przedstawiony przy pomocy równania (3.1) w postaci macierzowej. W podrozdziałach 3.2 oraz 3.3 opisano formę, jaką przyjmują składniki równania (3.1), w kolejnych etapach rozszerzania modelu wielopoziomowego, zarówno klasycznego jak i z dwoma kryteriami grupowania. Chociaż modele te różniły się pod względem strategii ich budowy, parametry modelu szacowane są dokładnie w taki sam sposób dla obu przypadków (por. Bates, 2010).

W celu estymacji współczynników modelu wielopoziomowego najczęściej stosuje się metodę największej wiarygodności (ML maximum likelihood) bądź resztową metodę największej wiarygodności (REML lub RML restricted/residual maximum likelihood). W rzeczywistości do klasycznego modelu regresji liniowej, estymatory resztowej największej wiarygodności, w przypadku modeli wielopoziomowych mogą być obciążone. Uważa się jednak, że obciążenie to w większości wypadków jest mniejsze, niż przy zastosowaniu metody największej wiarygodności (por. Biecek, 2011). Z drugiej strony, estymatory parametrów szacowane przy pomocy resztowej metody największej wiarygodności, pomimo takiej nazwy, nie charakteryzują się największą wiarygodnością. Ponadto modele szacowane tą metodą nie mogą być traktowane jako rozwinięcia modeli „uboższych”, jeżeli modyfikacja dotyczyła części stałej modelu, zatem nie można ich porównywać przy pomocy testu ilorazu wiarygodności (por. Frątczak i in., 2009; Bates, 2010; Hox, 2010; Biecek, 2011). Jest tak dlatego, że estymatory REML obliczane są po przeniesieniu części stałej modelu na lewą stronę równia (3.1). W związku z tym przy każdej modyfikacji części stałej modelu, zmienia się część objaśniana modelu, Zatem kolejne modele nie mogą być traktowane jako rozszerzenia wcześniejszych. Ponadto

w literaturze podaje się, że wpływ wyboru metody ML lub REML na otrzymane oszacowania dla modeli wielopoziomowych jest mniejszy niż się powszechnie uważa (por. Bates, 2010; Hox, 2010; Kreft, de Leeuw, 1998). W związku z powyższymi uwagami w następnym, czwartym rozdziale prezentowanej pracy wykorzystano, przy konstruowaniu modeli metodę największej wiarygodności (ML).

Ponadto w dostępnym oprogramowaniu estymacja parametrów modelu, zarówno w przypadku klasycznej jak i resztowej metody największej wiarygodności, przeprowadzona może zostać przy pomocy różnych algorytmów. Niektóre z nich to algorytm Newtona-Raphsona (wykorzystywana w programie SAS, por. Wolfinger, Tobias, Sall, 1994) oraz metoda z wykorzystaniem operacji na macierzach rzadkich (w pakiecie lme4 programu R, por. Bates, 2010). W prezentowanej pracy obliczenia wykonane zostały przy pomocy pakietu lme4 programu R. W dalszej części tego podrozdziału opisano algorytm szacowania parametrów modelu stosowany w tym pakiecie.

ESTYMATORY NAJWIĘKSZEJ WIAROGODNOŚCI

Rozważmy ogólny wielopoziomowy model liniowy opisany równaniem (3.1). Jak zauważono wcześniej, przy pomocy tego równania zapisać można dowolny model wielopoziomowy, zarówno klasyczny jak i z dwoma kryteriami grupowania. Ponieważ estymacja parametrów modelu przebiegać będzie dla poszczególnych elementów tego zapisu, został on powtórzony, ponadto określone zostały wymiary poszczególnych składowych

$$Y = X\beta + Ze + r \quad , \quad (3.1)$$

gdzie:

Y – wektor zmiennej objaśnianej o wymiarze n ,

X – macierz zmiennych objaśniających dla efektów stałych modelu o wymiarze $n \times p$,

β – wektor współczynników kierunkowych o wymiarze p ,

Z – macierz zmiennych objaśniających dla efektów losowych modelu o wymiarze $n \times q$,

e – wektor efektów losowych o wymiarze q , $e \sim N_q(\mathbf{0}; \Sigma)$,

r – składnik resztowy modelu o wymiarze n , $r \sim N_n(\mathbf{0}; \sigma^2 \mathbf{I}_n)$.

Należy zwrócić uwagę, że n , podobnie jak w poprzednich podrozdziałach, określa liczbę jednostek pierwszego poziomu, jednak p i q są odpowiednio wymiarami macierzy

efektów stałych oraz losowych. Nie można utożsamiać ich z liczbą zmiennych objaśniających czy liczbą poziomów modelu. W podrozdziałach 3.2 oraz 3.3 przedstawiono te wymiary dla modeli opisywanych w kolejnych etapach.

W równaniu (3.1) znane są wartości: wektora zmiennej objaśnianej \mathbf{Y} (tylko z próby uczącej), macierzy zmiennych objaśniających dla efektów stałych \mathbf{X} oraz macierzy \mathbf{Z} . Celem rozważań jest estymacja wektora $\boldsymbol{\beta}$ oraz macierzy kowariancji wektora efektów losowych \mathbf{e} , czyli $\boldsymbol{\Sigma}$. Parametry te oszacowane zostaną przy pomocy metody największej wiarygodności.

W pierwszej kolejności, przeanalizowana zostanie postać macierzy kowariancji $\boldsymbol{\Sigma}$. Załóżmy, że omawiany model charakteryzuje się strukturą P poziomową. Ponadto ze względu na kolejne poziomy uwzględnia się k_i (dla $i=1, \dots, P$) efektów losowych (wliczając w to efekt losowy wyrazu wolnego oraz współczynników kierunkowych przy zmiennych z niższych poziomów). Tak więc dla każdego poziomu należy oszacować k_i wartości wariancji oraz $k_i \times (k_i - 1) / 2$ kowariancji, czyli w sumie $k_i \times (k_i + 1) / 2$ wartości w macierzy kowariancji, związanych z danym poziomem. Ponieważ, na każdym poziomie jest określona liczba jednostek, z których efekty losowe dla każdej z nich charakteryzują się takim samym rozkładem normalnym, opisane powyżej wartości pojawiają się w macierzy $\boldsymbol{\Sigma}$ tyle razy, z ilu jednostek składa się dany poziom (dwukrotnie więcej w przypadku kowariancji). Ponieważ zakłada się, że nie występują korelacje losowe pomiędzy efektami losowymi z różnych poziomów, pozostałe wartości macierzy $\boldsymbol{\Sigma}$ są zerami. Tak więc, przy optymalnej kolejności kolumn w macierzy \mathbf{Z} (którą przyjmując można bez straty ogólności, a która w procedurze lmer uzyskiwana jest przy pomocy prostej transformacji), macierz $\boldsymbol{\Sigma}$ jest macierzą blokową diagonalną, wokół diagonali której pojawiają się kolejno macierze kowariancji z kolejnych poziomów, każda tyle razy, ile jednostek określonych jest na danym poziomie. Zatem przy potencjalnie dużej liczbie elementów macierzy $\boldsymbol{\Sigma}$ (q^2), w celu jednoznacznego jej określenia konieczne jest oszacowanie relatywnie małej liczby jej elementów $\sum_{i=1}^P k_i(k_i + 1) / 2$. Wektor tych parametrów oznaczamy będziemy $\boldsymbol{\theta}$. W związku z tym również macierz kowariancji $\boldsymbol{\Sigma}$ powinna być oznaczana jako $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$.

O zmiennej objaśnianej zakłada się, że ma rozkład normalny, zatem jeżeli przyjmując konkretną wartość wektora efektów losowych, \mathbf{e} , zmienna objaśniana charakteryzuje się następującym warunkowym rozkładem normalnym:

$$(\mathbf{y}|\mathbf{e}) \sim N_n(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{e}; \sigma^2 \mathbf{I}_n) . \quad (3.25)$$

Z kolei wektor efektów losowych charakteryzuje się następującym rozkładem normalnym:

$$\mathbf{e} \sim N_q(\mathbf{0}; \boldsymbol{\Sigma}_\theta) . \quad (3.26)$$

O macierzy $\boldsymbol{\Sigma}_\theta$, jako o macierzy kowariancji, wiadomo, że jest macierzą symetryczną, ponadto jest dodatnio półokreślona, tzn:

$$\bigwedge_{\mathbf{b} \in R^q / \{0\}} \mathbf{b}^T \boldsymbol{\Sigma}_\theta \mathbf{b} \geq 0 .$$

Niech macierz $\boldsymbol{\Lambda}_\theta$ będzie macierzą trójkątną dolną, spełniającą warunek:

$$\boldsymbol{\Sigma}_\theta = \sigma^2 \boldsymbol{\Lambda}_\theta \boldsymbol{\Lambda}_\theta^T . \quad (3.27)$$

Wiadomo, że macierz spełniająca powyższy warunek istnieje, ponieważ macierz $\boldsymbol{\Sigma}_\theta$ jest macierzą symetryczną.

Następnie wektor efektów losowych \mathbf{e} przedstawiony zostaje jako:

$$\mathbf{e} = \boldsymbol{\Lambda}_\theta \mathbf{u} . \quad (3.28)$$

Łatwo można wykazać, że wektor \mathbf{u} charakteryzuje się następującym rozkładem normalnym:

$$\mathbf{u} \sim N_q(\mathbf{0}; \sigma^2 \mathbf{I}_q) . \quad (3.29)$$

Jednocześnie równanie (3.1) oraz rozkład warunkowy zmiennej objaśnianej zapisać można:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u} + \mathbf{r} , \quad (3.30)$$

$$(\mathbf{y}|\mathbf{u}) \sim N_n(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u}; \sigma^2 \mathbf{I}_n) . \quad (3.31)$$

W wyniku powyższego zapisu, obiektem zainteresowania stała się macierz $\boldsymbol{\Lambda}_\theta$ zamiast macierzy $\boldsymbol{\Sigma}_\theta$ oraz wektor \mathbf{u} zamiast wektora \mathbf{e} . Zamiana taka jest korzystna ze względu na prostszy rozkład wektora \mathbf{u} .

Aby zastosować metodę największej wiarygodności konieczna jest znajomość rozkładu łącznego wektorów \mathbf{u} oraz \mathbf{y} , w celu wyznaczenia funkcji wiarygodności opisanej wzorem:

$$L(\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma | \mathbf{y}_{obs}) = \int_{R^q} f_{u,y} du . \quad (3.32)$$

Korzystając z twierdzenia Bayesa, odpowiednią gęstość warunkową można zapisać w postaci:

$$f_{u,y} = f_{y|u} f_u = f_u |y f_y \quad (3.33)$$

Ponieważ znane są rozkłady $(\mathbf{y}|\mathbf{u})$ oraz \mathbf{u} (por. odpowiednio wzory (3.31) oraz (3.29)), w celu wyznaczenia gęstości łącznej wyznaczane są odpowiednie gęstości (por. Krzyśko, 2009):

$$\begin{aligned} f_{\mathbf{y}|\mathbf{u}} &= (2\pi)^{-\frac{n}{2}} |\sigma^2 \mathbf{I}_n|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u})^T (\sigma^2 \mathbf{I}_n)^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u}) \right] = \\ &= (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u}) \right] = \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u}) \right], \end{aligned} \quad (3.34)$$

$$\begin{aligned} f_{\mathbf{u}} &= (2\pi)^{-\frac{q}{2}} |\sigma^2 \mathbf{I}_q|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \mathbf{u}^T (\sigma^2 \mathbf{I}_q)^{-1} \mathbf{u} \right] = (2\pi)^{-\frac{q}{2}} (\sigma^2)^{-\frac{q}{2}} \exp \left[-\frac{1}{2\sigma^2} \mathbf{u}^T \mathbf{u} \right] = \\ &= (2\pi\sigma^2)^{-\frac{q}{2}} \exp \left[-\frac{1}{2\sigma^2} \mathbf{u}^T \mathbf{u} \right]. \end{aligned} \quad (3.35)$$

Podstawiając tak wyznaczone gęstości do równania (3.33) wyznaczono gęstość łączną jako:

$$f_{\mathbf{u},\mathbf{y}} = (2\pi\sigma^2)^{-\frac{n+q}{2}} \exp \left\{ -\frac{1}{2\sigma^2} [(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u}) + \mathbf{u}^T \mathbf{u}] \right\}. \quad (3.36)$$

Niech $\tilde{\mathbf{u}}$ będzie wartością wektora \mathbf{u} maksymalizującą powyższą funkcję gęstości łącznej. Ponadto niech \mathbf{y}_{obs} oznacza wektor wartości (stały), które zmienna objaśniana przyjmuje w próbie.

$$\tilde{\mathbf{u}} = \arg \max_{\mathbf{u}} f_{\mathbf{u}|\mathbf{y}}(\mathbf{u}|\mathbf{y}_{obs}) = \arg \max_{\mathbf{u}} f_{\mathbf{u},\mathbf{y}}(\mathbf{u}, \mathbf{y}_{obs}) = \arg \max_{\mathbf{u}} f_{\mathbf{y}|\mathbf{u}}(\mathbf{y}_{obs}|\mathbf{u}) f_{\mathbf{u}}(\mathbf{u}). \quad (3.37)$$

Powyższa równość jest prawdziwa, ponieważ jest wartością funkcji w punkcie, a więc $f_{\mathbf{y}}(\mathbf{y}_{obs})$ jest stałą.

Następnie wyznaczany jest logarytm naturalny łącznej gęstości pomnożony przez (-2):

$$\begin{aligned} -2 \ln f_{\mathbf{u},\mathbf{y}}(\mathbf{u}, \mathbf{y}_{obs}) &= -2 \ln \left\{ (2\pi\sigma^2)^{-\frac{n+q}{2}} \exp \left\{ -\frac{1}{2\sigma^2} [(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \right. \right. \\ &\left. \left. \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u}) + \mathbf{u}^T \mathbf{u}] \right\} \right\} = (n+q) \ln(2\pi\sigma^2) + \frac{1}{\sigma^2} ((\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\boldsymbol{\Lambda}_\theta \mathbf{u}) + \\ &\mathbf{u}^T \mathbf{u}). \end{aligned} \quad (3.38)$$

Ponieważ logarytm naturalny jest funkcją rosnącą w swojej dziedzinie, powyższa funkcja jest malejąca. W związku z tym argument, dla którego funkcja $f_{\mathbf{u},\mathbf{y}}$ osiąga maksimum, to ten sam argument, dla którego funkcja $-2 \ln f_{\mathbf{u},\mathbf{y}}$ osiąga minimum. Ponadto we wzorze (3.38) od \mathbf{u} zależy wyłącznie drugi składnik sumy, a dokładnie druga część iloczynu wchodzącego w jego skład. W związku z tym równanie (3.37) można przekształcić do postaci:

$$\tilde{\mathbf{u}} = \arg \min_{\mathbf{u}} ((\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\Lambda_{\theta}\mathbf{u})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\Lambda_{\theta}\mathbf{u}) + \mathbf{u}^T \mathbf{u}). \quad (3.39)$$

W powyższym wzorze wyrażenie, z którego szukane jest minimum, nazywane jest sumą kwadratów reszt z karą (PRSS penalized residual sum of squares). Z kolei wartość \mathbf{u} , dla którego to minimum jest uzyskiwane, metodą najmniejszych kwadratów z karą (PLS penalized least squares solution) (por. Bates, 2013c).

Niech:

$$r_{\theta, \boldsymbol{\beta}}^2 = \min_{\mathbf{u}} ((\mathbf{y} - \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\Lambda_{\theta}\mathbf{u})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\Lambda_{\theta}\mathbf{u}) + \mathbf{u}^T \mathbf{u}) \quad (3.40)$$

$$r_{\theta}^2 = \min_{\mathbf{u}, \boldsymbol{\beta}} ((\mathbf{y} - \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\Lambda_{\theta}\mathbf{u})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\Lambda_{\theta}\mathbf{u}) + \mathbf{u}^T \mathbf{u}) \quad (3.41)$$

Obydwa powyższe minima wyznaczone mogą zostać bez użycia procedury iteracyjnej. Ponadto wartość wektora $\boldsymbol{\beta}$ (zależna od parametru θ) dla której osiągnane jest minimum (3.40) oznaczana będzie dalej jako $\hat{\boldsymbol{\beta}}_{\theta}$.

W celu rozwiązania równania (3.39) minimalizowaną sumę przedstawić można jako:

$$\begin{aligned} \tilde{\mathbf{u}} = \\ \arg \min_{\mathbf{u}} \left(\begin{bmatrix} \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \\ \mathbf{u} \end{bmatrix}^T \begin{bmatrix} \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \\ \mathbf{u} \end{bmatrix} \right) = \\ \arg \min_{\mathbf{u}} \left(\left(\begin{bmatrix} \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{Z}\Lambda_{\theta} \\ \mathbf{I}_q \end{bmatrix} \mathbf{u} \right)^T \left(\begin{bmatrix} \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{Z}\Lambda_{\theta} \\ \mathbf{I}_q \end{bmatrix} \mathbf{u} \right) \right). \end{aligned} \quad (3.42)$$

Dla tak zapisanego problemu rozwiązanie znaleźć można przy pomocy klasycznej metody najmniejszych kwadratów. Jest to wartość spełniająca następujące równanie:

$$\begin{aligned} \begin{bmatrix} \mathbf{Z}\Lambda_{\theta} \\ \mathbf{I}_q \end{bmatrix}^T \begin{bmatrix} \mathbf{Z}\Lambda_{\theta} \\ \mathbf{I}_q \end{bmatrix} \tilde{\mathbf{u}} &= \begin{bmatrix} \mathbf{Z}\Lambda_{\theta} \\ \mathbf{I}_q \end{bmatrix}^T \begin{bmatrix} \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \\ \mathbf{0} \end{bmatrix}, \\ ((\mathbf{Z}\Lambda_{\theta})^T \mathbf{Z}\Lambda_{\theta} + \mathbf{I}_q) \tilde{\mathbf{u}} &= \mathbf{Z}\Lambda_{\theta} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}), \\ (\Lambda_{\theta}^T \mathbf{Z}^T \mathbf{Z}\Lambda_{\theta} + \mathbf{I}_q) \tilde{\mathbf{u}} &= \mathbf{Z}\Lambda_{\theta} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ \tilde{\mathbf{u}} &= (\Lambda_{\theta}^T \mathbf{Z}^T \mathbf{Z}\Lambda_{\theta} + \mathbf{I}_q)^{-1} \mathbf{Z}\Lambda_{\theta} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \end{aligned} \quad (3.43)$$

Rozwiązanie tego równania nie jest oczywiście wyznaczone przez obliczanie macierzy odwrotnej, ale przy pomocy metody Choleskiego. Tak więc, przy pomocy dekompozycji Choleskiego dla macierzy rzadkich wyznaczana jest macierz trójkątna dolna \mathbf{L}_{θ} spełniająca warunek:

$$\mathbf{L}_{\theta} \mathbf{L}_{\theta}^T = \Lambda_{\theta}^T \mathbf{Z}^T \mathbf{Z}\Lambda_{\theta} + \mathbf{I}_q. \quad (3.44)$$

Podstawiając (3.44) do wzoru (3.43) otrzymujemy:

$$\mathbf{L}_\theta \mathbf{L}_\theta^T \tilde{\mathbf{u}} = \mathbf{Z} \Lambda_\theta (\mathbf{y} - \mathbf{X} \boldsymbol{\beta}) . \quad (3.45)$$

Przy pomocy metody Choleskiego możliwe jest rozwiązanie powyższego równania.

Następnie rozwiązanie to podstawione zostanie do równania (3.40):

$$r_{\theta, \beta}^2 = (\mathbf{y} - \mathbf{X} \boldsymbol{\beta} + \mathbf{Z} \Lambda_\theta \tilde{\mathbf{u}})^T (\mathbf{y} - \mathbf{X} \boldsymbol{\beta} + \mathbf{Z} \Lambda_\theta \tilde{\mathbf{u}}) + \tilde{\mathbf{u}}^T \tilde{\mathbf{u}} . \quad (3.46)$$

Ponadto dla dowolnego \mathbf{u} prawdziwa jest równość:

$$(\mathbf{y} - \mathbf{X} \boldsymbol{\beta} + \mathbf{Z} \Lambda_\theta \mathbf{u})^T (\mathbf{y} - \mathbf{X} \boldsymbol{\beta} + \mathbf{Z} \Lambda_\theta \mathbf{u}) + \mathbf{u}^T \mathbf{u} = r_{\theta, \beta}^2 + \left(\mathbf{L}_\theta^T (\mathbf{u} - \tilde{\mathbf{u}}) \right)^T \left(\mathbf{L}_\theta^T (\mathbf{u} - \tilde{\mathbf{u}}) \right) . \quad (3.47)$$

Podstawiając powyższe równanie do równania (3.36) otrzymujemy:

$$f_{u,y} = (2\pi\sigma^2)^{-\frac{n+q}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \left[r_{\theta, \beta}^2 + \left(\mathbf{L}_\theta^T (\mathbf{u} - \tilde{\mathbf{u}}) \right)^T \left(\mathbf{L}_\theta^T (\mathbf{u} - \tilde{\mathbf{u}}) \right) \right] \right\} . \quad (3.48)$$

Na potrzeby dalszych rozważań zdefiniowano następującą zmienną pomocniczą \mathbf{z} :

$$\mathbf{z} = \frac{\mathbf{L}_\theta^T (\mathbf{u} - \tilde{\mathbf{u}})}{\sigma} . \quad (3.49)$$

Tak zdefiniowana zmienna ma następującą własność:

$$\left| \frac{d\mathbf{z}}{d\mathbf{u}} \right| = \left| \frac{\mathbf{L}_\theta^T}{\sigma} \right| = \frac{|\mathbf{L}_\theta|}{\sigma} . \quad (3.50)$$

W tym momencie możliwe jest wyznaczenie wiarygodności:

$$\begin{aligned} L(\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma | \mathbf{y}_{obs}) &= \int_{R^q} f_{u,y} d\mathbf{u} = \\ &= \int_{R^q} (2\pi\sigma^2)^{-\frac{n+q}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \left[r_{\theta, \beta}^2 + \left(\mathbf{L}_\theta^T (\mathbf{u} - \tilde{\mathbf{u}}) \right)^T \left(\mathbf{L}_\theta^T (\mathbf{u} - \tilde{\mathbf{u}}) \right) \right] \right\} d\mathbf{u} = \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left(-\frac{r_{\theta, \beta}^2}{2\sigma^2} \right) \int_{R^q} (2\pi)^{-\frac{q}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \left(\mathbf{L}_\theta^T (\mathbf{u} - \tilde{\mathbf{u}}) \right)^T \left(\mathbf{L}_\theta^T (\mathbf{u} - \tilde{\mathbf{u}}) \right) \right\} \frac{|\mathbf{L}_\theta|}{|\mathbf{L}_\theta| \sigma^q} d\mathbf{u} = \\ &= \left| \frac{\mathbf{L}_\theta^T (\mathbf{u} - \tilde{\mathbf{u}})}{\sigma} \right| = \left(2\pi\sigma^2 \right)^{-\frac{n}{2}} \exp \left(-\frac{r_{\theta, \beta}^2}{2\sigma^2} \right) |\mathbf{L}_\theta|^{-1} \int_{R^q} (2\pi)^{-\frac{q}{2}} \exp \left(-\frac{\mathbf{z}^T \mathbf{z}}{2} \right) d\mathbf{z} = \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left(-\frac{r_{\theta, \beta}^2}{2\sigma^2} \right) |\mathbf{L}_\theta|^{-1} \mathbf{1} = (2\pi\sigma^2)^{-\frac{n}{2}} \exp \left(-\frac{r_{\theta, \beta}^2}{2\sigma^2} \right) |\mathbf{L}_\theta|^{-1} \end{aligned} \quad (3.51)$$

Podwojony logarytm naturalny wiarygodności nazywany jest dewiacją i oznaczany:

$$d(\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma | \mathbf{y}_{obs}) = -2 \ln L(\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma | \mathbf{y}_{obs}) = n \ln(2\pi\sigma^2) + 2 \ln |\mathbf{L}_\theta| + \frac{r_{\theta, \beta}^2}{\sigma^2} . \quad (3.52)$$

Estymatorami największej wiarygodności nazywamy te oszacowania parametrów, które minimalizują powyższą dewiację. Niech $\widehat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}$ oznacza tę wartość $\boldsymbol{\beta}$, która minimalizuje sumę kwadratów reszt z karą (PRSS) (por. Bates, 2010) dla ustalonego $\boldsymbol{\theta}$ oraz niech $r_{\boldsymbol{\theta}}^2$ oznacza sumę kwadratów reszt z karą dla $\boldsymbol{\beta} = \widehat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}$. Wtedy:

$$\check{d}(\boldsymbol{\theta}, \sigma | \mathbf{y}_{obs}) = n \ln(2\pi\sigma^2) + 2 \ln|\mathbf{L}_{\boldsymbol{\theta}}| + \frac{r_{\boldsymbol{\theta}}^2}{\sigma^2}. \quad (3.53)$$

Ponadto niech $\widehat{\sigma}_{\boldsymbol{\theta}}^2$ będzie wartością σ^2 , która minimalizuje dewiację dla ustalonego $r_{\boldsymbol{\theta}}^2$.

Wtedy $\widehat{\sigma}_{\boldsymbol{\theta}}^2 = \frac{r_{\boldsymbol{\theta}}^2}{n}$ oraz:

$$\check{d}(\boldsymbol{\theta} | \mathbf{y}_{obs}) = n \ln\left(\frac{2\pi r_{\boldsymbol{\theta}}^2}{n}\right) + 2 \ln|\mathbf{L}_{\boldsymbol{\theta}}| + n = 2 \ln|\mathbf{L}_{\boldsymbol{\theta}}| + n \left(1 + \ln\left(\frac{2\pi r_{\boldsymbol{\theta}}^2}{n}\right)\right). \quad (3.54)$$

W celu wyznaczenia $\tilde{\mathbf{u}}$ oraz $\widehat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}$ zastosować można następujący zapis:

$$\begin{aligned} \begin{bmatrix} \tilde{\mathbf{u}} \\ \widehat{\boldsymbol{\beta}}_{\boldsymbol{\theta}} \end{bmatrix} = \\ \arg \min_{\mathbf{u}, \boldsymbol{\beta}} \left(\begin{bmatrix} \mathbf{y} - \mathbf{Z}\boldsymbol{\Lambda}_{\boldsymbol{\theta}}\mathbf{u} - \mathbf{X}\boldsymbol{\beta} \\ \mathbf{u} \end{bmatrix}^T \begin{bmatrix} \mathbf{y} - \mathbf{Z}\boldsymbol{\Lambda}_{\boldsymbol{\theta}}\mathbf{u} - \mathbf{X}\boldsymbol{\beta} \\ \mathbf{u} \end{bmatrix} \right) = \\ \arg \min_{\mathbf{u}} \left(\begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{Z}\boldsymbol{\Lambda}_{\boldsymbol{\theta}} & \mathbf{X} \\ \mathbf{I}_q & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \boldsymbol{\beta} \end{bmatrix} \right)^T \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{Z}\boldsymbol{\Lambda}_{\boldsymbol{\theta}} & \mathbf{X} \\ \mathbf{I}_q & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \boldsymbol{\beta} \end{bmatrix} \right). \end{aligned} \quad (3.55)$$

Zgodnie z metodą najmniejszych kwadratów rozwiązaniem powyższego problemu są oszacowania spełniające następujący warunek:

$$\begin{bmatrix} \mathbf{Z}\boldsymbol{\Lambda}_{\boldsymbol{\theta}} & \mathbf{X} \\ \mathbf{I}_q & \mathbf{0} \end{bmatrix}^T \begin{bmatrix} \mathbf{Z}\boldsymbol{\Lambda}_{\boldsymbol{\theta}} & \mathbf{X} \\ \mathbf{I}_q & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{u}} \\ \widehat{\boldsymbol{\beta}}_{\boldsymbol{\theta}} \end{bmatrix} = \begin{bmatrix} \mathbf{Z}\boldsymbol{\Lambda}_{\boldsymbol{\theta}} & \mathbf{X} \\ \mathbf{I}_q & \mathbf{0} \end{bmatrix}^T \begin{bmatrix} \mathbf{y} \\ \mathbf{0} \end{bmatrix}, \quad (3.56)$$

co na drodze elementarnych przekształceń daje:

$$\begin{aligned} \begin{bmatrix} \boldsymbol{\Lambda}_{\boldsymbol{\theta}}^T \mathbf{Z}^T \mathbf{Z} \boldsymbol{\Lambda}_{\boldsymbol{\theta}} + \mathbf{I}_q & \boldsymbol{\Lambda}_{\boldsymbol{\theta}}^T \mathbf{Z}^T \mathbf{X} \\ \mathbf{X}^T \mathbf{Z} \boldsymbol{\Lambda}_{\boldsymbol{\theta}} & \mathbf{X}^T \mathbf{X} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{u}} \\ \widehat{\boldsymbol{\beta}}_{\boldsymbol{\theta}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Lambda}_{\boldsymbol{\theta}}^T \mathbf{Z}^T \mathbf{y} \\ \mathbf{X}^T \mathbf{y} \end{bmatrix}, \\ \begin{bmatrix} \tilde{\mathbf{u}} \\ \widehat{\boldsymbol{\beta}}_{\boldsymbol{\theta}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Lambda}_{\boldsymbol{\theta}}^T \mathbf{Z}^T \mathbf{Z} \boldsymbol{\Lambda}_{\boldsymbol{\theta}} + \mathbf{I}_q & \boldsymbol{\Lambda}_{\boldsymbol{\theta}}^T \mathbf{Z}^T \mathbf{X} \\ \mathbf{X}^T \mathbf{Z} \boldsymbol{\Lambda}_{\boldsymbol{\theta}} & \mathbf{X}^T \mathbf{X} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{\Lambda}_{\boldsymbol{\theta}}^T \mathbf{Z}^T \mathbf{y} \\ \mathbf{X}^T \mathbf{y} \end{bmatrix}. \end{aligned} \quad (3.57)$$

Podobnie jak wcześniej, rozwiązanie powyższego problemu znalezione zostaje przy pomocy metody Choleskiego. Przeprowadzana zostaje następująca dekompozycja Choleskiego (dla macierzy rzadkich):

$$\begin{bmatrix} \boldsymbol{\Lambda}_{\boldsymbol{\theta}}^T \mathbf{Z}^T \mathbf{Z} \boldsymbol{\Lambda}_{\boldsymbol{\theta}} + \mathbf{I}_q & \boldsymbol{\Lambda}_{\boldsymbol{\theta}}^T \mathbf{Z}^T \mathbf{X} \\ \mathbf{X}^T \mathbf{Z} \boldsymbol{\Lambda}_{\boldsymbol{\theta}} & \mathbf{X}^T \mathbf{X} \end{bmatrix} = \begin{bmatrix} \mathbf{L}_{\boldsymbol{\theta}} & \mathbf{0} \\ \mathbf{R}_{\mathbf{Z}\mathbf{X}} & \mathbf{R}_{\mathbf{X}} \end{bmatrix} \begin{bmatrix} \mathbf{L}_{\boldsymbol{\theta}}^T & \mathbf{R}_{\mathbf{Z}\mathbf{X}}^T \\ \mathbf{0} & \mathbf{R}_{\mathbf{X}}^T \end{bmatrix}.$$

W powyższym zapisie \mathbf{L}_θ , tak jak wcześniej, oznacza trójkątną dolną macierz wyznaczoną przy pomocy dekompozycji Choleskiego spełniającą warunek określony wzorem (3.44). Macierze \mathbf{R}_{ZX} oraz \mathbf{R}_X są odpowiednio wymiaru $p \times q$ oraz $p \times p$ spełniają poniższe warunki:

$$\mathbf{L}_\theta \mathbf{R}_{ZX}^T = \Lambda_\theta^T \mathbf{Z}^T \mathbf{X} \quad (3.58)$$

oraz

$$\mathbf{R}_X \mathbf{R}_X^T = \mathbf{X}^T \mathbf{X} - \mathbf{R}_{ZX} \mathbf{R}_{ZX}^T. \quad (3.59)$$

Oszacowania parametrów równania (3.57) wyznaczone zostaną więc przy pomocy metody Choleskiego na podstawie równania:

$$\begin{bmatrix} \mathbf{L}_\theta & \mathbf{0} \\ \mathbf{R}_{ZX} & \mathbf{R}_X \end{bmatrix} \begin{bmatrix} \mathbf{L}_\theta^T & \mathbf{R}_{ZX}^T \\ \mathbf{0} & \mathbf{R}_X^T \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{u}} \\ \hat{\boldsymbol{\beta}}_\theta \end{bmatrix} = \begin{bmatrix} \Lambda_\theta^T \mathbf{Z}^T \mathbf{y} \\ \mathbf{X}^T \mathbf{y} \end{bmatrix}. \quad (3.60)$$

Powyższe macierze uzyskane z dekompozycji Choleskiego zostały zapisane w formie blokowej, ponieważ jeden z wyznaczonych bloków posłuży w dalszych rozważaniach do wyznaczenia estymatorów resztowej metody największej wiarygodności.

ESTYMATORY RESZTOWEJ METODY NAJWIĘKSZEJ WIAROGODNOŚCI

Funkcję resztowej największej wiarygodności zapisać można:

$$L_R(\boldsymbol{\theta}, \sigma | \mathbf{y}_{obs}) = \int_{RP} L(\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma | \mathbf{y}_{obs}) d\boldsymbol{\beta}. \quad (3.61)$$

Resztowa dewiacja wyznaczana jest zatem jako:

$$d_R(\boldsymbol{\theta}, \sigma | \mathbf{y}_{obs}) = -2 \ln L_R(\boldsymbol{\theta}, \sigma | \mathbf{y}_{obs}) = -2 \ln \int_{RP} L(\boldsymbol{\theta}, \boldsymbol{\beta}, \sigma | \mathbf{y}_{obs}) d\boldsymbol{\beta}. \quad (3.62)$$

Sformułować można następującą równość:

$$r_{\boldsymbol{\theta}, \boldsymbol{\beta}}^2 = r_{\boldsymbol{\theta}}^2 + \left(\mathbf{R}_X^T (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_\theta) \right)^T \left(\mathbf{R}_X^T (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_\theta) \right). \quad (3.63)$$

Podstawiając równanie (3.51) do równania (3.61), resztową wiarygodność można przedstawić jako:

$$L_R(\boldsymbol{\theta}, \sigma | \mathbf{y}_{obs}) = \int_{RP} (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{r_{\boldsymbol{\theta}, \boldsymbol{\beta}}^2}{2\sigma^2}\right) |\mathbf{L}_\theta|^{-1} d\boldsymbol{\beta}. \quad (3.64)$$

Do powyższego równania podstawiając (3.63) można wyznaczyć resztową wiarygodność:

$$\begin{aligned}
L_R(\boldsymbol{\theta}, \sigma | \mathbf{y}_{obs}) &= \int_{R^p} (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{r_{\hat{\boldsymbol{\theta}}}^2 + (\mathbf{R}_X^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}))^T (\mathbf{R}_X^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}))}{2\sigma^2}\right) |\mathbf{L}_{\boldsymbol{\theta}}|^{-1} d\boldsymbol{\beta} = \\
& \exp\left(-\frac{r_{\hat{\boldsymbol{\theta}}}^2}{2\sigma^2}\right) (2\pi\sigma^2)^{-\frac{n-p}{2}} |\mathbf{L}_{\boldsymbol{\theta}}|^{-1} \int_{R^p} (2\pi\sigma^2)^{-\frac{p}{2}} \exp\left(-\frac{(\mathbf{R}_X^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}))^T (\mathbf{R}_X^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}))}{2\sigma^2}\right) \frac{|\mathbf{R}_X| d\boldsymbol{\beta}}{|\mathbf{R}_X| \sigma^p} = \\
& \left| \frac{z = \frac{\mathbf{R}_X^T(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}})}{\sigma}}{\left| \frac{dz}{d\boldsymbol{\beta}} \right| = \frac{|\mathbf{R}_X|}{\sigma^p}} \right| = \exp\left(-\frac{r_{\hat{\boldsymbol{\theta}}}^2}{2\sigma^2}\right) (2\pi\sigma^2)^{-\frac{n-p}{2}} |\mathbf{L}_{\boldsymbol{\theta}}|^{-1} |\mathbf{R}_X|^{-1} \int_{R^p} (2\pi\sigma^2)^{-\frac{p}{2}} \exp\left(-\frac{z^T z}{2}\right) dz = \\
& \exp\left(-\frac{r_{\hat{\boldsymbol{\theta}}}^2}{2\sigma^2}\right) (2\pi\sigma^2)^{-\frac{n-p}{2}} |\mathbf{L}_{\boldsymbol{\theta}}|^{-1} |\mathbf{R}_X|^{-1}. \tag{3.65}
\end{aligned}$$

Zatem resztowa dewiacja wynosi:

$$\begin{aligned}
d_R(\boldsymbol{\theta}, \sigma | \mathbf{y}_{obs}) &= -2 \ln\left(\exp\left(-\frac{r_{\hat{\boldsymbol{\theta}}}^2}{2\sigma^2}\right) (2\pi\sigma^2)^{-\frac{n-p}{2}} |\mathbf{L}_{\boldsymbol{\theta}}|^{-1} |\mathbf{R}_X|^{-1}\right) = \\
& (n-p) \ln(2\pi\sigma^2) + 2 \ln(|\mathbf{L}_{\boldsymbol{\theta}}| |\mathbf{R}_X|) + \frac{r_{\hat{\boldsymbol{\theta}}}^2}{\sigma^2}. \tag{3.66}
\end{aligned}$$

Estymatorem resztowej największej wiarygodności parametru σ jest $\widehat{\sigma^2}_{\boldsymbol{\theta}} = \frac{r_{\hat{\boldsymbol{\theta}}}^2}{n-p}$.

Zatem:

$$\begin{aligned}
d_R(\boldsymbol{\theta} | \mathbf{y}_{obs}) &= (n-p) \ln\left(\frac{2\pi r_{\hat{\boldsymbol{\theta}}}^2}{n-p}\right) + 2 \ln(|\mathbf{L}_{\boldsymbol{\theta}}| |\mathbf{R}_X|) + n-p = \\
& = 2 \ln(|\mathbf{L}_{\boldsymbol{\theta}}| |\mathbf{R}_X|) + (n-p) \left(1 + \ln\left(\frac{2\pi r_{\hat{\boldsymbol{\theta}}}^2}{n-p}\right)\right). \tag{3.67}
\end{aligned}$$

Oszacowanie parametru $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}_R$ wyznaczone jest jako wartość minimalizująca resztową dewiację, zaś oszacowaniem σ jest $\widehat{\sigma^2}_{\hat{\boldsymbol{\theta}}_R}$. Jako estymator $\boldsymbol{\beta}$ przyjmuje się $\hat{\boldsymbol{\beta}}_{\hat{\boldsymbol{\theta}}_R}$, gdzie $\hat{\boldsymbol{\beta}}_{\boldsymbol{\theta}}$ jest estymatorem wyznaczonym metodą największej wiarygodności dla danego $\boldsymbol{\theta}$.

3.5. Kryteria oceny dopasowania modelu

Jak wspomiano w podrozdziałach 3.2 oraz 3.3, podczas wprowadzania do modelu kolejnych parametrów należy każdorazowo weryfikować zasadność takiego postępowania, aby nie komplikować modelu jeżeli nie poprawia to w istotnym stopniu jakości jego dopasowania. Bates (2010) nazywa to zasadą oszczędności i powołuje się przy tym na znane stwierdzenie Alberta Einsteina: „Wszystko powinno być wykonane tak prosto, jak to możliwe, ale nie prościej”. Nadmierne rozszerzanie zbioru szacowanych parametrów nie tylko utrudnia interpretację wyników oraz wydłuża czas obliczeń (nawet do wielu dni), ale może uniemożliwić oszacowanie tych parametrów. Na przykład

uwzględnienie korelacji pomiędzy efektami losowymi w sytuacji, gdy jest ona zbliżona do zera może powodować brak zbieżności algorytmu, a więc brak jakichkolwiek oszacowań. Z drugiej strony jest oczywiste, że należy uwzględnić te zależności, które umożliwiają poprawę precyzji późniejszych szacunków.

W związku z powyższym, niezmiernie ważną kwestią jest ocena poprawy dopasowania modelu po każdym rozszerzeniu. W tym celu posłużyć można się omówioną w poprzednim podrozdziale funkcją wiarygodności i porównać wiarygodności obu modeli (lub resztowe wiarygodności, jeżeli parametry oszacowano metodą REML). Im bardziej zwiększy się wiarygodność modelu (lub jej logarytm naturalny, podawany przez większość programów), tym większa poprawa. W celu oceny, czy określona poprawa modelu jest istotna statystycznie zastosować można test ilorazu wiarygodności.

TEST ILORAZU WIAROGODNOŚCI χ^2

Przy pomocy testu χ^2 porównuje się jakość dopasowania dwóch różnych modeli: modelu A oraz jego rozszerzenia B. O modelach tych zakładamy, że model B, jako rozszerzenie A, jest lepiej dopasowany do danych empirycznych. W przypadku szacowania parametrów metodą największej wiarygodności (ML), test ten przeprowadzony może zostać po dowolnym rozszerzeniu modelu. Jednak w przypadku resztowej metody największej wiarygodności (REML) można go stosować tylko wówczas, jeżeli dodano parametr bądź parametry w losowej części modelu. W przeciwnym wypadku, czyli kiedy dokonano zmian w części stałej, zmienia się wyrażenie będące częścią objaśnianą w tej metodzie, zatem rozszerzonego modelu nie można traktować jako rozwinięcia modelu bazowego.

Weryfikacja tej hipotezy sprowadza się do sprawdzenia, czy różnica w jakości obu modeli jest statystycznie istotna (por. Lin, 1997; Goldstein, 2003, Krzyśko, 2009). Tak więc hipoteza zerowa zakłada, że model B nie jest istotnie lepszy od modelu A. Natomiast hipoteza alternatywna głosi, że wiarygodność modelu B jest statystycznie istotnie wyższa niż wiarygodność modelu A (por. wzór (3.68)). Odrzucenie hipotezy zerowej świadczy o tym, że warto rozszerzyć model A do modelu B i przyjąć model B jako obowiązujący w dalszych analizach. W przypadku braku podstaw do odrzucenia hipotezy zerowej nie należy wprowadzać rozważanej zmiany do modelu A i pozostawić go obowiązującym w dalszych rozważaniach.

Układ hipotez:

$$\begin{cases} H_0: \sigma_A = \sigma_B \\ H_1: \sigma_A > \sigma_B \end{cases}, \quad (3.68)$$

gdzie:

σ_A – odchylenie standardowe składnika losowego modelu A,

σ_B – odchylenie standardowe składnika losowego modelu B.

W celu weryfikacji hipotezy zerowej stosuje się statystykę testową o następującej postaci:

$$\chi^2 = 2 * \ln L_B - 2 * \ln L_A, \quad (3.69)$$

gdzie:

L_A - supremum funkcji wiarygodności w modelu A,

L_B - supremum funkcji wiarygodności w modelu B.

Można wykazać, że tak wyznaczona statystyka, przy prawdziwości hipotezy zerowej ma rozkład χ^2 , z liczbą stopni swobody obliczaną jako różnica pomiędzy liczbą parametrów szacowanych w modelu B a liczbą parametrów szacowaną w modelu A:

$$\chi^2|_{H_0} \sim \chi^2(p_B - p_A), \quad (3.70)$$

gdzie:

p_A - liczba parametrów szacowanych w modelu A,

p_B - liczba parametrów szacowanych w modelu B.

Należy pamiętać, że na liczbę szacowanych elementów składa się suma długości wektora współczynników części stałej modelu β , długość wektora θ , w skład którego wchodzi wszystkie wariancje i kowariancje efektów losowych wchodzące w skład macierzy kowariancji efektów losowych oraz jeden parametr odpowiadający wariancji składnika resztowego, σ^2 .

Obszar krytyczny, służący do weryfikacji powyższej hipotezy zerowej, jest więc następującej postaci:

$$B = \{ \underline{Y}: \chi^2 = 2 * \ln L_B - 2 * \ln L_A > \chi^2(1 - \alpha; p_B - p_A) \}, \quad (3.71)$$

gdzie:

α - przyjęty poziom istotności,

$\chi^2(a; b)$ - wartość kwantyla rozkładu χ^2 z b stopniami swobody z prawdopodobieństwa a.

W sytuacji, gdy statystyka testowa przyjmuje wartość należącą do obszaru krytycznego, hipotezę zerową odrzucamy na korzyść hipotezy alternatywnej. Oznacza to, że model B

charakteryzuje się istotnie lepszym od modelu A dopasowaniem do danych empirycznych.

KRYTERIA INFORMACJNE

Oprócz testu ilorazu wiarygodności, oceny poprawy modelu dokonać można w oparciu o liczne kryteria informacyjne. Są to charakterystyki modelu obliczane na podstawie jego wiarygodności oraz liczby szacowanych parametrów. Można je stosować zarówno w przypadku szacowania parametrów metodą największej wiarygodności jak i resztowej metody największej wiarygodności. Poniżej przedstawiono dwa najpopularniejsze kryteria informacyjne.

Kryterium informacyjne Akaike'a

Pierwszym z kryteriów opartych na funkcji wiarygodności jest kryterium informacyjne Akaike'a (AIC) (por. wzór (3.72)).

$$AIC = -2 * \ln(L) + 2 * p , \quad (3.72)$$

gdzie:

L – maksymalna wartość funkcji wiarygodności modelu,

p - liczba szacowanych w modelu parametrów.

Wysoka wartość funkcji wiarygodności informuje o dobrym dopasowaniu modelu. Ponieważ nadmierny wzrost liczby szacowanych parametrów uznawany jest za niekorzystny, kryterium AIC uwzględnia fakt, iż zbyt ich duża liczba obniża wartość modelu. Model z minimalną wartością AIC jest uznawany, według tego kryterium, za najlepiej dopasowany do danych (por. Sakamoto, Ishiguro, Kitagawa, 1986).

Bayesowskie kryterium informacyjne

Kolejnym kryterium pozwalającym ocenić jakość dopasowania modelu jest Bayesowskie kryterium informacyjne (BIC). Wartość jaką przyjmuje współczynnik BIC wyznacza się zgodnie ze wzorem (3.73) (por. Schwarz, 1978).

$$BIC = -2 * \ln(L) + \ln(n) * p , \quad (3.73)$$

gdzie:

L – maksymalna wartość funkcji wiarygodności modelu,

n - liczba obserwacji w próbie (z pierwszego poziomu),

p - liczba szacowanych w modelu parametrów.

Porównując kryterium BIC z AIC, można stwierdzić, że podobnie jak poprzednio uwzględnia ono dodatni wpływ wysokiej wartości funkcji wiarygodności modelu oraz ujemne oddziaływanie zbyt dużej liczby szacowanych parametrów. Jednak znaczenie liczby szacowanych parametrów uzależniono od liczebności próby. Uznano je za istotniejsze, gdy próba jest liczna, a za mniej istotne, gdy próba jest mała, co związane jest z czasem obliczeń. Ponieważ $\ln(7,4) \approx 2$, dla sytuacji gdy liczebność próby jest większa bądź równa 8, kryterium BIC bardziej surowo „karze” model za zwiększoną liczbę szacowanych parametrów niż AIC. Oczywiście, tak jak w przypadku kryterium AIC, za najlepszy uznawany jest model o najmniejszym współczynniku BIC. Kryterium BIC jest znane również jako SBC (Kryterium Bayesowskie Schwarza).

Rozdział 4. Oszacowanie macierzy migracji dla roku 2011

Badanie przepływów ludności związanych z zatrudnieniem jest, jak do tej pory badaniem jednorazowym, którego wyniki aktualne są dla roku 2006²⁵. Informacje umożliwiające konstrukcję macierzy migracji codziennych, związanych z zatrudnieniem, w warunkach polskich, można pozyskać tylko z tego badania. W przeszłości dostępne były informacje z tzw. spisów kadrowych, które umożliwiały szacunek podobnej macierzy migracji. Ostatni raz spis taki przeprowadzony został w roku 1983. Zatem delimitację lokalnych rynków pracy w Polsce, bezpośrednio na podstawie dostępnych źródeł, przeprowadzić można aktualnie jedynie dla 2006 roku.

Celem rozprawy jest delimitacja lokalnych rynków pracy w Polsce. Zostanie on zrealizowany na podstawie dostępnych danych, tj. dla 2006 roku. Jednak podjęta zostanie również próba przeprowadzenia szacunków dla lat późniejszych, w szczególności dla roku 2011. Wymaga ona oszacowania danych odpowiadających zakresem badaniu przepływów ludności związanych z zatrudnieniem. Przeprowadzone studia literaturowe wskazały, że mobilność ludności związana z zatrudnieniem kształtowana jest przez szereg czynników charakteryzujących zarówno predyspozycje indywidualnych osób, jak również gospodarkę regionu. Uwzględniano determinanty wynikające z podaży pracy i popytu na nią w ujęciu terytorialnym, infrastrukturę komunikacyjną, łatwość dojazdu. Podział przestrzenny kraju, utożsamiany z podziałem administracyjnym, charakteryzuje się strukturą wielopoziomową. Uzasadnionym wydaje się więc prowadzenie oszacowań przy pomocy modelowania uwzględniającego wielopoziomą strukturę rynku pracy. Model wielopoziomowy, w którym objaśnione zostaną wartości odpowiednich charakterystyk dojazdów do pracy, skonstruowany zostanie dla roku 2006. Przyjmując założenie, że w roku 2011 kształt zależności pomiędzy dojazdami do pracy oraz zmiennymi objaśniającymi zastosowanymi w modelu nie uległ istotnym zmianom, oszacowane zostaną charakterystyki dojazdów do pracy dla 2011 roku.

Jak opisano w podrozdziale 1.3, dane z badania przepływów ludności związanych z zatrudnieniem dostarczają bezpośrednio informacji tylko o pracownikach najemnych

²⁵ Chodzi o badanie przeprowadzone z wykorzystaniem danych rejestru POLTAX (por. Dojazdy do pracy w Polsce, 2010). W ramach NSP 2011 również pytano o dojazdy do pracy, jednak dane te nie zostały jeszcze opublikowane, ponadto nie zawierają informacji o kierunkach dojazdów. Ponadto trwają prace nad przygotowaniem rozszerzenia merytorycznego i kontynuacją badania z wykorzystaniem rejestrów Ministerstwa Finansów.

zatrudnionych poza gminą zamieszkania. Informacja o ludności pracującej w gminie zamieszkania, znajdująca się na diagonalu macierzy migracji, musi być oszacowana oddzielnie. W tym celu wykorzystać można dane o liczbie osób pracujących w gminach dostępne ze sprawozdawczości statystycznej. Aby rozważania ograniczyć tylko do osób jednocześnie mieszkających i pracujących na terenie gminy, wielkość wynikającą ze sprawozdawczości należy umniejszyć o liczbę przyjeżdżających do pracy zaczerpniętą z badania przepływów ludności związanych z zatrudnieniem (mieszkających poza granicami badanej gminy).

W związku z koniecznością określenia dwóch różnych wielkości: pracujących w gminie zamieszkania i przepływów pomiędzy parami gmin, również estymacja macierzy migracji podzielona została na dwie oddzielne procedury. W pierwszej kolejności, w podrozdziale 4.1 oszacowane zostaną dane odpowiadające zakresowi diagonalu macierzy migracji, następnie,

w podrozdziale 4.2 przeprowadzona zostanie estymacja wartości odpowiadających pozostałej części macierzy migracji.

4.1. Oszacowanie wartości na diagonalu

OGÓLNE PODEJŚCIE ORAZ STRUKTURA POPULACJI

W podrozdziale tym oszacowana zostanie liczba ludności pracującej w gminie zamieszkania (z podziałem na część miejską i wiejską). Wartości te szacowane są jako różnica pomiędzy liczbą osób pracujących w danej gminie i liczbą ludności przyjeżdżającej do pracy. Ponieważ liczba ludności pracującej w przekroju gmin jest znana dla roku 2011, wystarczy oszacować liczbę ludności przyjeżdżającej do pracy w danej gminie. W tym celu podjęta zostanie konstrukcja modelu dla roku 2006, w którym objaśnione zostaną znane dla tego roku liczby osób przyjeżdżających do pracy. Następnie, w myśl sformułowanego założenia o stałości w czasie, zależności pomiędzy przyjazdami do pracy oraz zmiennymi objaśniającymi, skonstruowany model zostanie wykorzystany dla oszacowania liczby przyjeżdżających do pracy w 2011 r. Ponieważ jednostki bazowe utożsamiać można z jednostkami podziału administracyjnego kraju, ich struktura jest wielopoziomowa. W związku z tym spełniony jest warunek konieczny konstrukcji modelu trypoziomowego, w którym jednostki poziomu pierwszego

określone są jako gminy z podziałem na część miejską i wiejską, jednostki drugiego poziomu jako powiaty, zaś jednostki poziomu trzeciego jako województwa.

ZMIENNA OBJAŚNIANA

Celem postępowania jest oszacowanie liczby przyjeżdżających do pracy. Ze względu na duże zróżnicowanie wielkości gmin (mierzonej liczbą ludności), zapewnienie porównywalności pod względem stopnia rozwoju gospodarczego regionów, aktywności ekonomicznej ludności czy intensywności dojazdów do pracy, wymaga zastosowania wielkości względnych. Można je traktować jako pewnego rodzaju mierniki sytuacji społeczno-gospodarczej, które różnią się znacznie pomiędzy jednostkami podziału administracyjnego. Zatem w konstruowanym modelu zmienną objaśnianą (Y_i) zdefiniowano jako udział przyjeżdżających do pracy wśród wszystkich pracujących w badanej gminie (dalej zmienna ta nazywana jest natężeniem przyjazdów). Jak podają w swojej pracy P.Korcelli, A.Potrykowska i D.Bodzak (1981), wskaźnik ten jest miarą stopnia przestrzennej koncentracji działalności gospodarczej i społecznej, rozpatrywanej w odniesieniu do podaży czynników pracy. Autorzy powyższego opracowania wskazują między innymi, że poziom natężenia przyjazdów różnił się istotnie pomiędzy ówczesnymi województwami (por. Korcelli, Potrykowska, Bodzak, 1981). Ze względu na postać wykorzystywanej w dalszej procedurze macierzy migracji, po zakończeniu oszacowań otrzymane wartości przemnożone zostały przez wartości mianownika, odpowiadające liczbie pracujących w danej gminie, które znane są ze sprawozdawczości, również dla roku 2011.

ZMIENNE OBJAŚNIAJĄCE

W dalszej kolejności podjęto próbę doboru zmiennych objaśniających do modelu. Powinien być to zestaw charakterystyk w jak największym stopniu opisujących sytuację na rynku pracy. W literaturze podaje się, że w jego skład powinny wchodzić zmienne określające zarówno popyt jak i podaż pracy (por. Gołata, 2004). Na podstawie literatury, w rozdziale pierwszym wskazano szereg czynników, którym w tym miejscu przypisana zostanie rola potencjalnych zmiennych objaśniających sytuację na rynku pracy. Zostaną one wykorzystane podczas konstrukcji modelu szacowania liczby przyjeżdżających do pracy. Przedstawiając zestaw 39 potencjalnych zmiennych objaśniających, uwzględniono strukturę wielopoziomową jednostek podziału administracyjnego, tj. przekrój gmin, powiatów i województw (por. tab. 4.1, 4.2, 4.3). Odzwierciedla to sposób zapisu

zmiennych, w których symbolach użyto dwóch subskryptów: pierwszy indeks określa poziom podziału terytorialnego (gmina=1, powiat=2, województwo=3), natomiast drugi indeks wskazuje kolejny numer zmiennej. Wyodrębniono pięć grup czynników, z których pierwsze trzy opisują rynek pracy z uwzględnieniem strony podażowej (pierwsza grupa), popytowej (druga grupa) i cen (trzecia grupa). Dwie pozostałe grupy czynników dotyczą odległości, na której podejmowane są dojazdy oraz położenia gmin.

Jako charakterystyki strony podażowej, zaproponowano stosunek osób pracujących do ludności w wieku produkcyjnym ($X_{1,10}$) i liczbę osób pracujących na kilometr kwadratowy ($X_{1,17}$) w przekroju gmin, stosunek osób pracujących do ludności w wieku produkcyjnym ($X_{2,6}$) w przekroju powiatów oraz: stosunek osób pracujących do ludności w wieku produkcyjnym ($X_{3,2}$), udział absolwentów publicznych uczelni wyższych wśród ludności w wieku produkcyjnym ($X_{3,7}$) oraz udział absolwentów uczelni wyższych łącznie wśród ludności w wieku produkcyjnym ($X_{3,10}$) w przekroju województw. Ponadto na poziomie gmin rozważono zmienną określoną jako liczba osób w wieku produkcyjnym na kilometr kwadratowy ($X_{1,18}$).

Z kolei strona popytowa opisana została przy pomocy udziału bezrobotnych wśród ludności w wieku produkcyjnym określonego dla gmin²⁶, powiatów i województw (odpowiednio $X_{1,3}$, $X_{2,1}$, $X_{3,1}$). Charakterystyka strony popytowej ujęta została także przy pomocy natężenia różnych rodzajów przedsiębiorstw w gminach (por. tab. 4.1), udziału osób pracujących w warunkach zagrożenia wśród wszystkich pracujących ($X_{2,9}$) w przekroju powiatów oraz nakładów na działalność innowacyjną w przedsiębiorstwach przemysłowych i łącznie w tys. zł na osobę w wieku produkcyjnym (odpowiednio $X_{3,4}$, $X_{3,9}$) w przekroju województw.

Trzecia grupa potencjalnych zmiennych objaśniających opisuje cenę pracy. Została ona określona przez średnie wynagrodzenie²⁷ obliczone w przekroju powiatów oraz województw (odpowiednio $X_{2,10}$, $X_{3,11}$).

²⁶ Należy zaznaczyć, że ta charakterystyka bezrobocia wyznaczona została dla gmin bez podziału na część miejską i wiejską. Ponieważ w literaturze dotyczącej modelowania wielopoziomowego nie zaleca się tworzenia poziomu, którego jednostki zawierają po dwie lub mniej jednostek poziomu niższego, wartości omawianej charakterystyki bezrobocia, przypisano zarówno części miejskiej i wiejskiej w przypadku gmin charakteryzujących się takim podziałem.

²⁷ Stosunek sumy wynagrodzeń osobowych brutto, honorariów wypłaconych niektórym grupom pracowników za prace wynikające z umowy o pracę, wypłat z tytułu udziału w zysku lub w nadwyżce

Tabela 4.1. Współczynniki korelacji liniowej Pearsona między udziałem przyjeżdżających do pracy a potencjalnymi zmiennymi objaśniającymi, poziom gmin, 2006

Symbol zmiennej ¹	Nazwa zmiennej	r_{xy}
X _{1,1}	Odległość od najbliższego z ośrodków centralnych	-0,2722
X _{1,2}	Liczba przedsiębiorstw z sekcji D zarejestrowanych w REGON na osobę w wieku produkcyjnym	0,2245
X _{1,3}	Udział bezrobotnych wśród ludności w wieku produkcyjnym	-0,1971
X _{1,4}	Liczba osób fizycznych prowadzących działalność gospodarczą na osobę w wieku produkcyjnym	0,1758
X _{1,5}	Liczba przedsiębiorstw z sektora prywatnego na osobę w wieku produkcyjnym	0,1731
X _{1,6}	Liczba przedsiębiorstw zarejestrowanych w REGON na osobę w wieku produkcyjnym	0,1636
X _{1,7}	Liczba przedsiębiorstw z sekcji I zarejestrowanych w REGON na osobę w wieku produkcyjnym	0,1634
X _{1,8}	Liczba przedsiębiorstw z sekcji G zarejestrowanych w REGON na osobę w wieku produkcyjnym	0,1566
X _{1,9}	Liczba przedsiębiorstw z sekcji F zarejestrowanych w REGON na osobę w wieku produkcyjnym	0,1546
X _{1,10}	Stosunek liczby osób pracujących do ludności w wieku produkcyjnym	0,1501
X _{1,11}	Liczba spółek handlowych z sektora prywatnego na osobę w wieku produkcyjnym	0,1434
X _{1,12}	Odległość od najbliższej stolicy województwa	-0,1108
X _{1,13}	Liczba przedsiębiorstw z sekcji O zarejestrowanych w REGON na osobę w wieku produkcyjnym	0,1035
X _{1,14}	Liczba przedsiębiorstw z sekcji K zarejestrowanych w REGON na osobę w wieku produkcyjnym	0,0656
X _{1,15}	Liczba spółek handlowych z udziałem kapitału zagranicznego z sektora prywatnego na osobę w wieku produkcyjnym	0,0653
X _{1,16}	Liczba spółek handlowych z udziałem kapitału zagranicznego z sektora publicznego na osobę w wieku produkcyjnym	0,0608
X _{1,17}	Liczba osób pracujących na kilometr kwadratowy	0,0112
X _{1,18}	Liczba osób w wieku produkcyjnym na kilometr kwadratowy	-0,0105

Uwaga:

¹ pierwszy indeks w symbolu zmiennej określa poziom podziału terytorialnego (gmina=1), natomiast drugi indeks wskazuje kolejny numer zmiennej

Źródło: Opracowanie własne na podstawie danych GUS

Atrakcyjność lokalizacji miejsca pracy została scharakteryzowana poprzez odległość²⁸ od dużych miast. W pierwszej kolejności dokonano wyboru ośrodków centralnych. W tym celu uwzględniono wszystkie miasta na prawach powiatu i uszeregowano je

bilansowej w spółdzielniach oraz dodatkowych wynagrodzeń rocznych dla pracowników jednostek sfery budżetowej do przeciętnej liczby zatrudnionych w danym okresie; po wyeliminowaniu osób wykonujących pracę nakładczą oraz zatrudnionych za granicą (por. http://www.stat.gov.pl/gus/definicje_PLK_HTML.htm?id=POJ-1234.htm).

²⁸ Uwzględniono odległość w kilometrach w linii prostej pomiędzy centroidami gmin, obliczoną przy pomocy programu R (por. Kopczevska, 2006).

w kolejności malejącej ze względu na liczbę osób zatrudnionych²⁹. W pierwszym kroku zbiór ośrodków centralnych zdefiniowano jako miasto o największej liczbie zatrudnionych. Następnie zbiór ten powiększono o kolejne miasta na prawach powiatu, uszeregowane także według liczby zatrudnionych. Dla tak zdefiniowanych w kolejnych etapach zbiorów wielkich miast badano korelację pomiędzy odległością a intensywnością wyjazdów do pracy. Następnie weryfikowano, dla którego ze zdefiniowanych zbiorów wielkich miast określonych w poszczególnych krokach, zależność pomiędzy odległością a intensywnością wyjazdów do pracy była najsilniejsza. Zgodnie z teorią Thunena głoszącą, że ośrodki centralne stymulują rozwój obszarów je otaczających, założono, że odległość od nich powinna wpływać na nasilenie wyjazdów do pracy. Ostatecznie wybrano zbiór wielkich miast, dla którego zależność korelacyjna była najsilniejsza. Według powyższego kryterium rolę ośrodków centralnych pełniły: Warszawa, Kraków, Poznań, Wrocław, Łódź, Katowice, Gdańsk, Częstochowa, Rzeszów, Gdynia, Bielsko-Biała, Tarnów, Tychy, Dąbrowa Górnicza, Zielona Góra, Jastrzębie-Zdrój, Nowy Sącz, Krosno, Piotrków Trybunalski, Jaworzno oraz Leszno (por. Gruchociak, 2012c). Jako pierwszą charakterystykę atrakcyjności lokalizacji miejsca pracy przyjęto więc małą odległość centrum gminy od najbliższego z wymienionych miast ($X_{1,1}$). Kolejną zmienną z tej kategorii zdefiniowano jako odległość centrum każdej gminy od najbliższej ze stolic województw ($X_{1,12}$). Zaproponowano także sześć zmiennych dotyczących sieci dróg gminnych i powiatowych w powiatach (por. tab. 4.2) oraz długości linii kolejowych ($X_{3,3}$). Ponadto wskazano trzy charakterystyki miejskiej komunikacji publicznej ($X_{3,5}$, $X_{3,6}$, $X_{3,8}$) w przekroju województw.

Z przedstawionego zespołu potencjalnych zmiennych objaśniających, zmienne wprowadzone do modelu wybrano stosując metodę regresji krokowej w przód, podczas kolejnych etapów jego konstrukcji. W celu przedstawienia relacji wybranych do modelu zmiennych objaśniających ze zmienną szacowaną, obliczono współczynniki korelacji liniowej Pearsona. Należy podkreślić, że współczynnik korelacji liniowej Pearsona przeznaczony jest do badania zależności pomiędzy zmiennymi określonymi dla tej samej populacji, tzn. gmin, powiatów, czy województw. Ujęcie relacji pomiędzy różnie zdefiniowanymi jednostkami powoduje tzw. paradoks Simpsona polegający

²⁹ Wybrano liczbę osób zatrudnionych, ponieważ uznano, że charakteryzuje ona wielkość miasta, oraz jednocześnie uwzględnia w pewnym zakresie stopień rozwoju. Sama wielkość miasta mierzona liczbą mieszkańców nie świadczy bowiem o tym, że stymuluje ono rozwój otaczających je regionów.

na formułowania wniosków, jeśli dane z niejednorodnych populacji połączy się w całość i analizuje jakby pochodziły z jednej jednorodnej populacji (por. Klimanek, 2003). W przypadku obliczania współczynnika korelacji dla zmiennej objaśnianej z poziomu gmin oraz zmiennych objaśniających z poziomu powiatów czy województw konieczne było zintegrowanie tych zbiorowości. Zmienne z poziomu powiatów i województw przeistoczono sztucznie w zmienne określone na poziomie gmin przypisując każdej z nich wartości cech, odpowiadające powiatowi czy województwu, do którego dana gmina należała. W wyniku takiego postępowania należało się spodziewać, że wartości bezwzględne tak obliczonych współczynników korelacji liniowej Pearsona będą charakteryzowały się mniejszymi wartościami, niż analogiczne współczynniki dla par zmiennych określonych na tym samym poziomie gmin, co wynika chociażby z ich struktury zawierającej wiele duplikatów. W związku z tym wartości współczynnika korelacji liniowej Pearsona ze zmienną objaśnianą rozważono oddzielnie dla charakterystyk każdego z poziomów (por. tab. 4.1, 4.2, 4.3).

Tabela 4.2. Współczynniki korelacji liniowej Pearsona między udziałem przyjeżdżających do pracy a potencjalnymi zmiennymi objaśniającymi, poziom powiatów, 2006

Symbol zmiennej ¹	Nazwa zmiennej	r_{xy}
X _{2,1}	Udział bezrobotnych wśród ludności w wieku produkcyjnym	-0,1686
X _{2,2}	Drogi gminne o nawierzchni twardej w kilometrach na km ²	0,1453
X _{2,3}	Drogi gminne łącznie w kilometrach na km ²	0,1356
X _{2,4}	Drogi gminne o nawierzchni twardej ulepszonej w kilometrach na km ²	0,1231
X _{2,5}	Drogi powiatowe o nawierzchni twardej ulepszonej w kilometrach na km ²	0,1110
X _{2,6}	Stosunek liczby osób pracujących do ludności w wieku produkcyjnym	0,1050
X _{2,7}	Drogi powiatowe łącznie w kilometrach na km ²	0,0965
X _{2,8}	Drogi powiatowe o nawierzchni twardej w kilometrach na km ²	0,0963
X _{2,9}	Udział osób pracujących w warunkach zagrożenia wśród wszystkich pracujących	0,0903
X _{2,10}	Średnie wynagrodzenie	0,0559

Uwaga:

¹ pierwszy indeks w symbolu zmiennej określa poziom podziału terytorialnego (powiat=2), natomiast drugi indeks wskazuje kolejny numer zmiennej

Źródło: Opracowanie własne na podstawie danych GUS

Wśród zmiennych określonych dla poziomu gmin najsilniej skorelowana z natężeniem przyjazdów do pracy była odległość od najbliższego ośrodka centralnego. Nieznacznie mniejszą korelację zmiennej objaśnianej stwierdzono z natężeniem przedsiębiorstw z sekcji D oraz udziałem osób bezrobotnych (por. tab. 4.1). Wśród charakterystyk określonych dla powiatów, najsilniejszą zależność z natężeniem przyjazdów do pracy

do gmin stwierdzono w przypadku udziału osób bezrobotnych. Zależność o nieco mniejszej sile stwierdzono w przypadku charakterystyk dróg gminnych (por. tab. 4.2). Udział osób bezrobotnych w przekroju województw również jest stosunkowo mocno skorelowany z natężeniem przyjazdów do pracy (por. tab. 4.3).

Tabela 4.3. Współczynniki korelacji liniowej Pearsona między udziałem przyjeżdżających do pracy a potencjalnymi zmiennymi objaśniającymi, poziom województw, 2006

Symbol zmiennej ¹	Nazwa zmiennej	r_{xy}
X _{3,1}	Udział bezrobotnych wśród ludności w wieku produkcyjnym	-0,1604
X _{3,2}	Stosunek liczby osób pracujących do ludności w wieku produkcyjnym	0,1466
X _{3,3}	Linie kolejowe normalnotorowe na 100 km ²	0,1140
X _{3,4}	Nakłady na działalność innowacyjną w przedsiębiorstwach przemysłowych w tys. zł na osobę w wieku produkcyjnym	0,1055
X _{3,5}	Miejsca w wozach na 1000 ludności miast obsługiwanych przez komunikację	0,0750
X _{3,6}	Ludność w miastach obsługiwanych przez zakłady komunikacji miejskiej w % ogółu ludności miast	0,0735
X _{3,7}	Udział absolwentów publicznych uczelni wyższych wśród ludności w wieku produkcyjnym	0,0657
X _{3,8}	Miejsca w wozach komunikacji miejskiej naziemnej na 1 osobę w wieku produkcyjnym	0,0606
X _{3,9}	Nakłady na działalność innowacyjną w tys. zł na osobę w wieku produkcyjnym.	0,0387
X _{3,10}	Udział absolwentów uczelni wyższych wśród ludności w wieku produkcyjnym	0,0316
X _{3,11}	Średnie wynagrodzenie	0,0299

Uwaga:

¹ pierwszy indeks w symbolu zmiennej określa poziom podziału terytorialnego (województwo=3), natomiast drugi indeks wskazuje kolejny numer zmiennej

Źródło: Opracowanie własne na podstawie danych GUS

WERYFIKACJA ZAŁOŻEŃ

Konstrukcję modelu wielopoziomowego rozpoczęto od weryfikacji hipotezy o wielopoziomowej strukturze szacowanej zmiennej. Przypuszczenie to zweryfikowano przy zastosowaniu testu analizy wariancji. Tak więc hipoteza zerowa głosiła, że średni poziom zmiennej objaśnianej różni się w zależności od przynależności do grup utworzonych przez jednostki wyższych poziomów. Jak zauważono wcześniej, szacowana zmienna opisuje aktywność zawodową, która jest uwarunkowana czynnikami ekonomicznymi zróżnicowanymi w przekroju powiatów. W związku z tym istnieją przesłanki merytoryczne, aby podejrzewać, że średni poziom zmiennej objaśnianej różni się pomiędzy jednostkami podziału administracyjnego.

W pierwszej kolejności przeanalizowano zróżnicowanie poziomu natężenia przyjazdów do pracy w przekroju gmin w ramach powiatów (por. wykres 4.1). Średni poziom badanej

zmiennej w grupach gmin różni się znacznie pomiędzy powiatami. Najniższy średni poziom przyjazdów do pracy odnotowano w gminie na prawach powiatu, miasto Świnoujście, wynosił on 0,07. Oznacza to, że zaledwie 7% pracujących na terenie miasta Świnoujścia mieszkało poza gminą zatrudnienia. Z kolei w gminach powiatu śłupeckiego, średni poziom zmiennej objaśnianej wynosił aż 0,85, co oznacza, że aż 85% osób pracujących na terenie tego powiatu mieszkało poza gminą zatrudnienia (por. wykres 4.1). Statystyczną istotność omawianego zróżnicowania zweryfikowano przy pomocy testu analizy wariancji (por. Krzyśko, 1996). W tym celu sformułowano następujący układ hipotez:

$$\begin{cases} H_0: \mu_1 = \dots = \mu_{379} \\ H_1: \sim H_0 \end{cases} \quad (4.1)$$

gdzie:

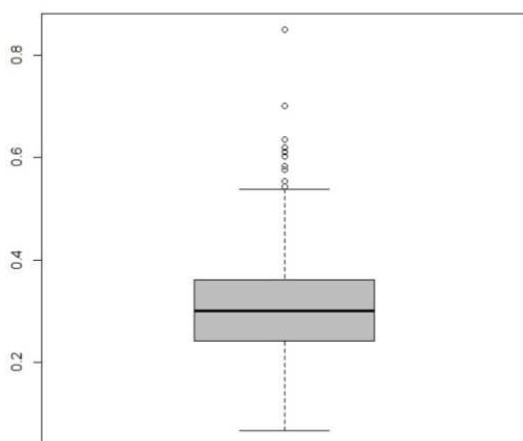
μ_j – średnia wartość natężenia przyjazdów do pracy w gminach należących do *j-tego* powiatu, $j=1, \dots, 379$.

Obliczono statystykę testową o następującej postaci:

$$F = \frac{\frac{\sum_{j=1}^J n_j (\bar{Y}_j - \bar{Y})^2}{J-1}}{\frac{\sum_{j=1}^J \sum_{i=1}^{n_j} (y_{ij} - \bar{Y}_j)^2}{n-J}} = 2,2365 \quad (4.2)$$

gdzie:

\bar{Y}_j – średni poziom natężenia przyjazdów do pracy w *j-tym* powiecie, $j=1, \dots, 379$.



Wykres 4.1. Rozkład średniego natężenia przyjazdów do pracy w gminach, przekrój powiatów, Polska, 2006

Źródło: Opracowanie własne na podstawie danych GUS

Na podstawie powyższej wartości statystyki testowej, należy na praktycznie dowolnym poziomie istotności ($p\text{-value} < 2,2e-16$) odrzucić hipotezę zerową. Zatem można twierdzić, że występuje statystycznie istotne zróżnicowanie powiatów ze względu na średnią wartość szacowanej zmiennej. Upoważnia to do przyjęcia założenia o dwupoziomowej strukturze udziału przyjeżdżających do pracy i konstrukcji modelu dwupoziomowego, w którym jednostka drugiego poziomu zmiennej Y , zdefiniowana jest jako powiat.

Następnie rozważono występowanie zróżnicowania poziomu natężenia przyjazdów do pracy w przekroju gmin w ramach województw (por. wykres 4.2). Najniższy średni poziom badanej zmiennej odnotowano w gminach województwa podlaskiego, wynosił on 0,22, co znaczy, że 22% pracujących w tym województwie mieszkało poza gminą zatrudnienia. Z kolei w gminach województwa wielkopolskiego średni poziom zmiennej objaśnianej wynosił aż 0,41. Oznacza to, że aż 41% osób pracujących w Wielkopolsce, czyli prawie dwukrotnie więcej niż w przypadku województwa podlaskiego, mieszkało poza gminą zatrudnienia (por. wykres 4.2). W celu sprawdzenia statystycznej istotności omawianego zróżnicowania posłużono się ponownie testem analizy wariancji. Weryfikacji poddano następujący układ hipotez:

$$\begin{cases} H_0: \mu_1 = \dots = \mu_{16} \\ H_1: \sim H_0 \end{cases} \quad (4.3)$$

gdzie:

μ_j – średnia wartość natężenia przyjazdów do pracy w gminach należących do j -tego województwa, $j=1, \dots, 16$.

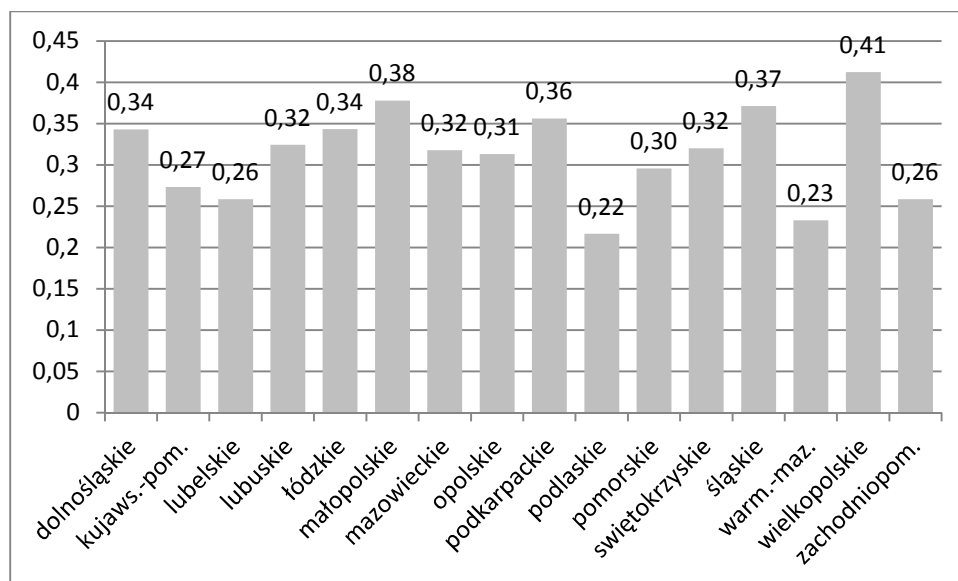
Obliczono statystykę testową następującej postaci:

$$F = \frac{\frac{\sum_{j=1}^{16} n_j (\bar{y}_j - \bar{y})^2}{J-1}}{\frac{\sum_{j=1}^J \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2}{n-J}} = 15,374 \quad (4.4)$$

gdzie:

\bar{Y}_j – średni poziom natężenia przyjazdów do pracy w j -tym województwie, $j=1, \dots, 16$.

W przypadku podziału na województwa, tak jak w przekroju powiatów, na praktycznie dowolnym poziomie istotności ($p\text{-value} < 2,2e-16$) można twierdzić, że występuje zróżnicowanie ze względu na średnią wartość natężenia przyjazdów w gminach. Upoważnia to do przyjęcia założenia o trzypoziomowej strukturze danych oraz do podjęcia próby konstrukcji modelu trzypoziomowego objaśniającego wartości zmiennej Y , gdzie jednostka drugiego poziomu zdefiniowana jest jako powiat, zaś jednostka poziomu trzeciego jako województwo.



Wykres 4.2. Średnie natężenie przyjazdów do pracy w gminach, przekrój województw, Polska, 2006

Źródło: Opracowanie własne na podstawie danych GUS

KONSTRUKCJA MODELU

W konstrukcji modelu trzypoziomowego uwzględniono wszystkie etapy komplikacji opisane w podrozdziale 3.2 wraz z weryfikacją istotności poprawy dopasowania modelu po każdym kroku. Obliczenia wykonano przy pomocy autorskiego kodu napisanego w programie R 2.15.2 z wykorzystaniem biblioteki lme4³⁰ (por. Bates, 2010; Bates, 2013a; Bates, 2013b; Biecek, 2011). Parametry oszacowano zgodnie z metodą największej wiarygodności. Przed rozpoczęciem obliczeń wartości wszystkich zmiennych objaśniających zostały scentrowane, poprzez odjęcie średniej globalnej. Wartości zmiennej objaśnianej nie poddano centrowaniu.

³⁰ Linear mixed-effects models using S4 classes

KLASYCZNA REGRESJA LINIOWA

W celu oceny wyników modelu wielopoziomowego, wyznaczone zostały dwie funkcje klasycznej regresji liniowej, które traktowano jako punkt odniesienia w porównaniach. Pierwsza z nich (w dalszym ciągu podrozdziału indeksowana jako etap 0.0) nie uwzględniała żadnych zmiennych objaśniających:

$$Y_i = \beta_{0,0,0}^{00} + r_i^{00}, \quad r_i^{00} \sim N(0; 0,04198) \quad , \quad (4.5)$$

Szacowanie natężenia przyjazdów do pracy przy pomocy powyższego modelu sprowadza się do przypisania każdej gminie średniej wartości tej zmiennej. Model ten wyznaczony został tylko i wyłącznie w celu porównania w późniejszych etapach, jest on słabej jakości (por. tab. 4.5). Z oszacowanych parametrów wynika, że średnie natężenie przyjazdów do pracy w gminach w 2006 roku wynosiło w Polsce około 32% (por. tab.4.4).

Tabela 4.4. Ocena parametru strukturalnego oraz błędu standardowego szacunku, klasyczny model regresji liniowej natężenia przyjazdów do pracy bez zmiennych objaśniających, Polska, 2006 (wzór 4.5)

Parametr	$\beta_{0,0,0}^{00}$
Współczynnik regresji	0,321149
Błąd standardowy szacunku	0,003702

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.5. Wybrane kryteria oceny jakości dopasowania klasycznego modelu regresji liniowej natężenia przyjazdów do pracy bez zmiennych objaśniających (wzór 4.5)

Kryteria dopasowania		
lnL	AIC	BIC
509,6189	-1015,238	-1003,183

Źródło: Opracowanie własne na podstawie danych GUS

W drugiej funkcji klasycznej regresji liniowej (w dalszej części podrozdziału indeksowanej jako etap 0.1), jako zmienne objaśniające przyjęto zmienne z pierwszego poziomu (por. tab. 4.1). Doboru zmiennych objaśniających dokonano przy zastosowaniu regresji krokowej w przód. Każdorazowo zmienne szeregowano w ten sposób, aby maksymalizować wiarygodność modelu i kolejno dołączano do modelu, dopóki test ilorazu wiarygodności wskazywał na istotną (na poziomie istotności 0,05) poprawę jakości modelu. Jako model wyjściowy przyjęto model bez zmiennych objaśniających opisany powyżej.

$$Y_i = \sum_{q=0}^{18} \beta_q^{01} X_{1,q,i} + r_i^{01} \quad , \quad r_i^{01} \sim N(0; 0,04198) \quad , \quad (4.6)$$

W tabeli 4.6 podano wartości współczynników stojących przy zmiennych, które uznane zostały za istotne zmienne objaśniające. Wartości pozostałych współczynników traktować można jako zera. Ponadto kolejność zmiennych objaśniających w tabeli jest konsekwencją kolejności dołączania ich do modelu w wyniku zastosowania procedury regresji krokowej w przód. Podobne podejście stosowane będzie w dalszej części tego podrozdziału oraz w podrozdziale 4.2.

Tabela 4.6. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, klasyczny model regresji liniowej natężenia przyjazdów do pracy, Polska, 2006 (wzór 4.6)

Parametr	$\beta_{0,0,0}^{01}$	$\beta_{1,0,0}^{01}$	$\beta_{2,0,0}^{01}$	$\beta_{10,0,0}^{01}$	$\beta_{18,0,0}^{01}$	$\beta_{7,0,0}^{01}$	$\beta_{3,0,0}^{01}$	$\beta_{4,0,0}^{01}$
Zmienna	-	$X_{1,1}$	$X_{1,2}$	$X_{1,10}$	$X_{1,18}$	$X_{1,7}$	$X_{1,3}$	$X_{1,4}$
Współczynnik regresji	0,3211	-0,0008	2,6160	0,1164	-0,0001	2,0240	-0,0031	0,4352
Błąd standardowy szacunku	0,0034	0,0001	0,5014	0,0204	0,0000	1,0800	0,0008	0,1489

Źródło: Opracowanie własne na podstawie danych GUS

Warto zauważyć, że w pierwszej kolejności do modelu wprowadzone zostały dwie zmienne najsilniej skorelowane z natężeniem przyjazdów, czyli odległość od ośrodków centralnych oraz natężenie przedsiębiorstw z sekcji D. W następnej kolejności dodany został wskaźnik zatrudnienia, który był stosunkowo słabo skorelowany ze zmienną objaśnianą. Stało się tak, ponieważ wskaźnik ten był jednocześnie słabiej skorelowany z dwiema pierwszymi zmiennymi wprowadzonymi do modelu, niż pozostałe potencjalne zmienne objaśniające, przez co wnosił do modelu nowe informacje.

Otrzymane współczynniki można interpretować następująco. Wyraz wolny (tab. 4.6, $\beta_{0,0,0}^{01}$) to wartość natężenia przyjazdów w gminie o przeciętnych wartościach wszystkich zmiennych objaśniających. Zatem współczynnik ten nie uległ zmianie w stosunku do modelu bez zmiennych objaśniających (por. tab. 4.4 i 4.6). Zwiększenie odległości centroidu gminy od najbliższego ośrodka centralnego o jeden kilometr powoduje spadek natężenia przyjazdów o 0,0008 (tab. 4.6, $\beta_{1,0,0}^{01}$), przy założeniu przeciętnego poziomu pozostałych zmiennych objaśniających. W podobnej sytuacji, wzrost natężenia przedsiębiorstw w sekcji D o 0,01 (przedsiębiorstwa tej sekcji na osobę w wieku produkcyjnym), wiąże się ze wzrostem zmiennej objaśnianej o 0,026 (tab. 4.6, $\beta_{2,0,0}^{01}$). Podobnie wzrost natężenia firm z sekcji I oraz osób fizycznych prowadzących działalność gospodarczą, związany jest ze wzrostem natężenia przyjazdów do pracy (tab. 4.6, $\beta_{7,0,0}^{01}$,

$\beta_{4,0,0}^{01}$). Ponadto w gminach o wysokich wartościach wskaźnika zatrudnienia (tab. 4.6, $\beta_{10,0,0}^{01}$), odnotowano również podwyższone natężenie przyjazdów. Z kolei wzrost natężenia zaludnienia (tab. 4.6, $\beta_{18,0,0}^{01}$) związany jest ze spadkiem wartości zmiennej objaśnianej. Podobną prawidłowość odnotowano w przypadku udziału bezrobotnych (tab. 4.6, $\beta_{3,0,0}^{01}$).

Tabela 4.7. Wybrane kryteria oceny jakości dopasowania klasycznego modelu regresji liniowej natężenia przyjazdów do pracy bez zmiennych objaśniających (wzór 4.6) oraz jego porównanie z etapem 00 (wzór 4.5)

Etap	Kryterium oceny			vs etap 00	
	lnL	AIC	BIC	Chi2	p-value
etap 01	731,4412	-1444,882	-1390,635	443,645	0,0000

Źródło: Opracowanie własne na podstawie danych GUS

Uwzględnienie zmiennych objaśniających w klasycznej regresji liniowej w znacznym stopniu poprawiło jakość dopasowania modelu, ponieważ na praktycznie dowolnym poziomie istotności stwierdzono wzrost jego wiarygodności (por. tab. 4.7).

ETAP 1 – Model bez zmiennych objaśniających

Podetap 1.1

Konstrukcję modelu trzypoziomowego rozpoczęto od uwzględnienia zróżnicowania poziomu natężenia przyjazdów w przekroju powiatów.

$$Y_i = \beta_{0,0,0}^{12} + e_{0(j)0}^{12} + r_i^{12} \quad , \quad (4.7)$$

gdzie:

$$r_i^{12} \sim N(0; 0,0360121) \quad , \quad e_{0(j)0}^{12} \sim N(0; 0,0058456)$$

Tabela 4.8. Ocena parametru strukturalnego oraz błędu standardowego szacunku, dwupoziomowy model natężenia przyjazdów do pracy, etap 1.1, Polska, 2006 (wzór 4.7)

Parametr	$\beta_{0,0,0}^{12}$
Współczynnik regresji	0,317316
Błąd standardowy szacunku	0,005486

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.9. Wybrane kryteria oceny jakości dopasowania modelu dwupoziomowego oszacowanego w etapie 1.1 (wzór 4.7) oraz jego porównanie z etapem 0.0 (wzór 4.5)

Etap	Kryterium oceny	vs etap 0.0
------	-----------------	-------------

	lnL	AIC	BIC	Chi2	p-value
etap 0.1	596,4428	-1186,886	-1168,803	173,648	0,0000

Źródło: Opracowanie własne na podstawie danych GUS

Na podstawie wyznaczonych kryteriów oceny dopasowania modelu, przede wszystkim testu ilorazu wiarygodności stwierdzono lepsze dopasowanie wyznaczonego w podetapie 1.1 modelu niż modelu klasycznej regresji liniowej bez zmiennych objaśniających (por. tab. 4.9).

Podetap 1.2

Następnie uwzględniono również zróżnicowanie natężenia przyjazdów w przekroju województw.

$$Y_i = \beta_{0,0,0}^{13} + e_{0(j)0}^{13} + e_{00(k)}^{13} + r_i^{13} \quad , \quad (4.8)$$

gdzie:

$$r_i^{13} \sim N(0; 0,0359137) \quad , \quad e_{0(j)0}^{13} \sim N(0; 0,0033868) \quad , \quad e_{00(k)}^{13} \sim N(0; 0,0025768)$$

Tabela 4.10. Ocena parametru strukturalnego oraz błędu standardowego szacunku, trzypoziomowy model natężenia przyjazdów do pracy, etap 1.2, Polska, 2006 (wzór 4.8)

Parametr	$\beta_{0,0,0}^{13}$
Współczynnik regresji	0,31149
Błąd standardowy szacunku	0,01363

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.11. Wybrane kryteria oceny jakości dopasowania modelu trzypoziomowego oszacowanego w etapie 1.2 (wzór 4.8) oraz jego porównanie z etapem 1.1 (wzór 4.7)

Etap	Kryterium oceny			vs etap 1.1	
	lnL	AIC	BIC	Chi2	p-value
etap 1.2	630,8902	-1253,78	-1229,67	68,895	0,0000

Źródło: Opracowanie własne na podstawie danych GUS

Uwzględnienie w podetapie 1.2 zróżnicowania natężenia przyjazdów w przekroju województw spowodowało statystycznie istotną poprawę dopasowania modelu w stosunku do podetapu poprzedniego (por. tab. 4.11).

ETAP 2 – model z losowym wyrazem wolnym

W etapie drugim do modelu wprowadzane są zmienne objaśniające z poziomu pierwszego, czyli charakterystyki określone dla gmin:

$$Y_i = \sum_{q=0}^{18} \beta_{q,0,0}^{21} X_{1,q,i} + e_{0(j)0}^{21} + e_{00(k)}^{21} + r_i^{21} \quad , \quad (4.9)$$

gdzie:

$$r_i^{21} \sim N(0; 0,03352814) \quad , \quad e_{0(j)0}^{21} \sim N(0; 0,00251155) \quad , \quad e_{00(k)}^{21} \sim N(0; 0,00026356).$$

Tabela 4.12. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 2, Polska, 2006 (wzór 4.9)

Parametr	$\beta_{0,0,0}^{21}$	$\beta_{2,0,0}^{21}$	$\beta_{1,0,0}^{21}$	$\beta_{18,0,0}^{21}$	$\beta_{10,0,0}^{21}$	$\beta_{7,0,0}^{21}$	$\beta_{3,0,0}^{21}$	$\beta_{8,0,0}^{21}$
Zmienna	-	$X_{1,2}$	$X_{1,1}$	$X_{1,18}$	$X_{1,10}$	$X_{1,7}$	$X_{1,3}$	$X_{1,8}$
Współczynnik regresji	0,31844	2,94168	-0,00076	-0,0001	0,10885	2,84056	-0,00386	0,86256
Błąd standardowy szacunku	0,00605	0,47909	0,00011	0,00001	0,02041	1,02479	0,00103	0,29809

Źródło: Opracowanie własne na podstawie danych GUS

Warto zauważyć, że zestaw zmiennych objaśniających z poziomu gmin jest zbliżony do zmiennych objaśniających uwzględnionych w liniowej funkcji regresji (por. wzór (4.6)), chociaż zmienne dodawane były w nieco innej kolejności. W szczególności w pierwszej kolejności uwzględnione zostało natężenie przedsiębiorstw z sekcji D, pomimo iż odległość od najbliższego ośrodka centralnego jest silniej skorelowana ze zmienną objaśnianą (por. tab. 4.1). Fakt ten można wytłumaczyć silniejszym zróżnicowaniem oddziaływania ośrodków centralnych na natężenie przyjazdów w gminach w przekroju powiatów czy województw, w porównaniu z oddziaływaniem natężenia przedsiębiorstw z sekcji D. Odległości ośrodków centralnych dołączona została do modelu jako druga.

Oszacowane współczynniki kierunkowe pierwszych sześciu zmiennych wprowadzonych do modelu są zbliżone do odpowiednich oszacowań parametrów w klasycznej funkcji regresji (po. tab. 4.12 i 4.6). Jako siódmą zmienną dodano natężenie przedsiębiorstw z sekcji G, co nie miało miejsca w klasycznym modelu regresji. Wzrost tego wskaźnika, przy stałych wartościach pozostałych zmiennych objaśniających, wiąże się ze wzrostem natężenia przyjazdów do pracy w gminach. W przeciwieństwie do klasycznej regresji liniowej,

do modelu trypoziomowego nie wprowadzono zmiennej objaśniającej określonej jako natężenie osób fizycznych prowadzących działalność gospodarczą.

Tabela 4.13. Wybrane kryteria oceny jakości dopasowania modelu trypoziomowego oszacowanego w etapie 2 (wzór 4.9) oraz jego porównanie z etapem 1.2 (wzór 4.8)

Etap	Kryterium oceny			vs etap 1.2	
	lnL	AIC	BIC	Chi2	p-value
etap 2	763,9143	-1505,829	-1439,526	266,05	0,0000

Źródło: Opracowanie własne na podstawie danych GUS

Tak wyznaczony model może być traktowany zarówno jako bezpośrednie rozwinięcie modelu opisanego w etapie 1.2 (por. wzór (4.8)) przez dodanie zmiennych objaśniających z pierwszego poziomu. Jest to również rozwinięcie modelu klasycznej regresji liniowej ze zmiennymi objaśniającymi z poziomu gmin (por. wzór (4.6)), poprzez uwzględnienie efektów losowych powiatu i województwa. W związku z powyższym sprawdzono przewagę modelu z etapu 2, nad każdym z wcześniejszych. W przypadku porównania z modelem bez zmiennych objaśniających test ilorazu wiarygodności wskazuje na istotną statystycznie poprawę wiarygodności modelu (por. tab. 4.13). Modelu oszacowanego w etapie drugim nie należy porównywać z modelem klasycznej regresji liniowej przy pomocy testu ilorazu wiarygodności, ponieważ w modelu klasycznej regresji liniowej dodano jedną zmienną, która nie została wprowadzona do modelu trzypoziomowego, dlatego nie można go traktować jako jego rozszerzenia. Zatem w tym wypadku, ocenę poprawy należy oprzeć tylko na kryteriach informacyjnych. Wartości obu kryteriów pozwoliły stwierdzić poprawę. Wartość kryterium Akaike'a zmniejszyła się o ponad 4%, a kryterium Schwarz'a o 3,5% (por. tab. 4.13 z tab. 4.6).

ETAP 3 – Model z losowym wyrazem wolnym zależnym od zmiennych z wyższych poziomów

Podetap 3.1

W pierwszym podetapie etapu trzeciego do modelu dodano zmienne z poziomu powiatów.

$$Y_i = \sum_{q=0}^{18} \beta_{q,0,0}^{32} X_{1,q,i} + \sum_{q=0}^{10} \beta_{0,q,0}^{32} X_{2,q,i} + e_{0(j)0}^{32} + e_{00(k)}^{32} + r_i^{32} \quad , \quad (4.10)$$

gdzie:

$$r_i^{32} \sim N(0; 0,0359137) \quad , \quad e_{0(j)0}^{32} \sim N(0; 0,0033868) \quad , \quad e_{00(k)}^{32} \sim N(0; 0,0025768)$$

Ponieważ model ten jest rozszerzeniem modelu z etapu 2 o zmienne z poziomu powiatów, jest on wyjaśniany przez dokładnie ten sam zestaw charakterystyk z poziomu gmin. Ponadto odpowiednie oszacowania współczynników kierunkowych przy zmiennych objaśniających z poziomu gmin przyjmują zbliżone wartości (por. Tab. 4.14 i 4.11). Wyjątkiem jest tu współczynnik kierunkowy przy udziale bezrobotnych, różniący się znacznie w stosunku do modelu bez zmiennych objaśniających z poziomu powiatów.

Różnica ta jest wynikiem dodania do modelu analogicznej charakterystyki określonej na poziomie powiatów i omówiona zostanie wraz ze współczynnikiem kierunkowym przy tej zmiennej.

Tabela 4.14. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 3.1, poziom gmin, Polska, 2006 (wzór 4.10)

Parametr	$\beta_{0,0,0}^{32}$	$\beta_{2,0,0}^{32}$	$\beta_{1,0,0}^{32}$	$\beta_{18,0,0}^{32}$	$\beta_{10,0,0}^{32}$	$\beta_{7,0,0}^{32}$	$\beta_{3,0,0}^{32}$	$\beta_{8,0,0}^{32}$
Zmienna	-	$X_{1,2}$	$X_{1,1}$	$X_{1,18}$	$X_{1,10}$	$X_{1,7}$	$X_{1,3}$	$X_{1,8}$
Współczynnik regresji	0,31830	3,08500	-0,00081	-0,00010	0,10000	2,92200	-0,00823	0,86870
Błąd standardowy szacunku	0,00579	0,47730	0,00011	0,00001	0,02047	1,02000	0,00159	0,29720

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.15. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 3.1, poziom powiatów, Polska, 2006 (wzór 4.10)

Parametr	$\beta_{0,1,0}^{32}$	$\beta_{0,9,0}^{32}$
Zmienna	$X_{2,1}$	$X_{2,9}$
Współczynnik regresji	0,00757	0,28860
Błąd standardowy szacunku	0,00208	0,09289

Źródło: Opracowanie własne na podstawie danych GUS

W pierwszej kolejności do modelu włączona została charakterystyka bezrobocia określona na poziomie powiatów. Należy podkreślić, że ponieważ model objaśniany jest również przez analogiczną zmienną określoną na poziomie gmin, obydwie te charakterystyki należy interpretować łącznie. Mianowicie bezrobocie jest niewątpliwie destymulantą natężenia przyjazdów, jednak gdy wszystkie gminy w okolicy (tzn. w powiecie) charakteryzują się wysokim poziomem bezrobocia, wpływ wysokiego poziomu tej zmiennej w konkretnej gminie w mniejszym stopniu wpływa na obniżenie natężenia przyjazdów, niż w przypadku gminy o wysokim udziale bezrobotnych leżącej w powiecie, gdzie udział bezrobotnych jest niski. Stąd, chociaż oszacowanie współczynnika kierunkowego przy udziale bezrobotnych na poziomie powiatów przyjmuje wartość dodatnią (por. tab. 4.15), to jego łączna wartość ze współczynnikiem kierunkowym tej zmiennej określonej na poziomie gmin (por. tab. 4.14) świadczy o ujemnym wpływie bezrobocia na natężenie przyjazdów. Drugą zmienną objaśniającą wprowadzoną do modelu jest udział osób pracujących w warunkach zagrożenia. Przy ustalonych wartościach pozostałych zmiennych objaśniających wzrost osób pracujących w warunkach zagrożenia o około cztery wiąże się z przyjazdem jednej dodatkowej osoby (por. tab. 4.15). Zależność tą tłumaczyć można tym, że stanowiska wymagające pracy w warunkach zagrożenia wymagają często również specyficznych kwalifikacji, z kolei

osoby posiadające szczególne kwalifikacje często zmuszone są dojeżdżać do pracy w której mogą je wykorzystać.

Tabela 4.16. Wybrane kryteria oceny jakości dopasowania modelu trzypoziomowego oszacowanego w etapie 3.1 (wzór 4.10) oraz jego porównanie z etapem 2 (wzór 4.9)

Etap	Kryterium oceny			vs etap 2	
	lnL	AIC	BIC	Chi2	p-value
etap 3.1	774,3858	-1522,772	-1444,414	20,9430	0,0000

Źródło: Opracowanie własne na podstawie danych GUS

Po wprowadzeniu dwóch zmiennych objaśniających stwierdzono statystycznie istotną poprawę jakości modelu w stosunku do etapu poprzedniego.

Podetap 3.2

W drugim podetapie etapu trzeciego podjęto próbę dodania do modelu zmiennych objaśniających z poziomu województw. Jednak dodanie żadnej ze zmiennych określonych na tym poziomie nie poprawiło w istotnym stopniu jakości modelu. W związku z tym jako model wyjściowy przy konstrukcji modelu w następnym etapie przyjęto model z etapu poprzedniego, 3.1 (por. wzór (4.10)). Nie wyklucza to jednak wykorzystania zmiennych objaśniających z poziomu województw w celu zróżnicowania zależności pomiędzy natężeniem przyjazdów a zmiennymi objaśniającymi z poziomu gmin oraz powiatów w przekroju województw w etapie piątym.

ETAP 4 - Model z losowymi współczynnikami regresji

Podetap. 4.1

W pierwszym podetapie etapu czwartego dopuszcza się uwzględnienie zróżnicowanie współczynników kierunkowych przy zmiennych objaśniających z poziomu gmin w przekroju powiatów. Należy zaznaczyć, że dodana może zostać część losowa współczynnika kierunkowego, w przypadku którego część stała nie była istotna i nie została wprowadzona do modelu w etapie drugim.

$$Y_i = \sum_{q=0}^{18} \beta_{q,0,0}^{4(1/2)} X_{1,q,i} + \sum_{q=0}^{10} \beta_{0,q,0}^{4(1/2)} X_{2,q,i} + \sum_{q=1}^{18} e_{q(j)0}^{4(1/2)} X_{1,q,i} + e_{0(j)0}^{4(1/2)} + e_{00(k)}^{4(1/2)} + r_i^{4(1/2)} \quad (4.11)$$

gdzie: $r_i^{4(1/2)} \sim N(0; 0,0248430)$, $e_{00(k)}^{4(1/2)} \sim N(0; 0,0001969)$,

$$\begin{bmatrix} e_{0(j)0}^{4(1/2)} \\ e_{16(j)0}^{4(1/2)} \\ e_{10(j)0}^{4(1/2)} \\ e_{4(j)0}^{4(1/2)} \\ e_{14(j)0}^{4(1/2)} \\ e_{2(j)0}^{4(1/2)} \end{bmatrix} \sim N \left(\mathbf{0}; \begin{bmatrix} 0,0022 & -0,1941 & -0,0058 & 0,0056 & -0,0425 & -0,0425 \\ -0,1941 & 99110 & -54,392 & 563,59 & -1366,8 & -684,08 \\ -0,0058 & -54,392 & 0,1393 & -0,4615 & 0,5574 & -0,0636 \\ 0,0056 & 563,59 & -0,4615 & 3,4274 & -7,5707 & -2,7905 \\ -0,0425 & -1366,8 & 0,5574 & -7,5707 & 21,792 & 4,1859 \\ 0,0689 & -684,08 & -0,0636 & -2,7905 & 4,1859 & 30,262 \end{bmatrix} \right)$$

Uwzględnienie losowego czynnika związanego z poziomem powiatów poprawiło dopasowanie modelu w przypadku współczynników kierunkowych pięciu zmiennych objaśniających z poziomu gmin. Były to: natężenie spółek handlowych z udziałem kapitału zagranicznego z sektora publicznego, wskaźnik zatrudnienia, natężenie osób fizycznych prowadzących działalność gospodarczą, natężenie przedsiębiorstw z sekcji K oraz natężenie przedsiębiorstw z sekcji D. Ponadto stwierdzono, że korelacja pomiędzy wprowadzonymi efektami losowymi jest istotna dla każdej pary. W związku z tym uwzględniono pełną macierz korelacji dla efektów losowych na poziomie powiatów.

Tabela 4.17. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 4.1, poziom gmin, Polska, 2006 (wzór 4.11)

Parametr	$\beta_{0,0,0}^{4(1/2)}$	$\beta_{2,0,0}^{4(1/2)}$	$\beta_{1,0,0}^{4(1/2)}$	$\beta_{18,0,0}^{4(1/2)}$	$\beta_{10,0,0}^{4(1/2)}$	$\beta_{7,0,0}^{4(1/2)}$	$\beta_{3,0,0}^{4(1/2)}$	$\beta_{8,0,0}^{4(1/2)}$
Zmienna	-	$X_{1,2}$	$X_{1,1}$	$X_{1,18}$	$X_{1,10}$	$X_{1,7}$	$X_{1,3}$	$X_{1,8}$
Współczynnik regresji	0,3192	4,6560	-0,0008	-0,0001	0,0356	2,5700	-0,0090	0,8006
Błąd standardowy szacunku	0,0056	0,7258	0,0001	0,0000	0,0373	0,9899	0,0015	0,3220

Źródło: Opracowanie własne na podstawie danych GUS

Należy rozumieć to w ten sposób, że zależność pomiędzy natężeniem przyjazdów a każdą z wymienionych powyżej zmiennych różni się w przekroju powiatów. Ponadto istnieje związek pomiędzy nasileniem się tej zależności ze zmienną objaśnianą w przypadku każdej pary zmiennych objaśniających. Tłumaczyć można to występowaniem pewnej ukrytej zmiennej charakteryzującej stopień rozwoju gospodarczego oraz sytuację ogólną w powiatach. W szczególności w indywidualny sposób traktowane są miasta na prawach powiatu.

Tabela 4.18. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 4.1, poziom powiatów, Polska, 2006 (wzór 4.11)

Parametr	$\beta_{0,1,0}^{4(1/2)}$	$\beta_{0,9,0}^{4(1/2)}$
Zmienna	$X_{2,1}$	$X_{2,9}$
Współczynnik regresji	0,0087	0,2557
Błąd standardowy szacunku	0,001948	0,0907

Źródło: Opracowanie własne na podstawie danych GUS

Cześć stała modelu, czyli zestaw zmiennych objaśniających pozostał taki sam jak wcześniej. W wyniku poszerzenia części losowej niewielkim zmianom uległy jedynie oszacowania wyrazu wolnego i współczynników kierunkowych (por. tab. 4.17 i 4.18). Wyjątkiem są współczynniki kierunkowe przy tych zmiennych z poziomu gmin, które objaśniane są zarówno przez część stałą wprowadzoną w etapie drugim, jak i składnik losowy z poziomu powiatów. Są to współczynniki kierunkowe przy natężeniu przedsiębiorstw z sekcji D oraz wskaźniku zatrudnienia, które uległy znacznym zmianom (por. tab. 4.17 i 4.14). Stało się tak dlatego, że liczba gmin wchodzących w skład powiatów różni się znacznie między powiatami. W szczególności miasta na prawach powiatu to powiaty składające się z jednej gminy. Co więcej, specyfika miast na prawach powiatu jest odmienna niż pozostałych gmin. W szczególności związek pomiędzy natężeniem przyjazdów a natężeniem przedsiębiorstw z sekcji D jest w tych miastach silniejszy niż w pozostałych gminach, zaś w przypadku wskaźnika zatrudnienia słabszy.

Tabela 4.19. Wybrane kryteria oceny jakości dopasowania modelu trypoziomowego oszacowanego w etapie 4.1 (wzór 4.11) oraz jego porównanie z etapem 3.1 (wzór 4.10)

Etap	Kryterium oceny			vs etap 3.1	
	lnL	AIC	BIC	Chi2	p-value
etap 4.1	970,7558	-1875,512	-1676,605	392,74	0,0000

Źródło: Opracowanie własne na podstawie danych GUS

Po uwzględnieniu omówionych wyżej zmian części losowej modelu uzyskano istotną poprawę jego jakości (por. tab. 4.19). Wskazuje na to test ilorazu wiarygodności oraz znaczne obniżenie wartości kryteriów informacyjnych (w przypadku kryterium Akaike'a o 23%, zaś w przypadku kryterium Schwarz'a o 16%).

Podetap 4.2

W drugim podetapie etapu czwartego dopuszcza się uwzględnienie zróżnicowanie współczynników kierunkowych przy zmiennych objaśniających z poziomu gmin w przekroju województw.

$$Y_i = \sum_{q=0}^{18} \beta_{q,0,0}^{4(1/3)} X_{1,q,i} + \sum_{q=0}^{10} \beta_{0,q,0}^{4(1/3)} X_{2,q,i} + \sum_{q=1}^{18} e_{q(j)0}^{4(1/3)} X_{1,q,i} + \sum_{q=1}^{18} e_{q,0(k)}^{4(1/3)} X_{1,q,i} + e_{0(j)0}^{4(1/3)} + e_{00(k)}^{4(1/3)} + r_i^{4(1/3)} \quad , \quad (4.12)$$

gdzie: $r_i^{4(1/3)} \sim N(0; 0,024843)$,

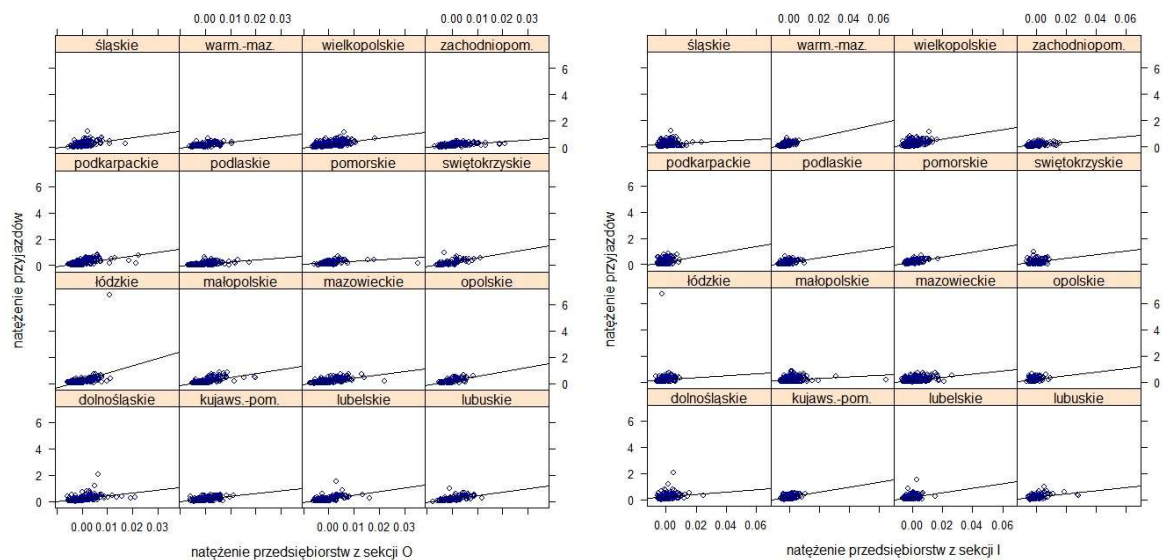
$$\begin{bmatrix} e_{0(j)0}^{4(1/3)} \\ e_{16(j)0}^{4(1/3)} \\ e_{10(j)0}^{4(1/3)} \\ e_{4(j)0}^{4(1/3)} \\ e_{14(j)0}^{4(1/3)} \\ e_{2(j)0}^{4(1/3)} \end{bmatrix} \sim N \left(\mathbf{0}; \begin{bmatrix} 0,0020 & -0,0939 & -0,0069 & 0,0089 & -0,0343 & 0,0402 \\ -0,0939 & 121620 & -59,535 & 619,55 & -1559 & -811,96 \\ -0,0069 & -59,535 & 0,1402 & -0,5170 & 0,8119 & -0,6398 \\ 0,0089 & 619,55 & -0,5170 & 3,5793 & -7,9727 & -1,9868 \\ -0,0343 & -1559 & 0,8119 & -7,9727 & 20,936 & 8,9531 \\ 0,0402 & -811,96 & -0,6398 & -1,9868 & 8,9531 & 24,539 \end{bmatrix} \right)$$

$$, \begin{bmatrix} e_{00(k)}^{4(1/3)} \\ e_{13,0(k)}^{4(1/3)} \\ e_{7,0(k)}^{4(1/3)} \end{bmatrix} \sim N \left(\mathbf{0}; \begin{bmatrix} 0,0003 & 25,291 & 0 \\ 25,291 & 27,746 & 0 \\ 0 & 0 & 23,054 \end{bmatrix} \right) .$$

Uwzględnienie losowego czynnika związanego z poziomem województw poprawiło dopasowanie modelu w przypadku współczynników kierunkowych dwóch zmiennych objaśniających z poziomu gmin. Były to: natężenie przedsiębiorstw z sekcji O ($X_{1,13}$) oraz natężenie przedsiębiorstw z sekcji I ($X_{1,7}$). Ponadto stwierdzono, że korelacja pomiędzy wprowadzoną w pierwszej kolejności zmienną, czyli natężeniem przedsiębiorstw z sekcji O a wyrazem wolnym efektu losowego z poziomu województw jest statystycznie istotna. Jednak nie można było tego stwierdzić w odniesieniu do korelacji współczynnika przy natężeniu przedsiębiorstw z sekcji I, w przypadku zależności zarówno z wyrazem wolnym jak i współczynnikiem przy natężeniu przedsiębiorstw z sekcji O. Z istotności badanych relacji wynikało więc, że we wprowadzonej macierzy kowariancji konieczne jest uwzględnienie tylko korelacji pomiędzy natężeniem przedsiębiorstw z sekcji O a wyrazem wolnym składnika losowego z poziomu województw.

Wyjaśniając oddziaływanie czynnika losowego z poziomu województw, przedstawiono relację między natężeniem przyjazdów do pracy a natężeniem przedsiębiorstw z sekcji O oraz natężeniem przedsiębiorstw z sekcji I, odrębnie w gminach każdego województwa. Ocenę zróżnicowania w tym zakresie ułatwia graficzne przedstawienie zależności (por. wykres 4.3). Szczególnie wyraźna jest odmienność relacji między zmienną $X_{1,13}$

w województwach opolskim i świętokrzyskim, w których jest ona silniejsza, niż w pozostałych. W województwie łódzkim natomiast wyraźny jest wpływ obserwacji odstającej, którą jest gmina Bedlno. Podobne zróżnicowanie w przekroju województw zaobserwować można dla między natężeniem przyjazdów do pracy a zmienną $X_{1,7}$. W tym przypadku, wyraźna odmienność dotyczy województw warmińsko-mazurskiego. Ponadto w wyniku przeprowadzonych analiz stwierdzono, że istnieje związek pomiędzy średnim poziomem natężenia przyjazdów oraz siłą zależności pomiędzy natężeniem przyjazdów a natężeniem przedsiębiorstw z sekcji O w przekroju województw. Nie stwierdzono jednak wpływu średniego poziomu natężenia przyjazdów w województwach na nasilenie zależności pomiędzy natężeniem przedsiębiorstw z sekcji I a zmienną objaśnianą.



Wykres 4.3. Zależność pomiędzy zmienną objaśnianą i zmiennymi objaśniającymi w przekroju gmin, dla każdego z województw rozważana indywidualnie

Wykres 4.3.a

Wykres 4.3.b

Linie regresji liniowej dla zmiennej $X_{1,13}$

Linie regresji liniowej dla zmiennej $X_{1,7}$

Źródło: Opracowanie własne na podstawie danych GUS

Cześć stała modelu, czyli zestaw zmiennych objaśniających pozostał taki sam jak wcześniej. W wyniku poszerzenia części losowej niewielkim zmianom uległy jedynie oszacowania wyrazu wolnego i współczynników kierunkowych (por. tab. 4.17 i 4.18). Należy zaznaczyć, że natężenie przedsiębiorstw z sekcji I zostało wprowadzone do modelu, jako zmienna objaśniająca, w etapie drugim. W bieżącym etapie odnotowano dodatkowo zróżnicowanie zależności tej zmiennej z natężeniem przyjazdów w przekroju województw. Jednakże oszacowanie części stałej stojącego przy niej współczynnika

kierunkowego nie uległo znacznej zmianie (por. tab. 4.20 i 4.17). Jest tak dlatego, że liczba gmin wchodzących w skład województw nie różni się wyraźnie pomiędzy województwami.

Tabela 4.20. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 4.2, poziom gmin, Polska, 2006 (wzór 4.12)

Parametr	$\beta_{0,0,0}^{4(1/3)}$	$\beta_{2,0,0}^{4(1/3)}$	$\beta_{1,0,0}^{4(1/3)}$	$\beta_{18,0,0}^{4(1/3)}$	$\beta_{10,0,0}^{4(1/3)}$	$\beta_{7,0,0}^{4(1/3)}$	$\beta_{3,0,0}^{4(1/3)}$	$\beta_{8,0,0}^{4(1/3)}$
Zmienna	-	$X_{1,2}$	$X_{1,1}$	$X_{1,18}$	$X_{1,10}$	$X_{1,7}$	$X_{1,3}$	$X_{1,8}$
Współczynnik regresji	0,3236	4,8890	-0,0008	-0,0001	0,0350	2,7280	-0,0091	0,8741
Błąd standardowy szacunku	0,0051	0,6970	0,0001	0,00001	0,0382	1,6200	0,0015	0,3314

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.21. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 4.2, poziom powiatów, Polska, 2006 (wzór 4.12)

Parametr	$\beta_{0,1,0}^{4(1/3)}$	$\beta_{0,9,0}^{4(1/3)}$
Zmienna	$X_{2,1}$	$X_{2,9}$
Współczynnik regresji	0,0085	0,2270
Błąd standardowy szacunku	0,0019	0,0885

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.22. Wybrane kryteria oceny jakości dopasowania modelu trzypoziomowego oszacowanego w etapie 4.2 (wzór 4.12) oraz jego porównanie z etapem 4.1 (wzór 4.11)

Etap	Kryterium oceny			vs etap 4.1	
	lnL	AIC	BIC	Chi2	p-value
etap 4.2	980,5272	-1889,054	-1672,065	19,543	0,0002

Źródło: Opracowanie własne na podstawie danych GUS

Po poszerzeniu modelu o omówione powyżej efekty losowe z poziomu województw, na poprawę jakości dopasowania modelu wskazywało kryterium Akaike'a oraz test ilorazu wiarygodności, jednak zgodnie z kryterium Schwarz'a wprowadzona komplikacja modelu nie była „opłacalna” (por. tab. 4.22 i 4.19). Ponieważ jako decydujące kryterium przyjęto test ilorazu wiarygodności, model oszacowany w tym etapie stał się modelem wyjściowym dla dalszego toku rozważań.

Podetap 4.3

W trzecim podetapie etapu czwartego podjęto próbę uwzględnienia w modelu zróżnicowania nasilenia zależności pomiędzy natężeniem przyjazdów a zmiennymi objaśniającymi z poziomu powiatów w przekroju województw. Jednak w przypadku żadnej zmiennej z poziomu powiatów nie uzyskano istotnej poprawy stopnia jakości modelu. Z tego powodu, jako model wyjściowy w następnym etapie przyjęto model z etapu poprzedniego, 4.2 (por. wzór (4.12)). Podejście to nie wyklucza wykorzystania zmiennych objaśniających z poziomu powiatów w celu wyjaśnienia zróżnicowania zależności pomiędzy natężeniem przyjazdów a zmiennymi objaśniającymi z poziomu gmin w etapie piątym.

ETAP 5 – Model z losowymi współczynnikami regresji zależnymi od zmiennych z wyższych poziomów

Podetap 5.1

W pierwszym podetapie etapu piątego możliwe jest uwzględnienie części zmienności zależności pomiędzy natężeniem przyjazdów a zmiennymi objaśniającymi z poziomu gmin przy pomocy charakterystyk z poziomu powiatów. Możliwe jest uwzględnienie takiego zróżnicowania zarówno w przypadku wszystkich zmiennych objaśniających z poziomu gmin, niezależnie od tego, czy zostały wprowadzone do modelu w etapie drugim oraz czy w etapie 4.1 stwierdzono ich istotne zróżnicowanie w ramach powiatów.

$$Y_i = \sum_{q=0}^{18} \beta_{q,0,0}^{5(1/2)} X_{1,q,i} + \sum_{q=0}^{10} \beta_{0,q,0}^{5(1/2)} X_{2,q,i} + \sum_{q=1}^{18} e_{q(j)0}^{5(1/2)} X_{1,q,i} + \sum_{q=1}^{18} \sum_{s=0}^{10} \beta_{q,s,0}^{5(1/2)} X_{1,q,i} X_{2,s,i} + \sum_{q=1}^{18} e_{q,0(k)}^{5(1/2)} X_{1,q,i} + e_{0(j)0}^{5(1/2)} + e_{00(k)}^{5(1/2)} + r_i^{5(1/2)}, \quad (4.13)$$

gdzie:

$$r_i^{5(1/2)} \sim N(0; 0,024604)$$

$$\begin{bmatrix} e_{0(j)0}^{5(1/2)} \\ e_{16(j)0}^{5(1/2)} \\ e_{10(j)0}^{5(1/2)} \\ e_{4(j)0}^{5(1/2)} \\ e_{14(j)0}^{5(1/2)} \\ e_{2(j)0}^{5(1/2)} \end{bmatrix} \sim N \left(\mathbf{0}; \begin{bmatrix} 0,0020 & 0,0580 & -0,0065 & 0,0063 & -0,0291 & 0,0369 \\ 0,0580 & 105790 & -55,686 & 549,53 & -1408,5 & -501,27 \\ -0,0065 & -55,686 & 0,1521 & -0,5061 & 0,7923 & -1,1537 \\ 0,0063 & 549,53 & -0,5061 & 3,2581 & -7,3266 & 0,0240 \\ -0,0291 & -1408,5 & 0,7923 & -7,3266 & 19,172 & 6,6706 \\ 0,0369 & -501,27 & -1,1537 & 0,0240 & 6,6706 & 19,613 \end{bmatrix} \right)$$

$$\begin{bmatrix} e_{00(k)}^{5(1/2)} \\ e_{13,0(k)}^{5(1/2)} \\ e_{7,0(k)}^{5(1/2)} \end{bmatrix} \sim N \left(\mathbf{0}; \begin{bmatrix} 0,0002 & 0,0855 & 0 \\ 0,0855 & 30,866 & 0 \\ 0 & 0 & 24,928 \end{bmatrix} \right)$$

Tabela 4.23. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.1, poziom gmin, Polska, 2006 (wzór 4.13)

Parametr	$\beta_{0,0,0}^{5(1/2)}$	$\beta_{2,0,0}^{5(1/2)}$	$\beta_{1,0,0}^{5(1/2)}$	$\beta_{18,0,0}^{5(1/2)}$	$\beta_{10,0,0}^{5(1/2)}$	$\beta_{7,0,0}^{5(1/2)}$	$\beta_{3,0,0}^{5(1/2)}$	$\beta_{8,0,0}^{5(1/2)}$
Zmienna	-	$X_{1,2}$	$X_{1,1}$	$X_{1,18}$	$X_{1,10}$	$X_{1,7}$	$X_{1,3}$	$X_{1,8}$
Współczynnik regresji	0,3248	4,8010	-0,0008	-0,0001	0,0408	3,1210	-0,0090	0,8034
Błąd standardowy szacunku	0,0050	0,6711	0,0001	0,00001	0,0389	1,6580	0,0015	0,3287

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.24. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.1, poziom powiatów, Polska, 2006 (wzór 4.13)

Parametr	$\beta_{0,1,0}^{5(1/2)}$	$\beta_{0,9,0}^{5(1/2)}$
Zmienna	$X_{2,1}$	$X_{2,9}$
Współczynnik regresji	0,00761	0,1556
Błąd standardowy szacunku	0,001936	0,09089

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.25. Oceny współczynników kierunkowych interakcji między zmiennymi objaśniającymi z różnych poziomów oraz błędy standardowe szacunku, model natężenia przyjazdów do pracy, etap 5.1, Polska, 2006 (wzór 4.13)

Parametr	$\beta_{5,6,0}^{5(1/2)}$	$\beta_{0,3,9}^{5(1/2)}$
Zmienna	$X_{1,5} \times X_{2,6}$	$X_{1,3} \times X_{2,9}$
Współczynnik regresji	-2,782	-0,04559
Błąd standardowy szacunku	1,03	0,01832

Źródło: Opracowanie własne na podstawie danych GUS

Poprawę jakości dopasowania modelu uzyskano przez wyjaśnienie części zależności pomiędzy natężeniem przyjazdów interakcją między natężeniem przedsiębiorstw z sektora prywatnego a wskaźnikiem zatrudnienia oraz relacją pomiędzy udziałem bezrobotnych oraz udziałem pracujących w warunkach zagrożenia (por. tab. 4.25).

Warto zauważyć, że w przypadku pierwszej ze wskazanych interakcji obydwie zmienne, z poziomu gmin i powiatów, nie były wcześniej uwzględnione w modelu. Natężenie przedsiębiorstw prywatnych określone dla gmin nie zostało wprowadzone do modelu w etapie drugim, ponieważ w wyniku korelacji z wcześniej uwzględnionymi zmiennymi nie wносиło nowej informacji, jednak gdyby charakterystyka ta została uwzględniona byłaby stymulantą natężenia przyjazdów (por. tab. 4.1). Ujemną wartość współczynnika przy interakcji tej zmiennej ze wskaźnikiem zatrudnienia w przekroju powiatów można uzasadniać słabszą relacją między natężeniem przyjazdów a zmiennymi skorelowanymi z natężeniem przedsiębiorstw z sektora prywatnego, w powiatach o wysokim wskaźniku zatrudnienia. Ujemną wartość oszacowania tego współczynnika niweluje znaczny wzrost współczynników kierunkowych charakteryzujących natężenie przedsiębiorstw z sekcji I oraz współczynnika zatrudnienia określonych dla gmin (por. tab. 4.23 i 4.20).

W przypadku interakcji pomiędzy udziałem bezrobotnych i udziałem osób pracujących w warunkach zagrożenia, obydwie zmienne były już wcześniej wprowadzone do modelu jako samodzielne zmienne objaśniające. Jak stwierdzono wcześniej udział bezrobotnych jest destymulantą natężenia przyjazdów. Zatem ujemną wartość współczynnika przy uwzględnionej interakcji należy interpretować tak, że w powiatach charakteryzujących się wysokim udziałem osób pracujących w warunkach zagrożenia zależność pomiędzy natężeniem przyjazdów a udziałem bezrobotnych zyskuje na sile. Warto zauważyć również, że uwzględnienie omówionej interakcji wpłynęło na obniżenie o ponad 30% współczynnika kierunkowego przy udziale osób pracujących w warunkach zagrożenia w powiatach zmiennej oraz spadek o ponad 10% współczynnika kierunkowego przy udziale bezrobotnych w przekroju powiatów (por. tab. 4.24 i 4.19).

Tabela 4.26. Wybrane kryteria oceny jakości dopasowania modelu trypoziomowego oszacowanego w etapie 5.1 (wzór 4.13) oraz jego porównanie z etapem 4.2 (wzór 4.12)

Etap	Kryterium oceny			vs etap 4.2	
	lnL	AIC	BIC	Chi2	p-value
etap 5.1	987,65	-1899,3	-1670,256	14,246	0,0008

Źródło: Opracowanie własne na podstawie danych GUS

Na podstawie testu ilorazu wiarygodności jako model wyjściowy w następnym podetapie przyjęto model z wprowadzonymi w tym podetapie interakcjami.

Podetap 5.2

W drugim podetapie etapu piątego możliwe jest wyjaśnienie części zmienności zależności pomiędzy natężeniem przyjazdów a zmiennymi objaśniającymi z poziomu gmin przy pomocy charakterystyk z poziomu województw. Możliwe jest uwzględnienie takiego zróżnicowania zarówno w przypadku wszystkich zmiennych objaśniających z poziomu gmin, niezależnie od tego, czy zostały wprowadzone do modelu w etapie drugim oraz czy w etapie 4.2 stwierdzono ich istotne zróżnicowanie w ramach województw.

$$\begin{aligned}
 Y_i = & \sum_{q=0}^{18} \beta_{q,0,0}^{5(1/3)} X_{1,q,i} + \sum_{q=0}^{10} \beta_{0,q,0}^{5(1/3)} X_{2,q,i} + \sum_{q=1}^{18} e_{q(j)0}^{5(1/3)} X_{1,q,i} + \\
 & \sum_{q=1}^{18} \sum_{s=0}^{10} \beta_{q,s,0}^{5(1/3)} X_{1,q,i} X_{2,s,i} + \sum_{q=1}^{18} \sum_{s=0}^{11} \beta_{q,0,s}^{5(1/3)} X_{1,q,i} X_{3,s,i} + \sum_{q=1}^{18} e_{q,0(k)}^{5(1/3)} X_{1,q,i} + \\
 & e_{0(j)0}^{5(1/3)} + e_{00(k)}^{5(1/3)} + r_i^{5(1/3)} \quad (4.14)
 \end{aligned}$$

gdzie: $r_i^{5(1/3)} \sim N(0; 0,024719)$

$$\begin{aligned}
 & \begin{bmatrix} e_{0(j)0}^{5(1/3)} \\ e_{16(j)0}^{5(1/3)} \\ e_{10(j)0}^{5(1/3)} \\ e_{4(j)0}^{5(1/3)} \\ e_{14(j)0}^{5(1/3)} \\ e_{2(j)0}^{5(1/3)} \end{bmatrix} \sim N \left(\mathbf{0}; \begin{bmatrix} 0,0019 & 0,2433 & -0,0072 & 0,0096 & -0,0353 & 0,0419 \\ 0,2433 & 94362 & -52,476 & 505,96 & -1288,6 & -392,86 \\ -0,0072 & -52,476 & 0,1557 & -0,4922 & 0,7569 & -1,2466 \\ 0,0096 & 505,96 & -0,4922 & 3,0702 & -6,9098 & 0,4187 \\ -0,0353 & -1288,6 & 0,7569 & -6,9098 & 18,175 & 5,6409 \\ 0,0419 & -392,86 & -1,2466 & 0,4187 & 5,6409 & 19,583 \end{bmatrix} \right) \\
 & , \begin{bmatrix} e_{00(k)}^{5(1/3)} \\ e_{13,0(k)}^{5(1/3)} \\ e_{7,0(k)}^{5(1/3)} \end{bmatrix} \sim N \left(\mathbf{0}; \begin{bmatrix} 0,0003 & 0,0836 & 0 \\ 0,0836 & 24,71 & 0 \\ 0 & 0 & 22,362 \end{bmatrix} \right) .
 \end{aligned}$$

Tabela 4.27. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.2, poziom gmin, Polska, 2006 (wzór 4.14)

Parametr	$\beta_{0,0,0}^{5(1/3)}$	$\beta_{2,0,0}^{5(1/3)}$	$\beta_{1,0,0}^{5(1/3)}$	$\beta_{18,0,0}^{5(1/3)}$	$\beta_{10,0,0}^{5(1/3)}$	$\beta_{7,0,0}^{5(1/3)}$	$\beta_{3,0,0}^{5(1/3)}$	$\beta_{8,0,0}^{5(1/3)}$
Zmienna	-	$X_{1,2}$	$X_{1,1}$	$X_{1,18}$	$X_{1,10}$	$X_{1,7}$	$X_{1,3}$	$X_{1,8}$
Współczynnik regresji	0,3244	4,8540	-0,0008	-0,0001	0,0346	3,1690	-0,0090	0,8555
Błąd standardowy szacunku	0,0052	0,6737	0,0001	0,00001	0,0390	1,6060	0,0015	0,3287

Zródło: Opracowanie własne na podstawie danych GUS

Tabela 4.28. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.2, poziom powiatów, Polska, 2006 (wzór 4.14)

Parametr	$\beta_{0,1,0}^{5(1/3)}$	$\beta_{0,9,0}^{5(1/3)}$
Zmienna	$X_{2,1}$	$X_{2,9}$
Współczynnik regresji	0,0081	0,1654
Błąd standardowy szacunku	0,0019	0,0906

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.29. Oceny współczynników kierunkowych interakcji między zmiennymi objaśniającymi z różnych poziomów oraz błędy standardowe szacunku, model natężenia przyjazdów do pracy, etap 5.2, Polska, 2006 (wzór 4.14)

Parametr	$\beta_{5,6,0}^{5(1/3)}$	$\beta_{3,9,0}^{5(1/3)}$	$\beta_{1,0,7}^{5(1/3)}$
Zmienna	$X_{1,5} \times X_{2,6}$	$X_{1,3} \times X_{2,9}$	$X_{1,1} \times X_{3,7}$
Współczynnik regresji	-2,828	-0,0466	-0,1371
Błąd standardowy szacunku	1,025	0,0182	0,0563

Źródło: Opracowanie własne na podstawie danych GUS

Istotną poprawę jakości modelu spowodowało uwzględnienie wpływu udziału absolwentów uczelni państwowych w województwach na nasilenie zależności pomiędzy natężeniem przyjazdów a odległością od najbliższego ośrodka centralnego (por. tab. 4.29). Mianowicie w województwach o wysokich wartościach tak określonego wskaźnika wykształcenia, wzrasta siła oddziaływania odległości od ośrodków centralnych na natężenie przyjazdów (por. tab. 4.27 i 4.29). Relację tą można wyjaśnić skłonnością absolwentów do podejmowania dojazdów do pracy w celu objęcia stanowiska odpowiadającego ich kwalifikacjom. Z drugiej strony, po ukończeniu studiów, absolwenci często pozostają w mieście, w którym studiowali lub jego okolicach, rzadko zaś wyprawdzają się w rejony peryferyjne, nawet jeżeli z nich pochodzą.

Tabela 4.30. Wybrane kryteria oceny jakości dopasowania modelu trypoziomowego oszacowanego w etapie 5.2 (wzór 4.14) oraz jego porównanie z etapem 5.1 (wzór 4.13)

Etap	Kryterium oceny			vs etap 5.1	
	lnL	AIC	BIC	Chi2	p-value
etap 5.2	991,1469	-1904,294	-1669,222	6,9937	0,0082

Źródło: Opracowanie własne na podstawie danych GUS

Uwzględnienie omówionej interakcji między zmiennymi spowodowało istotną, ze względu na wynik testu ilorazu wiarygodności, poprawę dopasowania modelu. Zatem

od modelu oszacowanego w tym etapie rozpoczęto konstrukcję modelu w etapie następnym.

Podetap 5.3

W trzecim podetapie etapu piątego możliwe jest uwzględnienie części zmienności zależności pomiędzy natężeniem przyjazdów a zmiennymi objaśniającymi z poziomu powiatów przy pomocy charakterystyk z poziomu województw. Możliwe jest uwzględnienie takiego zróżnicowania zarówno w przypadku wszystkich zmiennych objaśniających z poziomu powiatów, niezależnie od tego, czy zostały wprowadzone do modelu w etapie trzecim.

$$Y_i = \sum_{q=0}^{18} \beta_{q,0,0}^{5(2/3)} X_{1,q,i} + \sum_{q=0}^{10} \beta_{0,q,0}^{5(2/3)} X_{2,q,i} + \sum_{q=1}^{18} e_{q(j)0}^{5(2/3)} X_{1,q,i} + \sum_{q=1}^{18} \sum_{s=0}^{10} \beta_{q,s,0}^{5(2/3)} X_{1,q,i} X_{2,s,i} + \sum_{q=1}^{18} \sum_{s=0}^{11} \beta_{q,0,s}^{5(2/3)} X_{1,q,i} X_{3,s,i} + \sum_{q=1}^{10} \sum_{s=0}^{11} \beta_{0,q,s}^{5(2/3)} X_{2,q,i} X_{3,s,i} + \sum_{q=1}^{18} e_{q,0(k)}^{5(2/3)} X_{1,q,i} + e_{0(j)0}^{5(2/3)} + e_{00(k)}^{5(2/3)} + r_i^{5(2/3)} \quad , \quad (4.15)$$

gdzie: $r_i^{5(2/3)} \sim N(0; 0,024614)$,

$$\begin{bmatrix} e_{0(j)0}^{5(2/3)} \\ e_{16(j)0}^{5(2/3)} \\ e_{10(j)0}^{5(2/3)} \\ e_{4(j)0}^{5(2/3)} \\ e_{14(j)0}^{5(2/3)} \\ e_{2(j)0}^{5(2/3)} \end{bmatrix} \sim N \left(\mathbf{0}; \begin{bmatrix} 0,0019 & -0,1880 & -0,0072 & 0,0071 & -0,0256 & 0,0417 \\ -0,1880 & 97147 & -52,201 & 519,62 & -1324,6 & -448,88 \\ -0,0072 & -52,201 & 0,1489 & -0,4871 & 0,7768 & -1,1919 \\ 0,0071 & 519,62 & -0,4871 & 3,1522 & -7,1252 & 0,2010 \\ -0,0256 & -1324,6 & 0,7768 & -7,1252 & 18,502 & 6,1168 \\ 0,0417 & -448,88 & -1,1919 & 0,2010 & 6,1168 & 20,51 \end{bmatrix} \right)$$

$$, \quad \begin{bmatrix} e_{00(k)}^{5(2/3)} \\ e_{13,0(k)}^{5(2/3)} \\ e_{7,0(k)}^{5(2/3)} \end{bmatrix} \sim N \left(\mathbf{0}; \begin{bmatrix} 0,0002 & 0,0689 & 0 \\ 0,0689 & 25,812 & 0 \\ 0 & 0 & 22,085 \end{bmatrix} \right) .$$

Tabela 4.31. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.3, poziom gmin, Polska, 2006 (wzór 4.15)

Parametr	$\beta_{0,0,0}^{5(2/3)}$	$\beta_{2,0,0}^{5(2/3)}$	$\beta_{1,0,0}^{5(2/3)}$	$\beta_{18,0,0}^{5(2/3)}$	$\beta_{10,0,0}^{5(2/3)}$	$\beta_{7,0,0}^{5(2/3)}$	$\beta_{3,0,0}^{5(2/3)}$	$\beta_{8,0,0}^{5(2/3)}$
Zmienna	-	$X_{1,2}$	$X_{1,1}$	$X_{1,18}$	$X_{1,10}$	$X_{1,7}$	$X_{1,3}$	$X_{1,8}$
Współczynnik regresji	0,3255	4,7810	-0,0008	-0,0001	0,0382	3,1400	-0,0092	0,8447
Błąd standardowy szacunku	0,0050	0,6772	0,0001	0,00001	0,0387	1,5980	0,0015	0,3271

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.32. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.3, poziom powiatów, Polska, 2006 (wzór 4.15)

Parametr	$\beta_{0,1,0}^{5(2/3)}$	$\beta_{0,9,0}^{5(2/3)}$
Zmienna	$X_{2,1}$	$X_{2,9}$
Współczynnik regresji	0,0074	0,2404
Błąd standardowy szacunku	0,0019	0,0922

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.33. Oceny współczynników kierunkowych interakcji między zmiennymi objaśniającymi z różnych poziomów oraz błędy standardowe szacunku, model natężenia przyjazdów do pracy, etap 5.3, Polska, 2006 (wzór 4.15)

Parametr	$\beta_{5,6,0}^{5(2/3)}$	$\beta_{3,9,0}^{5(2/3)}$	$\beta_{1,0,7}^{5(2/3)}$	$\beta_{0,9,3}^{5(2/3)}$	$\beta_{0,7,1}^{5(2/3)}$
Zmienna	$X_{1,5} \times X_{2,6}$	$X_{1,3} \times X_{2,9}$	$X_{1,1} \times X_{3,7}$	$X_{2,9} \times X_{3,3}$	$X_{2,7} \times X_{3,1}$
Współczynnik regresji	-2,8930	-0,0766	-0,1202	-0,0763	-0,0120
Błąd standardowy szacunku	1,0270	0,0204	0,0546	0,0220	0,0074

Źródło: Opracowanie własne na podstawie danych GUS

W bieżącym podetapie uwzględniono interakcje dla dwóch par zmiennych. W pierwszej kolejności stwierdzono wpływ poziomu ilości linii kolejowych w województwach na zależność pomiędzy natężeniem przyjazdów a udziałem osób pracujących w warunkach zagrożenia w powiatach. Mianowicie w województwach o lepiej rozwiniętej infrastrukturze kolejowej zależność pomiędzy natężeniem przyjazdów a udziałem osób pracujących w warunkach zagrożenia traciła na sile (por. tab. 4.32 i 4.33). Ponadto zaobserwowano wpływ udziału bezrobotnych w województwach na zależność pomiędzy natężeniem przyjazdów a długością dróg powiatowych w powiatach (por. tab. 4.33). Wpływ długości dróg powiatowych na natężenie dojazdów jest silniejszy, w przypadku gdy duża część osób dojeżdżających do pracy jako środek transportu wybiera samochód. Z kolei preferencje w kwestii środka transportu mogą różnić się istotnie pomiędzy województwami. W swojej pracy P.Rosik, M.Stępniak i R.Wiśniewski podają, że około połowy liczby dojeżdżających do pracy do Białegostoku przemieszczała się samochodem, podczas gdy 75% dojeżdżających do Warszawy wybiera ten środek transportu (por. Rosik, Stępniak, Wiśniewski, 2010).

Tabela 4.34. Wybrane kryteria oceny jakości dopasowania modelu trypoziomowego oszacowanego w etapie 5(2/3) (wzór (4.15) oraz jego porównanie z etapem 5(1/3) (wzór 4.14)

Etap	Kryterium oceny			vs etap 5.2	
	lnL	AIC	BIC	Chi2	p-value
etap 5.3	997,8573	-1913,715	-1666,588	13,421	0,0012

Źródło: Opracowanie własne na podstawie danych GUS

Uwzględnienie omówionych powyżej interakcji między zmiennymi z drugiego i trzeciego poziomu spowodowało istotną ze względu na wynik testu ilorazu wiarygodności poprawę dopasowania modelu. Ponieważ jest to ostatni etap konstrukcji modelu trzypoziomowego, model opisany równaniem (4.15) przyjęty został za obowiązujące i przy jego pomocy dokonane zostaną oszacowania natężenia przyjazdów w roku 2011.

4.2. Oszacowanie wartości poza diagonalą

OGÓLNE PODEJŚCIE ORAZ STRUKTURA POPULACJI

W celu pełnego oszacowania macierzy migracji należy jeszcze uzupełnić odpowiednie wartości poza jej diagonalą. Są to wartości codziennego przepływu związanego z zatrudnieniem dla każdej pary gmin z podziałem na część miejską i wiejską w Polsce, które dla roku 2006 zaczerpnięte zostały z badania przepływów ludności związanych z zatrudnieniem (Dojazdy do pracy w Polsce, red. Kruszka, 2010). Podobnie jak w przypadku oszacowań wartości na diagonalu macierzy migracji, dokonanych w podrozdziale 4.1, podjęta zostanie konstrukcja modelu dla roku 2006, w którym objaśnione zostaną wartości znanych dla tego roku przepływów związanych z zatrudnieniem. Następnie, przyjmując założenie, że w roku 2011 kształt zależności pomiędzy dojazdami oraz zmiennymi objaśniającymi nie uległ istotnej zmianie, skonstruowany model wykorzystany zostanie do oszacowania przepływów związanych z zatrudnieniem dla 2011 roku.

Ponieważ konstrukcja modelu ma na celu oszacowanie przepływów związanych z zatrudnieniem, naturalnym wydaje się, określenie jednostki pierwszego poziomu jako pary gmin. Należy przy tym podkreślić, że pary gmin, dla których przeprowadzono szacunki

są niesymetryczne, ponieważ określony jest kierunek przepływu. W związku z tym, badana populacja składa się z 3062×3061 jednostek bazowych, określonych jako uporządkowane pary gmin (każda para składa się z dwóch różnych gmin; liczba osób pracująca w gminie zamieszkania oszacowana została w podrozdziale 4.1). Ponieważ badana zbiorowość składa się z par jednostek podziału administracyjnego, charakteryzuje się strukturą wielopoziomową o charakterze krzyżowym. W związku z tym postanowiono podjąć próbę konstrukcji modelu wielopoziomowego z dwoma kryteriami grupowania. W modelu tym, jednostki poziomu pierwszego określono jako niesymetryczne pary gmin.

Natomiast na poziomie drugim określono dwa rodzaje jednostek, ze względu na gminę, z której pochodzi codzienny przepływ związany z zatrudnieniem oraz gminę, do której jest on skierowany. W dalszych rozważaniach stosowane będzie skrótowe określenie gminy wyjazdu oraz gminy przyjazdu. Zarówno dla gminy wyjazdu jak i przyjazdu, jako jednostki trzeciego oraz czwartego poziomu, określono powiaty i województwa, do których gminy te należą. W dalszych rozważaniach, jednostki te nazywane będą odpowiednio powiatem i województwem wyjazdu oraz powiatem i województwem przyjazdu. Kryterium ze względu na jednostkę administracyjną wyjazdu nazywane będzie dalej pierwszym kryterium grupowania i oznaczane we wzorach indeksem „a”, zaś ze względu na jednostkę administracyjną przyjazdu drugim kryterium grupowania i oznaczane indeksem „b”. Podejście to uważać można za analogiczne do modelu z jednym kryterium grupowania.

Podczas próby konstrukcji modelu uwzględniającego wszystkie pary gmin, niezależnie, czy w 2006 roku następował między nimi przepływ, napotkano poważne trudności obliczeniowe³¹. Wynikały one przede wszystkim z wielkości populacji, gdyż w pierwszym etapie podjęto konstrukcję modelu dla $3062 \cdot 3061 = 9\,372\,782$ jednostek bazowych. Dodatkową przyczyną trudności obliczeniowych, był z pewnością stopień komplikacji modelu. Z uwagi na brak możliwości oszacowania modelu uwzględniającego wszystkie pary gmin, postanowiono ograniczyć do tych par, pomiędzy którymi nastąpił przepływ ludności związany z zatrudnieniem w 2006 roku. Natomiast jako oszacowania przepływów które w 2006 roku były zerowe, również w roku 2011 przyjęto zera. Takie uproszczenie zagadnienia spowodowało zmniejszenie liczby jednostek bazowych do 25 375, co stanowi mniej niż jeden procent wszystkich uporządkowanych par gmin (około 0,27%). W szacunkach pozostawiono, wynikający ze struktury zbiorowości, poziom złożoności modelu.

Opisane powyżej podejście wymaga przyjęcia założenia, że jeżeli dla pary gmin, pomiędzy którymi nie odnotowano codziennego przepływu związanego z zatrudnieniem w roku 2006, również w macierzy migracji szacowanej dla roku 2011 w odpowiednim miejscu znajdzie się zero. Autorka świadoma jest niedoskonałości takiego podejścia. Przyjęte założenie może okazać się fałszywe na przykład, jeżeli w gminie wcześniej nieatrakcyjnej z punktu widzenia osób poszukujących pracy, otwarte zostałyby

³¹ Szerzej opisane w aneksie

przedsiębiorstwo zatrudniające dużo osób. Możliwe są też sytuacje, gdy w roku 2006 pomiędzy pewną parą gmin nie odnotowano przepływu związanego z zatrudnieniem, ponieważ nie przekraczał on 10 osób, zaś w roku 2011 liczba osób dojeżdżających do pracy nieznacznie wzrosła, w wyniku czego odpowiednia wartość nie powinna już być zerowa. Stać mogło się tak np. w wyniku przeprowadzek wielu mieszkańców miast na przedmieścia, przy jednoczesnym pozostaniu we wcześniejszym miejscu pracy. Niewielkie niedoszacowania nie powinny jednak wpłynąć istotnie na wynik przeprowadzonej w kolejnym rozdziale delimitacji lokalnych rynków pracy, ponieważ w procesie ich wyodrębniania kluczowe znaczenie mają tylko najsilniejsze przepływy.

Z drugiej strony z dużym prawdopodobieństwem przypuszczać można, że w gminie, w której codzienny przepływ związany z zatrudnieniem nie miał miejsca w 2006 r., nie nastąpi on również w latach późniejszych. Sytuacja taka zdeterminowana może być szczególną trudnością przejazdu pomiędzy gminami, nawet w przypadku małej odległości pomiędzy nimi. Trudność taka wynikać może między innymi z rzeźby geograficznej terenu, na przykład rzeki płynącej pomiędzy gminami, lub braku mostu, drogi czy łatwego połączenia środkami transportu publicznego. W związku z powyższym, przyjęcie założenia o braku przepływów w 2011 r., pomiędzy parami gmin w których nie następował on w roku 2006, choć nie jest w pełni słuszne z metodologicznego punktu widzenia, umożliwi uzyskanie pierwszego przybliżenia oszacowania macierzy przepływów przy zastosowaniu modelowania wielopoziomowego.

ZMIENNA OBJAŚNIANA

W następnej kolejności podjęto próbę zdefiniowania zmiennej objaśnianej w konstruowanym modelu. Podobnie jak w modelu dla diagonalii, celem postępowania było oszacowanie wartości bezwzględnych codziennych przepływów związanych z zatrudnieniem. Przyjmując, analogiczny sposób postępowania jak w podrozdziale 4.1, w pierwszej kolejności oszacowano wartości względne. Zróżnicowanie natężenia dojazdów do pracy zależy od stopnia atrakcyjności warunków zatrudnienia w gminie wyjazdu i przyjazdu. Z kolei wartości bezwzględne przepływów związanych z zatrudnieniem zdeterminowane mogą być również wielkością obu gmin. Natężenie dojazdów do pracy można traktować jako pewnego rodzaju miernik rozwoju gospodarczego regionów, zakładając, że jego poziom jest zróżnicowany dla jednostek podziału administracyjnego. Po zakończeniu oszacowań otrzymane wartości

przemnożono przez wartości mianownika, znanego również dla roku 2011. Ponieważ badana zmienna określona jest dla par gmin, rozważono pięć następujących sposobów zdefiniowania mianownika: liczba ludności pracującej w gminie przyjazdu, liczba ludności w wieku produkcyjnym w gminie wyjazdu, średnia arytmetyczna dwóch powyższych, liczba ludności pracującej w obu gminach łącznie oraz liczba osób w wieku produkcyjnym w obu gminach łącznie. Uznając, że najważniejsze jest silne skorelowanie zmiennych objaśniających ze zmienną objaśnianą, dające podstawę konstrukcji ‘najlepszego’ w świetle przyjętych kryteriów modelu, zbadano korelację z potencjalnymi zmiennymi objaśniającymi określonymi dla poziomu pierwszego (por. tab. 4.35). W przypadku większości potencjalnych zmiennych objaśniających, w szczególności tych najmocniej skorelowanych z potencjalnymi zmiennymi objaśnianymi, najsilniejsze korelacje zaobserwowano z charakterystyką przepływów określoną jako udział mieszkańców gminy wyjazdu wśród wszystkich pracujących w gminie przyjazdu. W dalszej analizie przyjęto zatem konstrukcję zmiennej objaśnianej, określonej jako natężenie dojazdów, przy uwzględnieniu jako wielkość odniesienia liczby wszystkich pracujących w gminie przyjazdu, analogicznie jak w podrozdziale 4.1.

ZMIENNE OBJAŚNIAJĄCE

Odległość między miejscem zamieszkania i miejscem pracy jest jednym z ważniejszych czynników kształtujących natężenie dojazdów do pracy (por. Gumuła, Socha, Wojciechowski, 2007). Tak więc odległość pomiędzy gminą wyjazdu a gminą przyjazdu, przyjęto jako zmienną objaśniającą z pierwszego poziomu. Natężenie dojazdów do pracy pomiędzy parami gmin jest zmienną o charakterze zbliżonym do natężenia przyjazdów do pracy szacowanego w podrozdziale 4.1. Z tego powodu zdecydowano, jako potencjalne zmienne objaśniające z poziomu gmin, powiatów i województw, zarówno wyjazdu jak i przyjazdu, przyjąć charakterystyki, na których oparto wcześniejszą konstrukcję modelu (podrozdział 4.1). Należy jednak zwrócić uwagę, że zmienne określone dla gmin w opisywanym aktualnie modelu są charakterystykami drugiego poziomu, dla powiatów poziomu trzeciego a dla województw poziomu czwartego (ze względu na obydwa kryteria grupowania).

Nieco odmienne podejście przyjęto podczas doboru zmiennych objaśniających w konstruowanym celem późniejszych porównań modelu klasycznej regresji liniowej. Mianowicie, oprócz odległości pomiędzy gminami jako zmienne objaśniające

wykorzystano w nim również różnice pomiędzy poziomem w gminie wyjazdu i przyjazdu wszystkich zmiennych objaśniających z poziomu gmin. Tak wyznaczone różnice w dużym stopniu określają relację sytuacji gospodarczej w obu gminach, a więc powinny być dobrymi charakterystykami przepływów związanych z zatrudnieniem. Wykorzystanie ich w modelu wielopoziomowym nie było jednak możliwe, ponieważ są one w oczywisty sposób, liniowo zależne ze zmiennymi z poziomu gmin, według obu kryteriów grupowania. Innymi słowy, różnice te nie wnoszą żadnej nowej informacji niż odjemna i odjemnik oddzielnie. Pomimo to, wykorzystano je również w celu wyznaczenia korelacji z potencjalnymi zmiennymi objaśnianymi, będącymi podstawą wyboru zmiennych objaśniających w konstruowanym modelu. W przeciwnym wypadku, ograniczając badanie korelacji do zmiennych objaśniających z pierwszego poziomu, uwzględniono by w nim wyłącznie odległość pomiędzy gminami.

Tabela 4.35. Współczynniki korelacji liniowej Pearsona między potencjalnymi charakterystykami natężenia dojazdów do pracy a zmiennymi objaśniającymi, poziom par gmin, Polska, 2006

	Nazwa zmiennej z poziomu pary gmin ¹	udział w pracujących w gminie przyjeździe	udział w ludności w wieku produkcyjnym w gminie wyjazdu	średnia arytmetyczna dwóch wcześniejszych	liczba ludności pracującej w obu gminach łącznie	liczba ludności w wieku produkcyjnym w obu gminach łącznie
X _{1,1}	Odległość pomiędzy gminami	-0,248	-0,174	-0,248	-0,237	-0,220
X _{1,2}	Różnica odległości od najbliższego z ośrodków centralnych	0,156	-0,003	0,120	0,120	0,111
X _{1,3}	Różnica liczby przedsiębiorstw z sekcji D zarejestrowanych w REGON na osobę w wieku produkcyjnym	-0,121	0,046	0,009	0,016	0,019
X _{1,4}	Różnica udziału bezrobotnych wśród ludności w wieku produkcyjnym	0,223	-0,044	0,108	0,129	0,102
X _{1,5}	Różnica liczby osób fizycznych prowadzących działalność gospodarczą na osobę w wieku produkcyjnym	-0,296	0,193	0,026	-0,001	0,015
X _{1,6}	Różnica liczby przedsiębiorstw sektora prywatnego na osobę w wieku produkcyjnym	-0,320	0,199	-0,006	-0,034	-0,008
X _{1,7}	Różnica liczby przedsiębiorstw zarejestrowanych w REGON na osobę w wieku produkcyjnym	-0,319	0,205	0,004	-0,024	0,002
X _{1,8}	Różnica liczby przedsiębiorstw z sekcji I zarejestrowanych w REGON na osobę w wieku produkcyjnym	-0,269	0,103	-0,086	-0,103	-0,084

	Nazwa zmiennej z poziomu pary gmin ¹	udział w pracujących w gminie przyjeździe	udział w ludności w wieku produkcyjnym w gminie wyjazdu	średnia arytmetyczna dwóch wcześniejszych	liczba ludności pracującej w obu gminach łącznie	liczba ludności w wieku produkcyjnym w obu gminach łącznie
X _{1,9}	Różnica liczby przedsiębiorstw z sekcji G zarejestrowanych w REGON na osobę w wieku produkcyjnym	-0,307	0,217	0,054	0,019	0,041
X _{1,10}	Różnica liczby przedsiębiorstw z sekcji F zarejestrowanych w REGON na osobę w wieku produkcyjnym	-0,188	0,074	-0,010	-0,021	-0,011
X _{1,11}	Różnica wskaźników zatrudnienia	-0,229	0,211	0,061	0,015	0,169
X _{1,12}	Różnica liczby spółek handlowych z sektora prywatnego na osobę w wieku produkcyjnym	-0,286	0,117	-0,136	-0,152	-0,113
X _{1,13}	Różnica odległości od najbliższej stolicy województwa	0,204	-0,012	0,162	0,163	0,148
X _{1,14}	Różnica liczby przedsiębiorstw z sekcji O zarejestrowanych w REGON na osobę w wieku produkcyjnym	-0,295	0,212	0,015	-0,001	0,033
X _{1,15}	Różnica liczby przedsiębiorstw z sekcji K zarejestrowanych w REGON na osobę w wieku produkcyjnym	-0,358	0,214	-0,055	-0,100	-0,065
X _{1,16}	Różnica liczby spółek handlowych z udziałem kapitału zagranicznego z sektora prywatnego na osobę w wieku produkcyjnym	-0,204	0,050	-0,129	-0,134	-0,105
X _{1,17}	Różnica liczby spółek handlowych z udziałem kapitału zagranicznego z sektora publicznego na osobę w wieku produkcyjnym	-0,185	0,205	0,082	0,041	0,140
X _{1,18}	Różnica w liczbie osób pracujących na km ²	-0,362	0,234	-0,072	-0,123	-0,074
X _{1,19}	Różnica w liczbie osób w wieku produkcyjnym na km ²	-0,375	0,245	-0,038	-0,094	-0,062

Uwaga: ¹ Wartości zmiennych X_{1,2} do X_{1,19} wyznaczane są jako różnica odpowiedniej wartości w gminie przyjeździe i wyjeździe

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.36. Współczynniki korelacji liniowej Pearsona między natężeniem dojazdów do pracy a potencjalnymi zmiennymi objaśniającymi, poziom gmin wyjazdu oraz gmin przyjeździe, 2006

Nazwa zmiennej	Gmina wyjazdu		Gmina przyjeździe	
	X _{a,2}	r _{xy}	X _{b,2}	r _{xy}
Odległość od najbliższego z ośrodków centralnych	X _{a,2,1}	0,036	X _{b,2,1}	0,138
Liczba przedsiębiorstw z sekcji D zarejestrowanych w REGON na osobę w wieku produkcyjnym	X _{a,2,2}	0,019	X _{b,2,2}	-0,122
Udział bezrobotnych wśród ludności w wieku produkcyjnym	X _{a,2,3}	0,018	X _{b,2,3}	0,256
Liczba osób fizycznych prowadzących działalność gospodarczą na osobę w wieku produkcyjnym	X _{a,2,4}	0,078	X _{b,2,4}	-0,298

Nazwa zmiennej	Gmina wyjazdu		Gmina przyjazdu	
	$X_{a,2}$	r_{xy}	$X_{b,2}$	r_{xy}
Liczba przedsiębiorstw z sektora prywatnego na osobę w wieku produkcyjnym	$X_{a,2,5}$	0,078	$X_{b,2,5}$	-0,324
Liczba przedsiębiorstw zarejestrowanych w REGON na osobę w wieku produkcyjnym	$X_{a,2,6}$	0,079	$X_{b,2,6}$	-0,325
Liczba przedsiębiorstw z sekcji I zarejestrowanych w REGON na osobę w wieku produkcyjnym	$X_{a,2,7}$	0,023	$X_{b,2,7}$	-0,295
Liczba przedsiębiorstw z sekcji G zarejestrowanych w REGON na osobę w wieku produkcyjnym	$X_{a,2,8}$	0,088	$X_{b,2,8}$	-0,314
Liczba przedsiębiorstw z sekcji F zarejestrowanych w REGON na osobę w wieku produkcyjnym	$X_{a,2,9}$	0,025	$X_{b,2,9}$	-0,198
Wskaźnik zatrudnienia	$X_{a,2,10}$	0,087	$X_{b,2,10}$	-0,216
Liczba spółek handlowych z sektora prywatnego na osobę w wieku produkcyjnym	$X_{a,2,11}$	0,035	$X_{b,2,11}$	-0,290
Odległość od najbliższej stolicy województwa	$X_{a,2,12}$	0,047	$X_{b,2,12}$	0,225
Liczba przedsiębiorstw z sekcji O zarejestrowanych w REGON na osobę w wieku produkcyjnym	$X_{a,2,13}$	0,080	$X_{b,2,13}$	-0,315
Liczba przedsiębiorstw z sekcji K zarejestrowanych w REGON na osobę w wieku produkcyjnym	$X_{a,2,14}$	0,088	$X_{b,2,14}$	-0,357
Liczba spółek handlowych z udziałem kapitału zagranicznego z sektora prywatnego na osobę w wieku produkcyjnym	$X_{a,2,15}$	0,005	$X_{b,2,15}$	-0,209
Liczba spółek handlowych z udziałem kapitału zagranicznego z sektora publicznego na osobę w wieku produkcyjnym	$X_{a,2,16}$	0,078	$X_{b,2,16}$	-0,176
Liczba osób pracujących na km ²	$X_{a,2,17}$	0,120	$X_{b,2,17}$	-0,357
Liczba osób w wieku produkcyjnym na km ²	$X_{a,2,18}$	0,121	$X_{b,2,18}$	-0,387

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.37. Współczynniki korelacji liniowej Pearsona między natężeniem dojazdów do pracy a potencjalnymi zmiennymi objaśniającymi, poziom powiatów wyjazdu oraz powiatów przyjazdu, 2006

Nazwa zmiennej	Powiat wyjazdu		Powiat przyjazdu	
	$X_{a,3}$	r_{xy}	$X_{b,3}$	r_{xy}
Udział bezrobotnych	$X_{a,3,1}$	0,028	$X_{b,3,1}$	0,235
Drogi gminne o nawierzchni twardej w kilometrach na km ²	$X_{a,3,2}$	0,009	$X_{b,3,2}$	-0,297
Drogi gminne łącznie w kilometrach na km ²	$X_{a,3,3}$	0,015	$X_{b,3,3}$	-0,309
Drogi gminne o nawierzchni twardej ulepszonej w kilometrach na km ²	$X_{a,3,4}$	0,009	$X_{b,3,4}$	-0,306
Drogi powiatowe o nawierzchni twardej ulepszonej w kilometrach na km ²	$X_{a,3,5}$	0,023	$X_{b,3,5}$	-0,278
Wskaźnik zatrudnienia	$X_{a,3,6}$	0,042	$X_{b,3,6}$	-0,317
Drogi powiatowe łącznie w kilometrach na km ²	$X_{a,3,7}$	0,024	$X_{b,3,7}$	-0,275
Drogi powiatowe o nawierzchni twardej w kilometrach na km ²	$X_{a,3,8}$	0,024	$X_{b,3,8}$	-0,274
Udział osób pracujących w warunkach zagrożenia wśród wszystkich pracujących	$X_{a,3,9}$	-0,012	$X_{b,3,9}$	-0,003
Średnie wynagrodzenie	$X_{a,3,10}$	-0,004	$X_{b,3,10}$	-0,249

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.38. Współczynniki korelacji liniowej Pearsona między natężeniem dojazdów do pracy a potencjalnymi zmiennymi objaśniającymi, poziom województw wyjazdu i województw przyjazdu, 2006

Nazwa zmiennej	Województwo wyjazdu		Województwo przyjazdu	
	$X_{a,4}$	r_{xy}	$X_{b,4}$	r_{xy}
Udział bezrobotnych	$X_{a,4,1}$	0,030	$X_{b,4,1}$	0,061
Wskaźnik zatrudnienia	$X_{a,4,2}$	-0,031	$X_{b,4,2}$	-0,094
Linie kolejowe normalnotorowe na 100 km ²	$X_{a,4,3}$	-0,063	$X_{b,4,3}$	-0,059
Nakłady na działalność innowacyjną w przedsiębiorstwach przemysłowych w tys. zł na osobę w wieku produkcyjnym. W takich samych jednostkach mierzone są nakłady na działalność innowacyjną łącznie	$X_{a,4,4}$	-0,051	$X_{b,4,4}$	-0,106
Miejsca w wozach na 1000 ludności miast obsługiwanych przez komunikację	$X_{a,4,5}$	0,007	$X_{b,4,5}$	-0,058
Ludność w miastach obsługiwanych przez zakłady komunikacji miejskiej w % ogółu ludności miast	$X_{a,4,6}$	-0,047	$X_{b,4,6}$	-0,066
Udział absolwentów publicznych uczelni wyższych wśród ludności w wieku produkcyjnym. Analogicznie wyznaczany jest udział absolwentów uczelni wyższych łącznie	$X_{a,4,7}$	0,022	$X_{b,4,7}$	0,018
Miejsca w wozach komunikacji miejskiej naziemnej na 1 osobę w wieku produkcyjnym	$X_{a,4,8}$	-0,034	$X_{b,4,8}$	-0,097
Nakłady na działalność innowacyjną łącznie	$X_{a,4,9}$	-0,030	$X_{b,4,9}$	-0,098
Udział absolwentów uczelni wyższych łącznie	$X_{a,4,10}$	0,010	$X_{b,4,10}$	-0,049
Średnie wynagrodzenie	$X_{a,4,11}$	-0,033	$X_{b,4,11}$	-0,102

Źródło: Opracowanie własne na podstawie danych GUS

Spośród wymienionych w tabelach 4.35 do 4.38 potencjalnych zmiennych objaśniających, dobór do modelu przeprowadzono przy pomocy regresji krokowej w przód, podczas kolejnych etapów konstrukcji modelu. W celu wstępnej weryfikacji doboru zmiennych objaśniających obliczono współczynniki korelacji liniowej Pearsona każdej z tych zmiennych ze zmienną objaśnianą. Podobnie jak w podrozdziale 4.1 zmienne określone dla poziomów wyższych niż pierwszy, czyli para gmin, sztucznie przekształcono na charakterystyki jednostek pierwszego poziomu.. W tym celu każdej parze gmin przyporządkowano wartości zmiennych określonych dla powiatów i województw przyjazdu oraz wyjazdu. Taka struktura zmiennych objaśniających, złożonych z wielokrotnych duplikatów spowodowała stosunkowo niskie wartości współczynników korelacji liniowej Pearsona, choć część z nich wносиła istotną informację o natężeniu dojazdów do pracy między parami gmin, co wykazano podczas konstrukcji modelu.

Warto zwrócić uwagę na fakt, że w przypadku większości zmiennych korelacja z natężeniem dojazdów do pracy pomiędzy parami gmin jest zdecydowanie silniejsza dla

charakterystyki jednostek terytorialnych do których dojazdy te są skierowane (por. tab. 4.36 do 4.38). Prawidłowość ta jest konsekwencją określeniem mianownika zmiennej objaśnianej. Mianowicie rozważany jest udział osób dojeżdżających do pracy z gminy wyjazdu do gminy przyjazdu wśród wszystkich zatrudnionych w gminie przyjazdu. Stąd naturalnym wydaje się, że tak określone natężenie dojazdów do pracy jest lepszą charakterystyką stopnia rozwoju gmin przyjazdu niż wyjazdu, ponieważ mianownik dotyczy jedynie gminy przyjazdu.

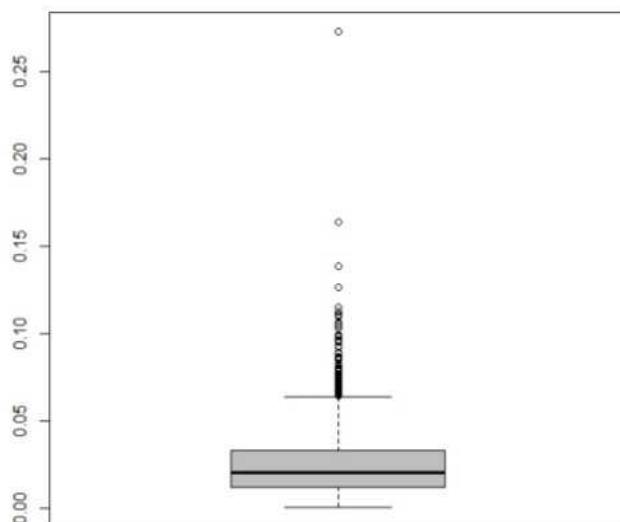
WERYFIKACJA ZAŁOŻEŃ

Konstrukcja modelu wielopoziomowego rozpoczęto od formalnej weryfikacji hipotezy głoszącej, że średni poziom natężenia dojazdów do pracy pomiędzy parami gmin różni się istotnie w zależności od przynależności do grup utworzonych przez jednostki wyższych poziomów. Jak zauważono wcześniej, szacowana zmienna charakteryzująca rynek pracy determinowana jest, między innymi, liczbą oferowanych miejsc pracy (w stosunku do liczby mieszkańców), wynikającą z poziomu rozwoju gospodarczego. W związku z tym istnieją przesłanki merytoryczne, aby przypuszczać, że średni poziom natężenia dojazdów do pracy różni się w przekroju jednostek podziału administracyjnego, w szczególności pomiędzy gminami, powiatami i województwami, zarówno ze względu na obszar z którego dojazd jest podejmowany jak i obszaru do którego jest skierowany.

W pierwszej kolejności przeanalizowano zróżnicowanie poziomu natężenia dojazdów do pracy w przekroju gmin wyjazdu (por. wykres 4.4) oraz przyjazdu (por. wykres 4.5). Sprawdzano więc, czy z jednych gmin wyjazdu do pracy podejmowane są częściej niż z innych oraz czy do niektórych gmin pracownicy dojeżdżają chętniej niż do innych. Na podstawie przeprowadzonych analiz stwierdzono, że średni poziom natężenia dojazdów do pracy różni się znacznie pomiędzy gminami, zarówno ze względu na pochodzenie jak i cel dojazdów.

Największe średnie natężenie przyjazdów do pracy odnotowano w mieście Witkowo z powiatu gnieźnieńskiego w Wielkopolsce. W dziesięciu gminach, do których dojeżdżali mieszkańcy Witkowa, stanowili oni przeciętnie ponad 27% zatrudnionych. Można przypuszczać, że duży wpływ na wartość względną przepływu związanego z zatrudnieniem pomiędzy parą gmin ma wielkość gminy wyjazdu, a także, ze względu na to, że w mianowniku tej zmiennej jest liczba osób pracujących w gminie przyjazdu, wielkość gminy przyjazdu. W większości gmin wyjazdu średni poziom przyjazdów

do pracy rozważanych z każdej gminy wyjazdu oddzielnie był znacznie niższy, w około połowie wypadków nie przekraczający 2%, zaś 578 gminach mniejszy niż 1%. Jedynie dla 14 gmin zaobserwowano przyjazdy do pracy przekraczające 10% osób pracujących w gminie przyjazdu.



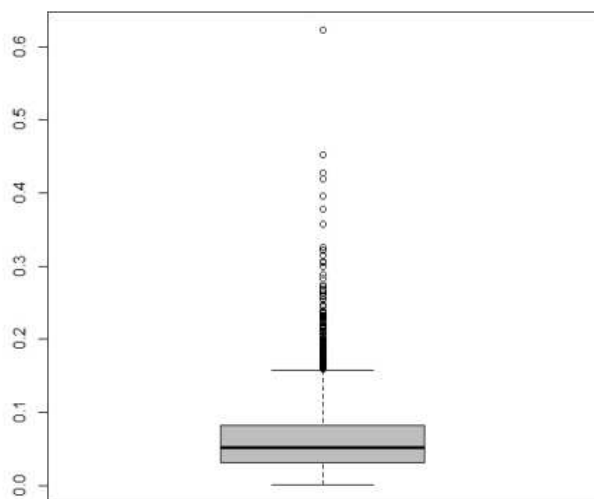
Wykres 4.4. Średnie natężenie przyjazdów do pracy pomiędzy parami gmin, przekrój gmin wyjazdu, Polska, 2006

Źródło: Opracowanie własne na podstawie danych GUS

Średni poziom natężenia wyjazdów do pracy z punktu widzenia gminy do której były one skierowane w dużym stopniu zależał od liczby gmin, z których był on skierowany. Mniejsza liczba gmin przekładała się oczywiście na większe natężenie dojazdów do pracy przypadające na jedną gminę wyjazdu, a więc również wyższą średnią. Najwyższe natężenie dojazdów do pracy, z punktu widzenia gminy przyjazdu, odnotowano w części wiejskiej gminy Pieniężno z powiatu braniewskiego położonego w województwie warmińsko-mazurskim. Do części wiejskiej tej gminy dojeżdżali tylko mieszkańcy jej części miejskiej i stanowili aż 62% zatrudnionych tam osób.

Największe średnie natężenie przyjazdów do pracy, ze względu na powiat wyjazdu, odnotowano w mieście Łomża, gminie na prawach powiatu z województwa podlaskiego. Przeciętnie przyjazdy do pracy do Łomży wynosiły nieco ponad 11% wszystkich osób pracujących w tym mieście. Ze względu na powiat przyjazdu największe średnie

natężenie wyjazdów do pracy odnotowano w powiecie suwalskim z województwa podlaskiego, wynosiło ono około 12% ludności pracującej w gminach do których dojeżdżali mieszkańcy Łomży.



Wykres 4.5. Średnie natężenie wyjazdów do pracy pomiędzy parami gmin, przekrój gmin przyjazdu, Polska, 2006

Źródło: Opracowanie własne na podstawie danych GUS

Należy podkreślić, że ze względu na przyjętą strukturę badania, pod uwagę brano tylko te pary gmin, pomiędzy którymi stwierdzono niezerowe przepływy związane z zatrudnieniem.

Statystyczną istotność omawianego zróżnicowania ze względu na gminę wyjazdu zweryfikowano przy pomocy testu analizy wariancji (por. Krzyśko, 1996). W tym celu sformułowano następujący układ hipotez:

$$\begin{cases} H_0: \mu_1 = \dots = \mu_{3048} \\ H_1: \sim H_0 \end{cases} \quad (4.16)$$

gdzie:

μ_j – średnia wartość natężenia dojazdów do pracy pochodzących z j -tej gminy, $j=1, \dots, 3048$.

Obliczono statystykę testową o następującej postaci:

$$F = \frac{\frac{\sum_{j=1}^J n_j (\bar{Y}_j - \bar{Y})^2}{J-1}}{\frac{\sum_{j=1}^J \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_j)^2}{n-J}} = \mathbf{1,0588} \quad (4.17)$$

gdzie:

\bar{Y}_j – średni poziom natężenia dojazdów w j -tej gminie wyjazdu, $j=1, \dots, 3048$,

$J = 3048$ – liczba gmin z których podejmowane były wyjazdy do pracy,

$n=25375$ – liczba par gmin o niezerowych dojazdach do pracy.

Na podstawie powyższej wartości statystyki testowej należy na poziomie istotności 0,05 (p-value=0,0173) odrzucić hipotezę zerową. Zatem można twierdzić, że występuje statystycznie istotne zróżnicowanie gmin ze względu na średnią wartość natężenia dojazdów do pracy, które z nich pochodzą.

W następnej kolejności statystycznej weryfikacji poddano hipotezę o zróżnicowaniu natężenia dojazdów do pracy pomiędzy gminami, do których są one skierowane.

$$\begin{cases} H_0: \mu_1 = \dots = \mu_{2843} \\ H_1: \sim H_0 \end{cases} \quad (4.18)$$

gdzie:

μ_j – średnia wartość natężenia dojazdów do pracy skierowanych do j -tej gminy, $j=1, \dots, 2843$.

Obliczono statystykę testową o następującej postaci:

$$F = \frac{\frac{\sum_{j=1}^J n_j (\bar{Y}_j - \bar{Y})^2}{J-1}}{\frac{\sum_{j=1}^J \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_j)^2}{n-J}} = \mathbf{5,9325} \quad (4.19)$$

gdzie:

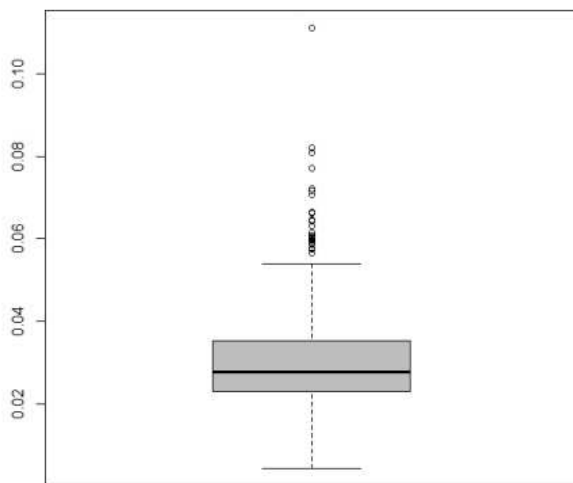
\bar{Y}_j – średni poziom natężenia dojazdów w j -tej gminie przyjazdu, $j=1, \dots, 2843$,

$J = 2843$ – liczba gmin z których podejmowane były wyjazdy do pracy,

$n=25375$ – liczba par gmin o niezerowych dojazdach do pracy.

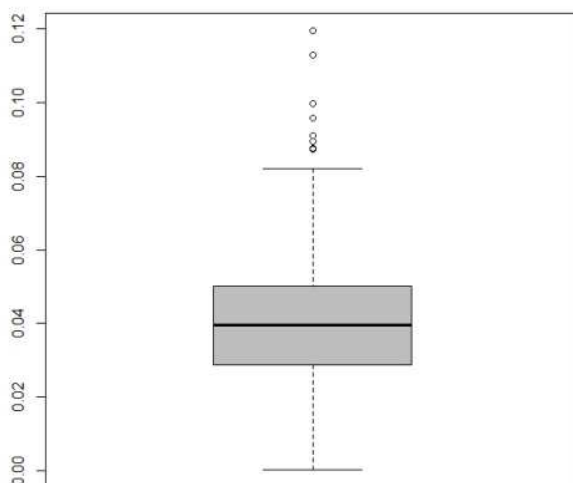
Na podstawie powyższej wartości statystyki testowej, należy na praktycznie dowolnym poziomie istotności (p-value <2,2e-16), odrzucić hipotezę zerową. Można zatem

twierdzić, że występuje statystycznie istotne zróżnicowanie gmin ze względu na średnią wartość natężenia dojazdów do pracy, do których są one skierowane.



Wykres 4.6. Średnie natężenie przyjazdów do pracy pomiędzy parami gmin, przekrój powiatów wyjazdu, Polska, 2006

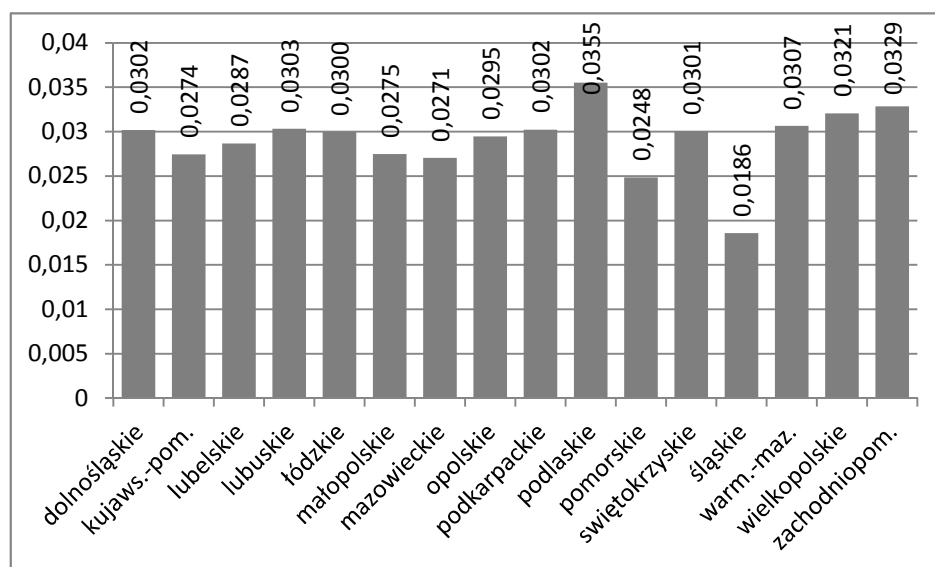
Źródło: Opracowanie własne na podstawie danych GUS



Wykres 4.7. Średnie natężenie wyjazdów do pracy pomiędzy parami gmin, przekrój powiatów przyjazdu, Polska, 2006

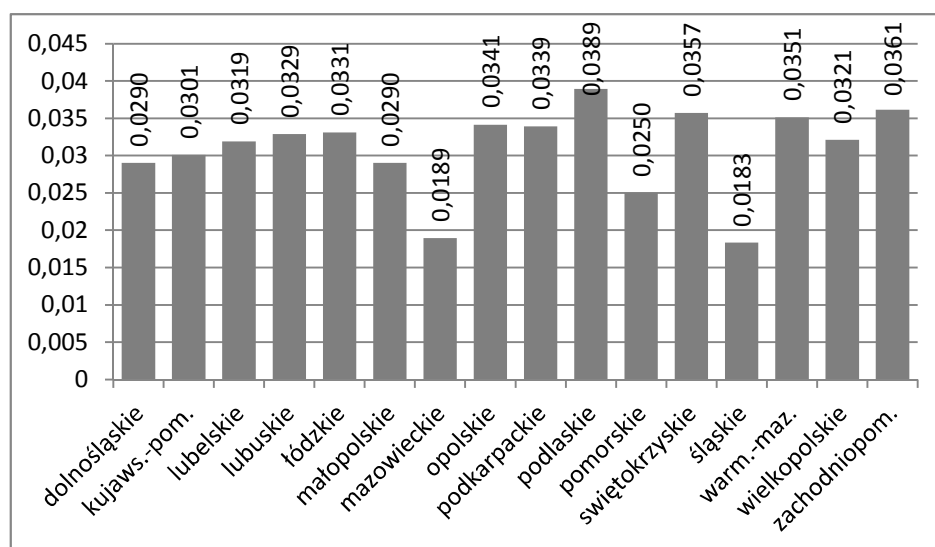
Źródło: Opracowanie własne na podstawie danych GUS

Analogiczne testy przeprowadzono dla weryfikacji zróżnicowania natężenia dojazdów do pracy pomiędzy parami gmin w przekroju powiatów wyjazdu i przyjazdu (por. odpowiednio wykresy 4.6 i 4.7) oraz województw wyjazdu i przyjazdu (por. odpowiednio wykresy 4.8 i 4.9 odpowiednio). Za każdym razem ich wyniki wskazywały na istotne statystycznie zróżnicowanie średniego poziomu badanej zmiennej (por. tab. 4.39).



Wykres 4.8. Średnie natężenie przyjazdów do pracy pomiędzy parami gmin, przekrój województw wyjazdu, Polska, 2006

Źródło: Opracowanie własne na podstawie danych GUS



Wykres 4.9. Średnie natężenie wyjazdów do pracy pomiędzy parami gmin, przekrój województw przyjazdu, Polska, 2006

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.39. Parametry testów analizy wariancji weryfikujących zróżnicowanie natężenia dojazdów do pracy we wszystkich rozważanych przekrojach terytorialnych, Polska, 2006

Parametr testów	Kryterium grupowania					
	Gmina wyjazdu	Gmina przyjazdu	Powiat wyjazdu	Powiat przyjazdu	Woje-wództwo wyjazdu	Woje-wództwo przyjazdu
Liczba wariantów kryterium grupowania	3048	2843	379	379	16	16
F	1,059	5,933	2,618	12,779	10,038	24,577
p-value	0,0173	<2,2e-16	<2,2e-16	<2,2e-16	<2,2e-16	<2,2e-16

Źródło: Opracowanie własne na podstawie danych GUS

Warto zauważyć, że silniejsze zróżnicowanie natężenia wyjazdów do pracy w przekroju jednostek administracyjnych, do których dojazd ten jest skierowany, niż w przekroju jednostek z których pochodzi (por. tab. 4.39) zdeterminowane jest określeniem mianownika natężenia dojazdów do pracy. Mianowicie rozważany jest udział osób dojeżdżających do pracy z gminy wyjazdu do gminy przyjazdu wśród wszystkich zatrudnionych w gminie przyjazdu. W związku z taką definicją zmiennej naturalnym wydaje się jej silniejsze zróżnicowanie w przekroju jednostek administracyjnych przyjazdu niż wyjazdu, ponieważ w drugim przypadku na poziom tej zmiennej duży wpływ może mieć wielkość gminy wyjazdu obok stopnia jej rozwoju.

KONSTRUKCJA MODELU

W konstrukcji krzyżowego modelu czteropoziomowego uwzględniono wszystkie etapy komplikacji modelu opisane w podrozdziale 3.3 wraz z weryfikacją istotności poprawy dopasowania modelu po każdym kroku. Obliczenia wykonano przy pomocy kodu autorskiego napisanego w programie R 2.15.2 z wykorzystaniem biblioteki lme4³² (por. Bates, 2010; Bates, 2013a; Bates, 2013b; Biecek, 2012). Parametry oszacowano zgodnie z metodą największej wiarygodności. Przed rozpoczęciem obliczeń wartości wszystkich zmiennych objaśniających zostały scentrowane, poprzez odjęcie średniej globalnej.

KLASYCZNA REGRESJA LINIOWA

W pierwszej kolejności wyznaczono dwie funkcje klasycznej regresji liniowej, które następnie wykorzystano w celu porównania z wynikami modelu wielopoziomowego.

³² Linear mixed-effects models using S4 classes

W pierwszej z nich (w dalszym ciągu podrozdziału indeksowanej jako etap 0.0) nie uwzględniono żadnych zmiennych objaśniających:

$$Y_i = \beta_{0,0,0}^{00} + r_i^{00}, \quad r_i^{00} \sim N(0; 0,05317), \quad (4.20)$$

Szacowanie natężenia przyjazdów do pracy przy pomocy powyższego modelu sprowadza się do przypisania każdej parze gmin średniej wartości natężenia dojazdów do pracy. Model ten wyznaczono tylko i wyłącznie w celu porównania w późniejszych etapach, jest on, podobnie jak w przypadku szacowania wartości na diagonalu macierzy migracji, bardzo słabej jakości (por. tab. 4.41). Z oszacowanych parametrów wynika, że średnie natężenie przyjazdów do pracy w gminach w 2006 roku wynosiło około 3% (por. tab.4.40).

Tabela 4.40. Oceny parametru strukturalnego oraz błędu standardowego szacunku, klasyczny model regresji liniowej natężenia dojazdów do pracy pomiędzy parami gmin bez zmiennych objaśniających, Polska, 2006 (wzór 4.20)

Parametr	$\beta_{0,0,0}^{00}$
Współczynnik regresji	0,0281782
Błąd standardowy szacunku	0,0003338

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.41. Wybrane kryteria oceny jakości dopasowania klasycznego modelu regresji liniowej natężenia dojazdów do pracy bez zmiennych objaśniających (wzór 4.5)

Kryteria dopasowania		
lnL	AIC	BIC
38449,97	-76895,93	-76879,65

Źródło: Opracowanie własne na podstawie danych GUS

W drugiej funkcji regresji (wzór 4.21, w dalszej części podrozdziału indeksowanej jako etap 0.1) przyjęto zmienne objaśniające określone dla par gmin. Wśród nich uwzględniono nie tylko odległość pomiędzy parami gmin, ale także wszystkie różnice pomiędzy charakterystykami gmin wyjazdu i przyjazdu (por. tab. 4.38). Dobór zmiennych objaśniających przeprowadzono metodą regresji krokowej w przód. Zmienne objaśniające każdorazowo szeregowano w taki sposób, aby maksymalizować wiarygodność modelu. Zmienne kolejno dołączano do modelu, do momentu kiedy test ilorazu wiarygodności wskazywał na istotną (na poziomie istotności 0,05) poprawę jakości. Jako model wyjściowy przyjęto model bez zmiennych objaśniających opisany powyżej.

$$Y_i = \sum_{q=0}^{19} \beta_q^{01} X_{1,q,i} + r_i^{01} \quad , \quad r_i^{01} \sim N(0; 0,04784) \quad , \quad (4.21)$$

Tabela 4.42. Oceny parametrów strukturalnych oraz błędów standardowych szacunku klasycznego modelu regresji liniowej natężenia dojazdów do pracy pomiędzy parami gmin ze zmiennymi objaśniającymi, Polska, 2006 (wzór 4.21)

Parametr	$\beta_{0,0,0,0}^{01}$	$\beta_{19,0,0,0}^{01}$	$\beta_{1,0,0,0}^{01}$	$\beta_{11,0,0,0}^{01}$	$\beta_{9,0,0,0}^{01}$	$\beta_{16,0,0,0}^{01}$	$\beta_{15,0,0,0}^{01}$
Zmienna	-	$X_{1,19}$	$X_{1,1}$	$X_{1,11}$	$X_{1,9}$	$X_{1,16}$	$X_{1,15}$
Współczynnik regresji	0,028178	-0,000018	-0,000116	-0,01532	-0,4824	0,5836	-1,012
Błąd standardowy szacunku	0,0003003	0,000001	0,000005	0,000899	0,03545	0,1378	0,05115

Parametr	$\beta_{7,0,0,0}^{01}$	$\beta_{5,0,0,0}^{01}$	$\beta_{4,0,0,0}^{01}$	$\beta_{18,0,0,0}^{01}$	$\beta_{6,0,0,0}^{01}$	$\beta_{2,0,0,0}^{01}$
Zmienna	$X_{1,7}$	$X_{1,5}$	$X_{1,4}$	$X_{1,18}$	$X_{1,6}$	$X_{1,2}$
Współczynnik regresji	1,234	-0,391	0,000536	0,000018	-0,6765	-0,000051
Błąd standardowy szacunku	0,08089	0,05861	0,000086	0,000003	0,09126	0,000012

Źródło: Opracowanie własne na podstawie danych GUS

W tabeli 4.42 podano wartości współczynników regresji przy zmiennych objaśniających, które jako istotne zostały wprowadzone do modelu. Wartości pozostałych współczynników traktować można jako zera. Ponadto kolejność zmiennych objaśniających w tabeli jest konsekwencją kolejności dołączania ich do modelu w rezultacie zastosowania procedury regresji krokowej w przód.

Jako pierwsza zmienna objaśniająca do modelu wprowadzona został różnica pomiędzy gęstością zaludnienia w gminie przyjazdu i wyjazdu ($X_{1,19}$). Jest ona destymulantą natężenia dojazdów do pracy, ponieważ duża liczba mieszkańców w wieku produkcyjnym na kilometr kwadratowy w gminie przyjazdu, wiąże się z nasyceniem jej lokalnego rynku pracy. Sytuacja dużej gęstości ‘zaludnienia’ stanowi czynnik zniechęcający dla mieszkańców gminy wyjazdu, zwłaszcza, jeżeli tamtejszy lokalny rynek pracy jest mniej ‘nasycony’ ludnością w wieku produkcyjnym. Kolejną destymulantą natężenia dojazdów do pracy jest odległość pomiędzy gminami ($X_{1,1}$). Z oczywistych powodów, przy stałych warunkach zatrudnienia, pracownicy wolą pracować jak najbliżej miejsca zamieszkania. Rolę destymulant natężenia dojazdów do pracy pełni również różnica wskaźników zatrudnienia ($X_{1,11}$) oraz różnica natężenia

przedsiębiorstw z sekcji G (handel hurtowy i detaliczny; naprawa pojazdów samochodowych, włączając motocykle), zmienna ($X_{1,9}$). Z kolei różnica w natężeniu spółek handlowych z udziałem kapitału zagranicznego z sektora prywatnego ($X_{1,16}$) jest stymulantą natężenia dojazdów do pracy. Im bardziej natężenie takich spółek w gminie przyjazdu przewyższa ich natężenie w gminie wyjazdu, tym więcej osób skłonnych jest dojeżdżać do pracy. Spowodowane jest to prawdopodobnie tym, że spółki handlowe z udziałem kapitału zagranicznego z sektora prywatnego zatrudniają często wykwalifikowanych fachowców na wyspecjalizowanych stanowiskach, które trudno znaleźć w najbliższej okolicy. Z drugiej strony, również potencjalni pracownicy skłonni są podjąć dojazdy do pracy aby zostać zatrudnionym na bardziej atrakcyjnych stanowiskach. Jednak już większe natężenie przedsiębiorstw z sekcji K (Obsługa nieruchomości, wynajem i usługi związane z prowadzeniem działalności gospodarczej) ($X_{1,15}$) w gminie przyjazdu niż wyjazdu wiąże się ze spadkiem natężenia dojazdów.

Tabela 4.43. Wybrane kryteria oceny jakości dopasowania klasycznego modelu regresji liniowej natężenia przyjazdów do pracy ze zmiennymi objaśniającymi (wzór 4.6) oraz jego porównanie z etapem 0.0 (wzór 4.5)

Etap	Kryterium oceny			vs etap 00	
	lnL	AIC	BIC	Chi2	p-value
Etap 01	41138,75	-82249,5	-82135,5	498,8988	0,0000

Źródło: Opracowanie własne na podstawie danych GUS

Uwzględnienie zmiennych objaśniających w znacznym stopniu poprawiło jakość dopasowania modelu (por. tab. 4.43).

ETAP 1 – Model bez zmiennych objaśniających

Konstrukcję modelu rozpoczęto od uwzględnienia w podetapie 1.1 zróżnicowania poziomu natężenia dojazdów do pracy pomiędzy parami gmin w przekroju gmin wyjazdu. W podetapie 1.2 uwzględniono również zróżnicowanie poziomu zmiennej objaśnianej w przekroju gmin przyjazdu. Następnie, w podetapie 1.3 podjęto próbę uwzględnienia zróżnicowania w przekroju powiatów wyjazdu. Jednak próba ta nie spowodowała istotnej poprawy jakości dopasowania modelu. Zatem efekt losowy z poziomu powiatów wyjazdu nie został uwzględniony w modelu. Istotną poprawę spowodowało natomiast wprowadzenie do modelu, w podetapie 1.4, składnika losowego z poziomu powiatów przyjazdu. W ramach podetapu 1.5 podjęto próbę rozszerzenia modelu o składnik losowy z poziomu województw wyjazdu. Jednak podobnie

jak w przypadku powiatów wyjazdu, nie spowodowało to statystycznie istotnej poprawy. W ostatnim podetapie (1.6) etapu pierwszego, uwzględniono zróżnicowanie natężenia dojazdów do pracy między parami gmin w przekroju województw przyjazdu.

$$Y_i = \beta_{0000}^1 + \sum_{p=2}^4 e_{a0\dots0(k)0\dots0}^1 + \sum_{p=2}^4 e_{b0\dots0(k)0\dots0}^1 + r_i^1, \quad (4.22)$$

gdzie: $r_i^1 \sim N(0; \sigma_1^2)$, $e_{a0\dots0(k)0\dots0}^1 \sim N(0; \tau_{a,p,0,0,0,0}^1)$, dla $p=2, \dots, 4$,
 $e_{b0\dots0(k)0\dots0}^1 \sim N(0; \tau_{b,p,0,0,0,0}^1)$, dla $p=2, \dots, 4$.

Tabela 4.44. Oceny parametru strukturalnego oraz błędu standardowego szacunku modeli natężenia przyjazdów do pracy, podetapy etapu pierwszego, Polska, 2006 (wzór 4.22)

Parametr:	Etap 1.1	Etap 1.2	Etap 1.4	Etap 1.6
$\beta_{0,0,0,0}^1$	0,02736	0,05669	0,05602	0,05693
$S(\beta_{0,0,0,0}^1)$	0,00038	0,00082	0,00105	0,00219

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.45. Oceny wariancji składników losowych modeli natężenia przyjazdów do pracy, podetapy etapu pierwszego, Polska, 2006 (wzór 4.22)

Parametr:	Etap 1.1	Etap 1.2	Etap 1.4	Etap 1.6
σ_1^2	0,002736	0,001818	0,001817	0,001817
$\tau_{a,2,0,0,0,0}^1$	0,000083	0,000032	0,000030	0,000029
$\tau_{b,2,0,0,0,0}^1$	-	0,001317	0,001187	0,001188
$\tau_{b,3,0,0,0,0}^1$	-	-	0,000152	0,000106
$\tau_{b,4,0,0,0,0}^1$	-	-	-	0,000060

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.46 Wybrane kryteria oceny jakości dopasowania modeli natężenia przyjazdów do pracy oszacowanych w etapie pierwszym (wzór 4.22) oraz ich porównanie z wcześniejszymi podetapami

Kryterium oceny	Etap 1.1	Etap 1.2	Etap 1.4	Etap 1.6
LnL	38538,84	41755,12	41780,93	41794,9
AIC	-77071,7	-83502,2	-83551,87	-83577,8
BIC	-77047,3	-83469,7	-83511,16	-83528,95
Vs	Etap 0.0	Etap 1.1	Etap 1.2	Etap 1.4
Chi2	177,74	6432,565	51,621	27,936
p-value	0,0000000	0,0000000	0,0000000	0,0000001

Źródło: Opracowanie własne na podstawie danych GUS

Na podstawie wyznaczonych kryteriów oceny dopasowania modelu, przede wszystkim testu ilorazu wiarygodności, stwierdzono systematyczną poprawę dopasowania wyznaczonego modelu w porównaniu z podetapem poprzednim. Jedynie porównanie podetapów 1.4 z 1.6 przy pomocy kryterium Schwarza nie wskazuje na zasadność

rozszerzenia modelu. Kierując się jednak kryterium Akaike'a oraz testem ilorazu wiarygodności uzyskaną poprawę należy uznać za statystycznie istotną na praktycznie dowolnym poziomie istotności (większym niż 0,0000001).

ETAP 2 – Model z losowym wyrazem wolnym

W etapie drugim do modelu wprowadzana została jedyna zmienna objaśniająca określona dla par gmin, czyli odległość pomiędzy gminą wyjazdu i przyjazdu.

$$Y_i = \beta_{0000}^2 + \beta_{1000}^2 X_{1,1,i} + \sum_{p=2}^4 e_{a0\dots0(k)0\dots0}^2 + \sum_{p=2}^4 e_{b0\dots0(k)0\dots0}^2 + r_i^2, \quad (4.23)$$

gdzie: $r_i^2 \sim N(0; \sigma_2^2)$, $e_{a0\dots0(k)0\dots0}^2 \sim N(0; \tau_{a,p,0,0,0,0}^2)$, dla $p=2, \dots, 4$,

$e_{b0\dots0(k)0\dots0}^2 \sim N(0; \tau_{b,p,0,0,0,0}^2)$, dla $p=2, \dots, 4$.

Tabela 4.47. Oceny parametrów strukturalnych oraz błędów standardowych szacunku modelu natężenia przyjazdów do pracy, etap drugi, Polska, 2006 (wzór 4.23)

Parametr:	Etap 2
$\beta_{0,0,0,0}^2$	0,05786
$S(\beta_{0,0,0,0}^2)$	0,002146
$\beta_{1,0,0,0}^2$	-0,000072
$S(\beta_{1,0,0,0}^2)$	0,000005

Źródło: Opracowanie własne na podstawie danych GUS

Odległość pomiędzy gminami okazała się istotną zmienną objaśniającą natężenie dojazdów do pracy między badanymi gminami. Zwiększenie odległości między gminami związane jest ze spadkiem natężenia dojazdów. W tym miejscu warto przypomnieć, że rozważane są tylko te pary gmin, pomiędzy którymi w 2006 roku nastąpił przepływ związany z zatrudnieniem, a więc położone w odległości pozwalającej na codzienny dojazd do pracy. Gdyby uwzględnione zostały wszystkie możliwe pary gmin, zależność pomiędzy odległością między gminami a natężeniem dojazdów do pracy byłaby niewątpliwie znacznie silniejsza. Z drugiej strony, duża liczba zerowych wartości zmiennej objaśnianej dla par gmin położonych w znacznej odległości od siebie, prawdopodobnie spowodowałaby przeszacowanie natężenia dojazdów do pracy pomiędzy parami gmin położonymi blisko siebie.

Tabela 4.48. Oceny wariancji składników losowych natężenia przyjazdów do pracy, etap drugi, Polska, 2006 (wzór 4.23)

Parametr:	Etap 2
σ_2^2	0,001797
$\tau_{a,2,0,0,0,0}^2$	0,000040
$\tau_{b,2,0,0,0,0}^2$	0,001155
$\tau_{b,3,0,0,0,0}^2$	0,000098
$\tau_{b,4,0,0,0,0}^2$	0,000057

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.49. Wybrane kryteria oceny jakości dopasowania modelu natężenia przyjazdów do pracy oszacowanego w etapie 2 (wzór 4.23) oraz ich porównanie z etapem 1.6

Kryterium oceny	Etap 2
lnL	41898,08
AIC	-83782,2
BIC	-83725,2
vs	Etap 1.6
Chi2	206,3603
p-value	0,000000

Źródło: Opracowanie własne na podstawie danych GUS

Na podstawie testu ilorazu wiarygodności stwierdzono statystycznie istotną poprawę jakości dopasowania modelu w stosunku do ostatniego podetapu etapu pierwszego (por. tab. 4.46 i 4.49).

Warto zauważyć, że zestaw zmiennych objaśniających z poziomu pierwszego, w przypadku oszacowań przeprowadzonych przy pomocy klasycznej regresji liniowej był znacznie szerszy. W konstruowanym modelu wielopoziomowym odpowiadające mu charakterystyki wprowadzone zostaną dopiero w dwóch pierwszych podetapach etapu trzeciego. W związku z tym, modelu z etapu 2 nie należy porównywać z klasyczną regresją liniową.

ETAP 3 – Model z losowym wyrazem wolnym zależnym od zmiennych z wyższych poziomów

W kolejnych podetapach etapu trzeciego, do modelu wprowadzone mogą zostać zmienne objaśniające z poziomów wyższych niż pierwszy. Tak więc, w pierwszej kolejności wprowadzane są charakterystyki określone dla gmin wyjazdu oraz przyjazdu (odpowiednio podetap 3.1 i 3.2), następnie zmienne określone dla powiatów wyjazdu oraz przyjazdu (odpowiednio podetap 3.3 i 3.4) a na końcu charakterystyki województw wyjazdu i przyjazdu (odpowiednio podetap 3.5 i 3.6).

$$Y_i = \beta_{0000}^3 + \beta_{1000}^3 X_{1,1,i} + \sum_{p=2}^4 \sum_{q=1}^{Q_{a,p}} \beta_{\underbrace{a_0 \dots 0 q 0 \dots 0}_{1 \dots p \dots 4}}^3 X_{a,p,q,i} + \sum_{p=2}^{P_b} \sum_{q=1}^{Q_{b,p}} \beta_{\underbrace{b_0 \dots 0 q 0 \dots 0}_{1 \dots p \dots 4}}^3 X_{b,p,q,i} + \sum_{p=2}^4 e_{\underbrace{a_0 \dots 0 (k) 0 \dots 0}_{1 \dots p \dots 4}}^3 + \sum_{p=2}^4 e_{\underbrace{b_0 \dots 0 (k) 0 \dots 0}_{1 \dots p \dots 4}}^3 + r_i^3, \quad (4.24)$$

gdzie: $Q_{a,2}=Q_{b,2}=18$, $Q_{a,3}=Q_{b,3}=10$, $Q_{a,4}=Q_{b,4}=11$, $r_i^3 \sim N(0; \sigma_3^2)$,
 $e_{\underbrace{a_0 \dots 0 (k) 0 \dots 0}_{1 \dots p \dots 4}}^3 \sim N(0; \tau_{a,p,0,0,0}^3)$, dla $p=2, \dots, 4$, $e_{\underbrace{b_0 \dots 0 (k) 0 \dots 0}_{1 \dots p \dots 4}}^3 \sim N(0; \tau_{b,p,0,0,0}^3)$,
dla $p=2, \dots, 4$.

Tabela 4.50. Oceny parametrów strukturalnych oraz błędów standardowych szacunku modelu natężenia przyjazdów do pracy, podetapy etapu trzeciego, Polska, 2006 (wzór 4.24)

Parametr:	Etap 3.1	Etap 3.2	Etap 3.3	Etap 3.4	Etap 3.5
$\beta_{0,0,0,0}^2$	0,052340	0,054480	0,051310	0,054400	0,051190
$S(\beta_{0,0,0,0}^2)$	0,002227	0,003101	0,003229	0,003436	0,004676
$\beta_{1,0,0,0}^2$	-0,000082	-0,000077	-0,000079	-0,000078	-0,000081
$S(\beta_{1,0,0,0}^2)$	0,000005	0,000005	0,000005	0,000005	0,000005
$\beta_{a,0,18,0,0}^2$	0,000012	0,000012	0,000011	0,000011	0,000011
$S(\beta_{a,0,18,0,0}^2)$	0,000001	0,000001	0,000001	0,000001	0,000001
$\beta_{a,0,16,0,0}^2$	4,515000	4,041000	4,389000	4,398000	4,121000
$S(\beta_{a,0,16,0,0}^2)$	1,677000	1,670000	1,672000	1,671000	1,673000
$\beta_{a,0,10,0,0}^2$	0,005035	0,004693	0,002529	0,002510	0,002548
$S(\beta_{a,0,10,0,0}^2)$	0,001971	0,001968	0,002046	0,002046	0,002046
$\beta_{b,0,17,0,0}^2$	-	0,000004	0,000004	0,000006	0,000006
$S(\beta_{b,0,17,0,0}^2)$	-	0,000011	0,000011	0,000011	0,000011
$\beta_{b,0,10,0,0}^2$	-	-0,032310	-0,032530	-0,032980	-0,032860
$S(\beta_{b,0,10,0,0}^2)$	-	0,003372	0,003379	0,003383	0,003381
$\beta_{b,0,14,0,0}^2$	-	-0,554000	-0,558100	-0,547900	-0,555900
$S(\beta_{b,0,14,0,0}^2)$	-	0,079490	0,079680	0,079180	0,079320
$\beta_{b,0,3,0,0}^2$	-	0,001095	0,001130	0,001056	0,001050
$S(\beta_{b,0,3,0,0}^2)$	-	0,000195	0,000196	0,000197	0,000197
$\beta_{b,0,18,0,0}^2$	-	-0,000023	-0,000023	-0,000022	-0,000022
$S(\beta_{b,0,18,0,0}^2)$	-	0,000004	0,000004	0,000004	0,000004
$\beta_{b,0,12,0,0}^2$	-	0,000110	0,000111	0,000109	0,000108
$S(\beta_{b,0,12,0,0}^2)$	-	0,000030	0,000030	0,000030	0,000030
$\beta_{a,0,0,6,0}^2$	-	-	0,014040	0,013930	0,014530
$S(\beta_{a,0,0,6,0}^2)$	-	-	0,003650	0,003649	0,003654
$\beta_{b,0,0,3,0}^2$	-	-	-	-0,002405	-0,002231
$S(\beta_{b,0,0,3,0}^2)$	-	-	-	0,001045	0,001049
$\beta_{a,0,0,0,8}^2$	-	-	-	-	-51,680000
$S(\beta_{a,0,0,0,8}^2)$	-	-	-	-	18,690000
$\beta_{a,0,0,0,7}^2$	-	-	-	-	0,605400
$S(\beta_{a,0,0,0,7}^2)$	-	-	-	-	0,279900

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.51. Oceny wariancji składników losowych modelu natężenia przyjazdów do pracy, podetapy etapu trzeciego, Polska, 2006 (wzór 4.24)

Parametr:	Etap 3.1	Etap 3.2	Etap 3.3	Etap 3.4	Etap 3.5
σ_z^2	0,001797	0,001813	0,001812	0,001812	0,001811
$\tau_{a,2,0,0,0,0}^2$	0,00000077	0	0,00000000000003	0	0
$\tau_{b,2,0,0,0,0}^2$	0,001112	0,000684	0,000687	0,000687	0,000686
$\tau_{b,3,0,0,0,0}^2$	0,000116	0,000029	0,000030	0,000030	0,000029
$\tau_{b,4,0,0,0,0}^2$	0,000061	0,000030	0,000031	0,000021	0,000024

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.52. Wybrane kryteria oceny jakości dopasowania modelu natężenia przyjazdów do pracy oszacowanego w etapie 3 (wzór 4.24) oraz ich porównanie z etapem 2

Kryterium oceny	Etap 3.1	Etap 3.2	Etap 3.3	Etap 3.4	v 3.5
lnL	42150,65	42532,04	42539,42	42541,80	42547,78
AIC	-84281,30	-85032,08	-85044,84	-85047,60	-85055,56
BIC	-84199,88	-84901,81	-84906,43	-84901,05	-84892,73
vs	Etap 2	Etap 3.1	Etap 3.2	Etap 3.3	Etap 3.4
Chi2	505,14	762,78	14,76	4,76	11,96
p-value	0,000000	0,000000	0,000122	0,029125	0,002531

Źródło: Opracowanie własne na podstawie danych GUS

Na podstawie testu ilorazu wiarygodności stwierdzono istotną na praktycznie dowolnym poziomie istotności poprawę jakości dopasowania modelu skonstruowanego w podetapie 3.1 w stosunku do modelu z drugiego etapu. Również wprowadzenie do modelu rozszerzeń opisywanych w podetapach 3.2, 3.3 i 3.5 spowodowało poprawę zarówno obydwóch kryteriów informacyjnych jak i statystycznie istotną poprawę według testu ilorazu wiarygodności. Również wprowadzenie charakterystyki określonej dla powiatów przyjazdu spowodowało poprawę dopasowania modelu według testu ilorazu wiarygodności, jednak tylko na poziomie ufności przekraczającym 0,029125 (por. tab. 4.52).

ETAP 4 – Model z losowymi współczynnikami regresji

W podetapach etapu czwartego dopuszcza się uwzględnienie zróżnicowania współczynników kierunkowych przy zmiennych objaśniających z poziomów niższych, w przekroju jednostek poziomów wyższych. W pierwszej kolejności uwzględnione zostało ewentualne zróżnicowanie zależności pomiędzy natężeniem dojazdów do pracy pomiędzy parami gmin a odległością między nimi, w przekroju gmin wyjazdu i przyjazdu, powiatów wyjazdu i przyjazdu oraz województw wyjazdu i przyjazdu (odpowiednio etapy 4.1, 4.2, 4.3, 4.4., 4.5 i 4.6). Następnie dopuszczono możliwość

występowania zróżnicowania pomiędzy zmienną objaśnianą a charakterystykami z poziomu gmin wyjazdu w przekroju powiatów wyjazdu oraz charakterystykami z poziomu gmin przyjazdu w przekroju powiatów przyjazdu (odpowiednio etapy 4.7 i 4.8) . W dalszej kolejności uwzględniono możliwość zróżnicowania charakteru zależności pomiędzy natężeniem dojazdów do pracy między parami gmin a zmiennymi określonymi dla gmin wyjazdu w przekroju województw wyjazdu oraz natężeniem dojazdów do pracy między parami gmin a zmiennymi określonymi dla gmin przyjazdu w przekroju województw przyjazdu (odpowiednio etapy 4.9 i 4.10). Na zakończenie sprawdzono wpływ uwzględnienia zróżnicowania zależności pomiędzy zmienną objaśnianą a charakterystykami powiatów wyjazdu w przekroju województw wyjazdu oraz charakterystykami powiatów przyjazdu w przekroju województw przyjazdu (odpowiednio etapy 4.11 i 4.12) .

Rozszerzeń modelu o kolejne efekty losowe dokonano przy pomocy regresji krokowej w przód. W celu weryfikacji poprawy jakości dopasowania modelu po wprowadzeniu każdej zmiany stosowano test ilorazu wiarygodności. Ponieważ w tym etapie zaleca się szczególną ostrożność przy rozszerzaniu modelu, poprawę uznawano za istotną jeżeli p-value

nie przekraczało 0,01 (w poprzednich podetapach przyjmowano poziom istotności 0,05).

$$\begin{aligned}
 Y_i = & \beta_{0000}^4 + \beta_{1000}^4 X_{1,1,i} + \sum_{p=2}^4 \sum_{q=1}^{Q_{a,p}} \beta_{\underbrace{a_0 \dots a_{p-1} q_0 \dots q_{p-1}}_{1 \dots p \dots 4}}^4 X_{a,p,q,i} + \sum_{p=2}^{P_b} \sum_{q=1}^{Q_{b,p}} \beta_{\underbrace{b_0 \dots b_{p-1} q_0 \dots q_{p-1}}_{1 \dots p \dots 4}}^4 X_{b,p,q,i} + \\
 & \sum_{p=2}^4 e_{\underbrace{a_0 \dots a_{p-1} (k) 0 \dots 0}_{1 \dots p \dots 4}}^4 + \sum_{p=2}^4 e_{\underbrace{b_0 \dots b_{p-1} (k) 0 \dots 0}_{1 \dots p \dots 4}}^4 + \sum_{p=2}^4 e_{\underbrace{1_0 \dots 1_{p-1} (k) 0 \dots 0}_{1 \dots p \dots 4}}^4 X_{1,1,i} + \\
 & \sum_{p=2}^4 \sum_{r=2}^{p-1} \sum_{q=1}^{Q_{a,r}} e_{\underbrace{a_0 \dots a_{p-1} q_0 \dots q_{p-1} (k) 0 \dots 0}_{1 \dots r \dots p \dots 4}}^4 X_{a,r,q,i} + \sum_{p=2}^4 \sum_{r=2}^{p-1} \sum_{q=1}^{Q_{b,r}} e_{\underbrace{a_0 \dots a_{p-1} q_0 \dots q_{p-1} (k) 0 \dots 0}_{1 \dots r \dots p \dots 4}}^4 X_{b,r,q,i} + r_i^4 \quad ,
 \end{aligned}
 \tag{4.25}$$

gdzie: $Q_{a,2}=Q_{b,2}=18$, $Q_{a,3}=Q_{b,3}=10$, $Q_{a,4}=Q_{b,4}=11$, $r_i^4 \sim N(0; \sigma_4^2)$,

$$e_{a,2}^4 = e_{a,0(k)00}^4 , e_{a,2}^4 \sim N(\mathbf{0}; \boldsymbol{\tau}_{a,2}^4) , e_{b,2}^4 = \begin{bmatrix} e_{b,0(k)00}^4 \\ e_{b,1(k)00}^4 \end{bmatrix} , e_{b,2}^4 \sim N(\mathbf{0}; \boldsymbol{\tau}_{b,2}^4) ,$$

$$e_{a,3}^4 = e_{a,0,18(k)0}^4 , e_{a,3}^4 \sim N(\mathbf{0}; \boldsymbol{\tau}_{a,3}^4) , e_{b,3}^4 = \begin{bmatrix} e_{b,0,0(k)0}^4 \\ e_{b,0,16(k)0}^4 \\ e_{b,0,12(k)0}^4 \end{bmatrix} , e_{b,3}^4 \sim N(\mathbf{0}; \boldsymbol{\tau}_{b,3}^4) ,$$

$$e_{a,4}^4 = e_{a,0,17,0(k)}^4, \mathbf{e}_{a,4}^4 \sim N(\mathbf{0}; \boldsymbol{\tau}_{a,4}^4), \mathbf{e}_{b,4}^4 = \begin{bmatrix} e_{b,0,0,0(k)}^4 \\ e_{b,0,10,0(k)}^4 \\ e_{b,0,5,0(k)}^4 \end{bmatrix}, \mathbf{e}_{b,4}^4 \sim N(\mathbf{0}; \boldsymbol{\tau}_{b,4}^4).$$

Tabela 4.53. Oceny parametrów strukturalnych oraz błędów standardowych szacunku modelu natężenia przyjazdów do pracy, podetapy etapu czwartego, Polska, 2006 (wzór 4.25)

Parametr:	Etap 4.2	Etap 4.7	Etap 4.8	Etap 4.9	Etap 4.10
$\beta_{0,0,0,0}^4$	0,089200	0,095770	0,100400	0,099800	0,103100
$S(\beta_{0,0,0,0}^4)$	0,004700	0,004524	0,004575	0,004534	0,005107
$\beta_{1,0,0,0}^4$	-0,003859	-0,003887	-0,003864	-0,003866	-0,003854
$S(\beta_{1,0,0,0}^4)$	0,000095	0,000095	0,000095	0,000095	0,000095
$\beta_{a,0,18,0,0}^4$	0,000016	0,000022	0,000022	0,000013	0,000015
$S(\beta_{a,0,18,0,0}^4)$	0,000001	0,000001	0,000001	0,000002	0,000002
$\beta_{a,0,16,0,0}^4$	2,623000	1,044000	1,097000	1,355000	1,339000
$S(\beta_{a,0,16,0,0}^4)$	1,648000	1,541000	1,541000	1,537000	1,534000
$\beta_{a,0,10,0,0}^4$	0,000726	-0,001064	-0,001126	-0,003961	-0,003669
$S(\beta_{a,0,10,0,0}^4)$	0,001986	0,001837	0,001838	0,001969	0,001946
$\beta_{b,0,17,0,0}^4$	0,000027	0,000024	0,000046	0,000045	0,000049
$S(\beta_{b,0,17,0,0}^4)$	0,000008	0,000008	0,000008	0,000008	0,000008
$\beta_{b,0,10,0,0}^4$	-0,016740	-0,016230	-0,039650	-0,039280	-0,054820
$S(\beta_{b,0,10,0,0}^4)$	0,002491	0,002446	0,003883	0,003872	0,008337
$\beta_{b,0,14,0,0}^4$	-0,389500	-0,342400	-0,263600	-0,256800	-0,315300
$S(\beta_{b,0,14,0,0}^4)$	0,064560	0,062100	0,063330	0,062880	0,072130
$\beta_{b,0,3,0,0}^4$	0,000445	0,000375	0,000143	0,000176	0,000215
$S(\beta_{b,0,3,0,0}^4)$	0,000170	0,000158	0,000163	0,000161	0,000157
$\beta_{b,0,18,0,0}^4$	-0,000023	-0,000022	-0,000027	-0,000027	-0,000027
$S(\beta_{b,0,18,0,0}^4)$	0,000003	0,000003	0,000003	0,000003	0,000003
$\beta_{b,0,12,0,0}^4$	0,000104	0,000097	0,000103	0,000103	0,000101
$S(\beta_{b,0,12,0,0}^4)$	0,000025	0,000023	0,000026	0,000026	0,000025
$\beta_{a,0,0,6,0}^4$	0,038890	0,020000	0,021600	0,023090	0,024440
$S(\beta_{a,0,0,6,0}^4)$	0,004150	0,004679	0,004683	0,004635	0,004571
$\beta_{b,0,0,3,0}^4$	-0,002561	-0,002782	-0,003112	-0,003097	-0,003081
$S(\beta_{b,0,0,3,0}^4)$	0,000819	0,000762	0,000766	0,000757	0,000678
$\beta_{a,0,0,0,8}^4$	-30,110000	-17,130000	-17,850000	-15,040000	-27,580000
$S(\beta_{a,0,0,0,8}^4)$	19,250000	18,780000	18,760000	18,880000	17,350000
$\beta_{a,0,0,0,7}^4$	0,543700	0,316400	0,366800	0,360700	0,457100
$S(\beta_{a,0,0,0,7}^4)$	0,293700	0,291300	0,291100	0,292000	0,265900

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.54. Oceny wariancji składników losowych modelu natężenia przyjazdów do pracy, podetapy etapu czwartego, Polska, 2006 (wzór 4.25)

wariancja	Etap 4.2		Etap 4.7		Etap 4.8	
σ_4^2	0,001242		0,0012465		0,001251	
$\tau_{a,2}^4$	0,000051		0,000000000016		0,00000000000008	
$\tau_{b,2}^4$	0,004529	-0,000249	0,004484	-0,000248	0,004381	-0,000250
	-0,000249	0,000015	-0,000248	0,000015	-0,000250	0,000015
$\tau_{a,3}^4$	-		0,000000000213		0,0000000002	
$\tau_{b,3}^4$	0,0000204		0,000000000005		0,00000000	00003
					0	0
					0	957,66
						0
						0,00000000
					0	0
						7
$\tau_{a,4}^4$	-		-		-	
$\tau_{b,4}^4$	0,000031		0,000018		0,00001662	

wariancja	Etap 4.9		Etap 4.10	
σ_4^2	0,001250		0,001249	
$\tau_{a,2}^4$	0,000000000000003		0,00000000005	
$\tau_{b,2}^4$	0,004360	-0,000249	0,004262	-0,000245
	-0,000249	0,000015	-0,000245	0,000015
$\tau_{a,3}^4$	0,0000000002		0,0000000002	
$\tau_{b,3}^4$	0,00000000000004	0	0,000001	0
	0	955,99	0	929,52
	0	0	0,00000007	0,00000001
$\tau_{a,4}^4$	0,000000001		0,000000001	
$\tau_{b,4}^4$	0,000012		0,000368	-0,000387
			-0,000387	0,000793
			-0,001744	0,000777
			0,000777	0,011175

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.55. Wybrane kryteria oceny jakości dopasowania modelu natężenia przyjazdów do pracy oszacowanego w etapie 4 (wzór 4.25) oraz ich porównanie z etapem 3.5

Kryterium oceny	Etap 4.2	Etap 4.7	Etap 4.8	Etap 4.9	Etap 4.10
lnL	44762,21	44919,71	44995,77	45010,97	45061,68
AIC	-89480,42	-89793,43	-89941,53	-89969,93	-90061,36
BIC	-89301,3	-89606,17	-89738	-89758,25	-89808,98
vs	Etap 3.5	Etap 4.2	Etap 4.7	Etap 4.8	Etap 4.9
Chi2	4428,86	315,01	152,10	30,40	101,43
p-value	0,000000	0,000000	0,000000	0,000000	0,000000

Źródło: Opracowanie własne na podstawie danych GUS

Na podstawie testu ilorazu wiarygodności stwierdzono statystycznie istotną na praktycznie dowolnym poziomie istotności poprawę jakości dopasowania modelu w stosunku do etapu trzeciego oraz po wprowadzeniu zmian w kolejnych z podetapów etapu czwartego (por. 4.55).

ETAP 5 – Model z losowymi współczynnikami regresji zależnymi od zmiennych z wyższych poziomów

W kolejnych podetapach etapu piątego możliwe jest wyjaśnienie części zmienności zależności pomiędzy natężeniem dojazdów do pracy dla par gmin a zmiennymi objaśniającymi, przy pomocy charakterystyk z wyższych poziomów. Możliwe jest uwzględnienie takiego zróżnicowania w przypadku wszystkich możliwych par zmiennych objaśniających z różnych poziomów ze względu na to samo kryterium grupowania. Wprowadzenie zmiennych nie jest uzależnione od tego, czy zostały one wcześniej wprowadzone oddzielnie do modelu oraz czy stwierdzono istotne zróżnicowanie charakteru zależności pomiędzy natężeniem dojazdów do pracy a zmienną objaśniającą z niższego poziomu, w przekroju jednostek terytorialnych, dla których określona jest charakterystyka wyższego poziomu. Wyjaśnienie części zmienności zależności pomiędzy zmienną objaśnianą a charakterystyką niższego poziomu przez charakterystykę poziomu wyższego odbywa się przez wprowadzenie do modelu interakcji pomiędzy obydwojema zmiennymi objaśniającymi (por wzór 4.26).

W pierwszej kolejności do modelu wprowadzono interakcję pomiędzy odległością między gminami a zmiennymi z poziomu gmin wyjazdu i przyjazdu, powiatów wyjazdu i przyjazdu oraz województw wyjazdu i przyjazdu (odpowiednio etapy 5.1, 5.2, 5.3, 5.4., 5.5 i 5.6). Następnie dopuszczono możliwość występowania interakcji pomiędzy charakterystykami gmin wyjazdu a zmiennymi z poziomu powiatów wyjazdu oraz pomiędzy charakterystykami gmin przyjazdu a zmiennymi z poziomu powiatów przyjazdu (odpowiednio etapy 5.7 i 5.8). W dalszej kolejności uwzględniono możliwość wyjaśnienia części zróżnicowania charakteru zależności pomiędzy natężeniem dojazdów do pracy a zmiennymi określonymi dla gmin wyjazdu w przy pomocy charakterystyk województw wyjazdu oraz natężeniem dojazdów do pracy a zmiennymi określonymi dla gmin przyjazdu przy pomocy zmiennych określonych dla województw przyjazdu (odpowiednio etapy 5.9 i 5.10). Na koniec zaś podjęto próbę wyjaśnienia części zróżnicowania charakteru zależności pomiędzy natężeniem dojazdów a zmiennymi z poziomu powiatów wyjazdu przy pomocy charakterystyk województw wyjazdu oraz zmiennych określonych dla powiatów przyjazdu przez charakterystyki województw przyjazdu (odpowiednio etapy 5.11 i 5.12). Ze względu na trudności interpretacyjne oraz obliczeniowe postanowiono ograniczyć rozszerzanie modelu do interakcji pomiędzy

parami zmiennych objaśniających. Teoretycznie możliwe jest jednak rozszerzenie modelu także o interakcje pomiędzy trzema i czterema zmiennymi objaśniającymi.

Podobnie jak w etapie poprzednim, w celu zachowania szczególnej ostrożności podczas rozszerzania modelu dodawano do niego tylko te interakcje pomiędzy zmiennymi objaśniającymi, które na poziomie istotności 0,01 poprawiały jakość jego dopasowania (według testu ilorazu wiarygodności).

$$\begin{aligned}
 Y_i = & \beta_{0000}^5 + \beta_{1000}^5 X_{1,1,i} + \sum_{p=2}^4 \sum_{q=1}^{Q_{a,p}} \beta_{\underbrace{a_0 \dots 0 q_0 \dots 0}_{1 \dots p \dots 4}}^5 X_{a,p,q,i} + \sum_{p=2}^{P_b} \sum_{q=1}^{Q_{b,p}} \beta_{\underbrace{b_0 \dots 0 q_0 \dots 0}_{1 \dots p \dots 4}}^5 X_{b,p,q,i} + \\
 & \sum_{p_1=1}^3 \sum_{p_2=p_1}^4 \sum_{q_1=1}^{Q_{a,p_1}} \sum_{q_2=1}^{Q_{a,p_2}} \beta_{\underbrace{a_0 \dots q_1 \dots q_2 \dots 0}_{1 \dots p_1 \dots p_2 \dots 4}}^5 X_{a,p_1,q_1,i} X_{a,p_2,q_2,i} + \\
 & \sum_{p_1=1}^3 \sum_{p_2=p_1}^4 \sum_{q_1=1}^{Q_{b,p_1}} \sum_{q_2=1}^{Q_{b,p_2}} \beta_{\underbrace{b_0 \dots q_1 \dots q_2 \dots 0}_{1 \dots p_1 \dots p_2 \dots 4}}^5 X_{b,p_1,q_1,i} X_{b,p_2,q_2,i} + \sum_{p=2}^4 e_{\underbrace{a_0 \dots 0(k) 0 \dots 0}_{1 \dots p \dots 4}}^5 + \\
 & \sum_{p=2}^4 e_{\underbrace{b_0 \dots 0(k) 0 \dots 0}_{1 \dots p \dots 4}}^5 + \sum_{p=2}^4 e_{\underbrace{1_0 \dots 0(k) 0 \dots 0}_{1 \dots p \dots 4}}^5 X_{1,1,i} + \sum_{p=2}^4 \sum_{r=2}^{p-1} \sum_{q=1}^{Q_{a,r}} e_{\underbrace{a_0 \dots 0 q_0 \dots 0(k) 0 \dots 0}_{1 \dots r \dots p \dots 4}}^5 X_{a,r,q,i} + \\
 & \sum_{p=2}^4 \sum_{r=2}^{p-1} \sum_{q=1}^{Q_{b,r}} e_{\underbrace{a_0 \dots 0 q_0 \dots 0(k) 0 \dots 0}_{1 \dots r \dots p \dots 4}}^5 X_{b,r,q,i} + r_i^5, \tag{4.26}
 \end{aligned}$$

gdzie: $Q_{a,2}=Q_{b,2}=18$, $Q_{a,3}=Q_{b,3}=10$, $Q_{a,4}=Q_{b,4}=11$, $r_i^5 \sim N(0; \sigma_5^2)$, $e_{a,2}^5 = e_{a,0(k)00}^5$,

$$e_{a,2}^5 \sim N(\mathbf{0}; \boldsymbol{\tau}_{a,2}^5), e_{b,2}^5 = \begin{bmatrix} e_{b,0(k)00}^5 \\ e_{b,1(k)00}^5 \end{bmatrix}, e_{b,2}^5 \sim N(\mathbf{0}; \boldsymbol{\tau}_{b,2}^5),$$

$$e_{a,3}^5 = e_{a,0,18(k)0}^5, e_{a,3}^5 \sim N(\mathbf{0}; \boldsymbol{\tau}_{a,3}^5), e_{b,3}^5 = \begin{bmatrix} e_{b,00(k)0}^5 \\ e_{b,0,16(k)0}^5 \\ e_{b,0,12(k)0}^5 \end{bmatrix}, e_{b,3}^5 \sim N(\mathbf{0}; \boldsymbol{\tau}_{b,3}^5),$$

$$e_{a,4}^5 = e_{a,0,17,0(k)}^5, e_{a,4}^5 \sim N(\mathbf{0}; \boldsymbol{\tau}_{a,4}^5), e_{b,4}^5 = \begin{bmatrix} e_{b,000(k)}^5 \\ e_{b,0,10,0(k)}^5 \\ e_{b,0,5,0(k)}^5 \end{bmatrix}, e_{b,4}^5 \sim N(\mathbf{0}; \boldsymbol{\tau}_{b,4}^5).$$

Tabela 4.56. Oceny parametrów strukturalnych oraz błędów standardowych szacunku modelu natężenia przyjazdów do pracy, podetapy etapu piątego, Polska, 2006 (wzór 4.26)

Parametr:	Etap 5.1	Etap 5.2	Etap 5.3	Etap 5.4	Etap 5.7	Etap 5.8	Etap 5.9
$\beta_{0,0,0,0}^5$	0,09951	0,08158	0,07814	0,07679	0,07857	0,08325	0,08239
$S(\beta_{0,0,0,0}^5)$	0,004965	0,005996	0,006031	0,006068	0,006138	0,006129	0,006138
$\beta_{1,0,0,0}^5$	-0,003926	-0,002862	-0,002823	-0,002915	-0,002919	-0,002901	-0,002882
$S(\beta_{1,0,0,0}^5)$	0,000096	0,000258	0,000257	0,000261	0,000260	0,000261	0,000261
$\beta_{a,0,18,0,0}^5$	0,000015	0,000015	0,000017	0,000017	0,000015	0,000015	0,000015
$S(\beta_{a,0,18,0,0}^5)$	0,000002	0,000002	0,000002	0,000002	0,000002	0,000002	0,000002
$\beta_{a,0,16,0,0}^5$	1,244	0,9407	0,9536	0,932	0,6211	0,6111	0,844
$S(\beta_{a,0,16,0,0}^5)$	1,519	1,52	1,521	1,52	1,517	1,517	1,519
$\beta_{a,0,10,0,0}^5$	-0,0049	-0,005098	-0,00465	-0,004626	-0,00465	-0,004585	-0,00529

Parametr:	Etap 5.1	Etap 5.2	Etap 5.3	Etap 5.4	Etap 5.7	Etap 5.8	Etap 5.9
$S(\beta_{a,0,10,0,0}^5)$	0,00196	0,001965	0,00195	0,001949	0,001972	0,001971	0,00199
$\beta_{b,0,17,0,0}^5$	0,00005	-0,000026	-0,000028	-0,000024	-0,000024	-0,000023	-0,000022
$S(\beta_{b,0,17,0,0}^5)$	0,000008	0,000012	0,000012	0,000012	0,000012	0,000012	0,000012
$\beta_{b,0,10,0,0}^5$	-0,0528	-0,05106	-0,04918	-0,05275	-0,05342	-0,05232	-0,05247
$S(\beta_{b,0,10,0,0}^5)$	0,007989	0,007589	0,007638	0,007669	0,007797	0,007584	0,007613
$\beta_{b,0,14,0,0}^5$	-0,3291	-0,4295	-0,433	-0,4346	-0,4313	-0,4076	-0,3995
$S(\beta_{b,0,14,0,0}^5)$	0,07252	0,07607	0,07507	0,07616	0,07625	0,07612	0,07622
$\beta_{b,0,3,0,0}^5$	0,000249	0,002568	0,002637	0,002713	0,002686	0,002607	0,002768
$S(\beta_{b,0,3,0,0}^5)$	0,000157	0,000341	0,00034	0,000343	0,000344	0,000343	0,000348
$\beta_{b,0,18,0,0}^5$	-0,000028	-0,000023	-0,000023	-0,000024	-0,000023	-0,00003	-0,000031
$S(\beta_{b,0,18,0,0}^5)$	0,000003	0,000003	0,000003	0,000003	0,000003	0,000004	0,000004
$\beta_{b,0,12,0,0}^5$	0,000124	0,000142	0,000139	0,000135	0,000121	0,000124	0,000127
$S(\beta_{b,0,12,0,0}^5)$	0,000025	0,000025	0,000024	0,000025	0,000025	0,000025	0,000025
$\beta_{a,0,0,6,0}^5$	0,02746	0,02797	0,04053	0,04184	0,04153	0,041	0,03951
$S(\beta_{a,0,0,6,0}^5)$	0,00455	0,004556	0,005154	0,005168	0,005197	0,005189	0,005218
$\beta_{b,0,0,3,0}^5$	-0,003539	-0,004188	-0,004132	-0,00385	-0,004297	-0,008066	-0,00814
$S(\beta_{b,0,0,3,0}^5)$	0,000686	0,000695	0,000679	0,000707	0,000721	0,001079	0,001079
$\beta_{a,0,0,0,8}^5$	-1,995	-4,096	-1,95	-0,06603	-12,85	-9,324	-11,3
$S(\beta_{a,0,0,0,8}^5)$	17,21	17,13	17,04	17,16	17,37	17,24	17,26
$\beta_{a,0,0,0,7}^5$	0,3575	0,3324	0,2848	0,3524	0,3304	0,2865	0,4658
$S(\beta_{a,0,0,0,7}^5)$	0,2634	0,2623	0,2604	0,2628	0,2633	0,2611	0,2705
$\beta_{a,1,17,0,0}^5$	-	-	-	-	-	-	-
	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
	16	16	03	03	03	03	03
$S(\beta_{a,1,17,0,0}^5)$	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
	03	03	04	04	04	04	04
$\beta_{a,1,18,0,0}^5$	-	-	-	-	-	-	-
	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
	05	05	08	08	08	08	08
$S(\beta_{a,1,18,0,0}^5)$	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
	01	01	02	02	02	02	02
$\beta_{b,1,17,0,0}^5$	-	0,000004	0,000004	0,000004	0,000004	0,000004	0,000004
$S(\beta_{b,1,17,0,0}^5)$	-	0,000000	0,000000	0,000000	0,000000	0,000000	0,000000
		5	5	5	5	5	5
$\beta_{b,1,3,0,0}^5$	-	-0,000153	-0,000152	-0,000199	-0,000198	-0,000198	-0,000206
$S(\beta_{b,1,3,0,0}^5)$	-	0,000002	0,000002	0,000026	0,000026	0,000025	0,000026
$\beta_{b,1,11,0,0}^5$	-	0,03432	0,03301	0,03495	0,03566	0,0313	0,0318
$S(\beta_{b,1,11,0,0}^5)$	-	0,008475	0,008428	0,008482	0,008507	0,00851	0,008515
$\beta_{a,1,0,6,0}^5$	-	-	-0,000241	-0,000245	-0,00024	-0,00024	-0,000245
$S(\beta_{a,1,0,6,0}^5)$	-	-	0,000046	0,000046	0,000046	0,000046	0,000046
$\beta_{b,1,0,1,0}^5$	-	-	-	0,005513	0,005459	0,005419	0,006110
$S(\beta_{b,1,0,1,0}^5)$	-	-	-	0,001894	0,001897	0,001886	0,001904
$\beta_{a,0,12,4,0}^5$	-	-	-	-	0,000086	0,000095	0,000095
$S(\beta_{a,0,12,4,0}^5)$	-	-	-	-	0,000022	0,000022	0,000022
$\beta_{a,0,17,1,0}^5$	-	-	-	-	0,000209	0,000207	0,000216
$S(\beta_{a,0,17,1,0}^5)$	-	-	-	-	0,000056	0,000056	0,000056
$\beta_{b,0,18,3,0}^5$	-	-	-	-	-	0,000004	0,000004
$S(\beta_{b,0,18,3,0}^5)$	-	-	-	-	-	0,000001	0,000001
$\beta_{a,0,3,0,7}^5$	-	-	-	-	-	-	-0,02026
$S(\beta_{a,0,3,0,10}^5)$	-	-	-	-	-	-	0,007761

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.57. Oceny wariancji składników losowych modelu natężenia przyjazdów do pracy, podetapy etapu piątego, Polska, 2006 (wzór 4.26)

wariancja	Etap 5.1		Etap 5.2		Etap 5.3	
σ_5^2	0,001220		0,001224		0,001224	
$\tau_{a,2}^5$	0		0		0,0000000007	
$\tau_{b,2}^5$	0,004319	-0,000250	0,003853	-0,000219	0,003872	-0,000218
	-0,000250	0,000015	-0,000219	0,000013	-0,000218	0,000013
$\tau_{a,3}^5$	0,0000000002		0,0000000002		0,0000000002	
$\tau_{b,3}^5$	0	0	0	0	0,000015	0
	0	950,95	0	950,38	0	461,8
	0	0	0	0	0	0
	0,0000000006		0,0000000007		0,0000000001	
$\tau_{a,4}^5$	0,0000000017		0,0000000018		0,0000000012	
$\tau_{b,4}^5$	0,000318		0,000301		0,000310	

wariancja	Etap 5.4		Etap 5.7		Etap 5.8	
σ_5^2	0,001222		0,001222		0,001222	
$\tau_{a,2}^5$	0		0,000000000000003		0	
$\tau_{b,2}^5$	0,003856	-0,000219	0,003854	-0,000219	0,003848	-0,000219
	-0,000219	0,000013	-0,000219	0,000013	-0,000219	0,000013
$\tau_{a,3}^5$	0,00000000017		0,00000000016		0,00000000016	
$\tau_{b,3}^5$	0	0	0	0	0	0
	0	969,05	0	971,56	0	963,11
	0	0	0	0	0	0
	0,0000000007		0,0000000007		0,0000000007	
$\tau_{a,4}^5$	0,0000000012		0,0000000003		0,0000000003	
$\tau_{b,4}^5$	0,000314		0,000337		0,000323	

wariancja	Etap 5.9	
σ_5^2	0,001222	
$\tau_{a,2}^5$	0,0000000000000000003	
$\tau_{b,2}^5$	0,003846	-0,000219
	-0,000219	0,000013
$\tau_{a,3}^5$	0,00000000016	
$\tau_{b,3}^5$	0,0000000000000004	0
	0	959,63
	0	0
	0,0000000007	
$\tau_{a,4}^5$	0,00000000033	
$\tau_{b,4}^5$	0,000321	

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.58. Wybrane kryteria oceny jakości dopasowania modelu natężenia przyjazdów do pracy oszacowanego w etapie 5 (wzór 4.26) oraz ich porównanie z etapem 4.1

Kryterium oceny	Etap 5.1	Etap 5.2	Etap 5.3	Etap 5.4	Etap 5.7	Etap 5.8	Etap 5.9
lnL	45278,06	45367,51	45371,87	45385,09	45398,45	45408,61	45412,8
AIC	-90490,12	-90663,02	-90669,74	-90694,18	-90716,89	-90735,21	-90740,01
BIC	-90221,45	-90369,92	-90368,5	-90384,81	-90391,23	-90401,41	-90398,06
vs	etap 4.10	etap 5.1	etap 5.2	etap 5.3	etap 5.4	etap 5.7	etap 5.8
Chi2	432,76	178,9	8,72	26,44	26,71	20,32	6,7933
p-value	0,000000	0,000000	0,003146	0,000000	0,000002	0,000007	0,00915

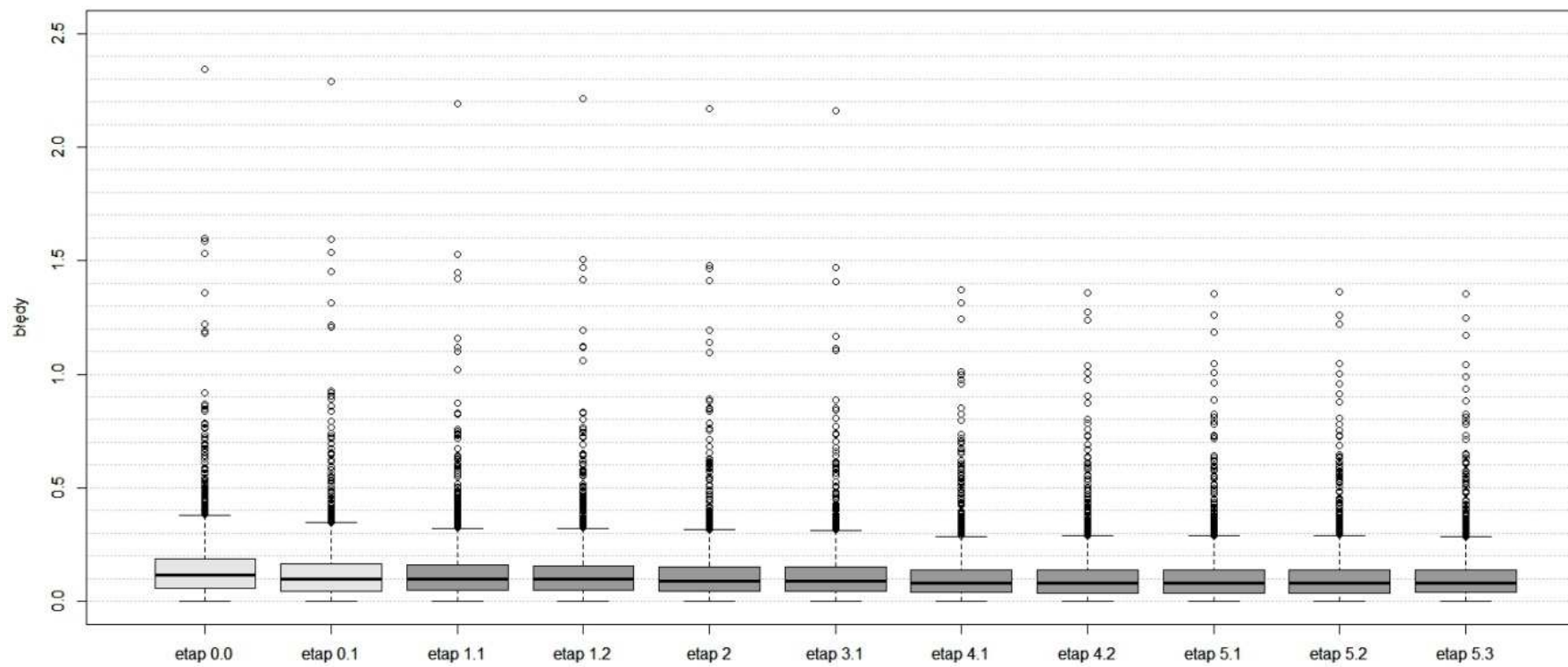
Źródło: Opracowanie własne na podstawie danych GUS

Na podstawie testu ilorazu wiarygodności stwierdzono istotną, na poziomie istotności przynajmniej 0,01, poprawę jakości dopasowania modelu po wprowadzeniu rozszerzeń opisanych w kolejnych podetapach etapu piątego. Ponieważ jest to ostatni etap konstrukcji modelu czteropoziomowego z dwoma kryteriami grupowania, model opisany równaniem (4.26) przyjęty zostanie do dalszej analizy. Przy jego pomocy, po podstawieniu zmiennych objaśniających aktualnych dla roku 2011, podjęta zostanie próba oszacowania elementów macierzy przepływów ludności związanych z zatrudnieniem w roku 2011.

4.3. Ocena jakości oszacowania elementów macierzy migracji

Dzięki zastosowaniu modelu trzypoziomowego przy szacowaniu wartości na diagonalu uzyskano znaczną poprawę jakości dopasowania modelu (wiarygodność wzrosła o 36%, wartość kryterium Akaike'a spadła o 32% zaś kryterium Schwarz'a o prawie 20%) w stosunku do klasycznej regresji liniowej (por. tab. 4.7 i 4.34). Ponadto szacując natężenie przyjazdów do pracy przy pomocy trzypoziomowego pełnego modelu oszacowanego w podrozdziale 4.2, popełnimy przeciętnie błąd o ponad 17% mniejszy niż w przypadku klasycznej regresji liniowej (por. tab. 4.59). Zauważyć można, że nastąpił równomierny spadek błędów szacunku w porównaniu z klasyczną regresją liniową, ponieważ obniżyły się wartości wszystkich decyli (por. tab. 4.60).

Również w trakcie wprowadzania kolejnych rozszerzeń zaobserwowano konsekwentne zmniejszanie się błędów oszacowań (por. wyk. 4.10 i tab. 4.59). Stwierdzono mianowicie konsekwentny spadek wszystkich kwartyli charakteryzujących popełniane błędy, a także wartości błędu, powyżej którego byłby on uznawany za nietypowo wysoki (por. wyk. 4.10). Również wartości największych błędów szacunku zmniejszyły się dzięki zastosowaniu modeli oszacowanych w kolejnych podetapach. Szczególny spadek silnie odstających błędów zaobserwowano po wprowadzeniu w etapie czwartym zmiennych objaśniających z poziomu powiatów do współczynników kierunkowych przy zmiennych z poziomu gmin (por. wyk. 4.10).



Wykres 4. 10. Rozkład ocen błędów standardowych szacunku natężenia przyjazdów dla różnych modeli, Polska, 2006

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.59. Wartości średnie ocen błędów standardowych szacunku natężenia przyjazdów do pracy w gminach, wybrane modele, Polska, 2006

Etap	0.1	1.2	2	3.1	4.2	5.3
Średnia ocena standardowego błędu szacunku	0,1236873	0,1222041	0,1155903	0,1155627	0,1032609	0,1024518

Źródło: Opracowanie własne na podstawie danych GUS

Tabela 4.60. Charakterystyki rozkładu błędów standardowych szacunku natężenia przyjazdów do pracy w gminach według modelu trypoziomowego w porównaniu z klasycznym modelem regresji liniowej, Polska, 2006

Model regresji	Wartość min	Decyle rozkładu błędów									Wartość max
		10	20	30	40	50	60	70	80	90	
regresja liniowa	0,00015	0,0179	0,0364	0,0555	0,0753	0,0973	0,1219	0,1507	0,1847	0,2377	4,2161
model trypoziomowy	0,00001	0,0143	0,0292	0,0470	0,0621	0,0804	0,0994	0,1225	0,1543	0,2007	1,3539

Źródło: Opracowanie własne na podstawie danych GUS

Dla połowy powiatów uzyskane z użyciem modelu trypoziomowego oszacowanie różniło się od faktycznej wartości o nie więcej niż 0,0804. W przypadku jednej piątej gmin uzyskano błąd nie większy niż 0,0292, z kolei tylko dla 10 procent gmin otrzymano oszacowanie różniące się od wartości empirycznej o więcej niż 0,2007 (por. tab. 4.60). Ponadto wszystkie wyznaczone kwantyle błędów oszacowań przyjmowały mniejsze wartości przy oszacowaniach wyznaczonych przy pomocy modelu trypoziomowego niż klasycznej regresji liniowej (por. tab. 4.60). Wartość największego błędu oszacowania była zaś aż trzykrotnie większa w przypadku nie uwzględnienia struktury trypoziomowej (por. wykres 4.10 i tab. 4.60). Zarówno w przypadku modelu trypoziomowego, jak i klasycznej regresji liniowej, dla żadnej z gmin nie oszacowano ujemnego natężenia przyjazdów. Ponieważ natężenie przyjazdów do pracy w naturalny sposób powinno przyjmować wartości nieujemne, jest to niewątpliwą zaletą skonstruowanych modeli.

Przy pomocy skonstruowanego modelu trypoziomowego oszacowano natężenie przyjazdów w gminach dla roku 2011³³. Otrzymane oszacowania przemnożono następnie przez liczbę osób pracujących w tych gminach w 2011 roku, w ten sposób uzyskując

³³ Pomiędzy rokiem 2006 a 2011 następujące gminy zostały podzielone na część miejską i wiejską: Wolbórz, Nowe Brzesko, Bobowa, Szczucin, Radłów, Łaszczów, Kołaczyce, Pruchnik, Brzostek, Przeclaw, Michałowo, Krynki, Czyżew, Szepietowo, Tychowo i Gościno. W celu dokonania oszacowań przy pomocy modelu trypoziomowego skonstruowanego dla roku 2006, konieczne było więc połączenie obu części dla roku 2011, w tym celu zsumowano odpowiednie dane bezwzględne.

oszacowania liczby osób pracujących na terenie gmin, jednak jednocześnie zamieszkałych poza ich granicami. Następnie wartości te odjęto od znanych dla roku 2011 liczebności ludności pracującej w gminach, otrzymując tym samym oszacowanie elementów na diagonalu macierzy przepływów aktualnej dla roku 2011.

Po dokonaniu oszacowań sprawdzono, ile procent zmienności zmiennej objaśnianej wyjaśnione zostało przy pomocy skonstruowanego modelu. W tym celu skorzystano z przybliżenia $R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$. Okazało się, że precyzja oszacowań natężenia przyjazdów dokonanych przy pomocy modelu trypoziomowego utrzymywała się na akceptowalnym poziomie ($R^2=0,499$), czego raczej nie można stwierdzić w przypadku modelu klasycznej regresji liniowej ($R^2=0,135$). Również w przypadku wartości bezwzględnych liczby osób przyjeżdżających do pracy oraz liczby osób pracujących w gminie zamieszkania zauważono znaczną przewagę modelu trypoziomowego nad klasyczną regresją liniową (por. tab. 4.61).

Tabela 4.61. Wartości współczynników determinacji modeli natężenia przyjazdów do pracy dla elementów na diagonalu macierzy migracji

Model	Natężenie przyjazdów	Liczba osób przyjeżdżających do pracy	Liczba osób pracujących w gminie zamieszkania ¹
Klasyczna regresja liniowa	0,1347983	0,5696773	0,9542779
Model wielopoziomowy	0,4993631	0,9092982	0,9903545

Uwaga:

¹ W przypadku klasycznej regresji liniowej w gminie Kleszczów z województwa łódzkiego oszacowano ujemną wartość mieszkańców zatrudnionych na jej terenie (-2367 osób), którą zastąpiono zerem. Z kolei zgodnie

z modelem wielopoziomowym ujemne wartości osób pracujących w obszarze zamieszkania oszacowano dla części wiejskiej gminy Olszyna z województwa Dolnośląskiego, części wiejskiej gminy Polkowice z województwa Dolnośląskiego, części miejskiej gminy Kleczew z Wielkopolski, gminie Ostrowite z Wielkopolski oraz gminie Powidz, również z Wielkopolski (odpowiednio -53, -1905, -542, -42 i -577).

Źródło: Opracowanie własne na podstawie danych GUS

Analogiczną weryfikację jakości modelu i błędów szacunku przeprowadzono dla modelu wykorzystanego do oszacowania elementów macierzy migracji poza diagonalą (por. tab. 4.62).

Tabela 4.62. Wartości współczynników determinacji modeli natężenia dojazdów do pracy dla elementów poza diagonalą macierzy migracji

Model	Natężenie przepływów	Liczba osób dojeżdżających do pracy pomiędzy parami gmin
Klasyczna regresja liniowa	0,2000348	0,3750672
Model wielopoziomowy	0,6312913	0,7119627

Źródło: Opracowanie własne na podstawie danych GUS

Również w przypadku modelu mającego na celu oszacowanie przepływów związanych z zatrudnieniem pomiędzy parami gmin zaobserwowano znaczną przewagę zastosowanego modelu czteropoziomowego z dwoma kryteriami grupowania nad klasyczną regresją liniową (por. tab. 4.62). Należy jednak zaznaczyć, że precyzja szacunków poza diagonalą jest istotnie niższa niż dla diagonal, pomimo zastosowania znacznie bardziej zaawansowanego modelu. Wynika to z faktu, że pary gmin, dla których szacowane są przepływy związane z zatrudnieniem są niezwykle trudną do badania populacją. Ponadto pogłębiona analiza otrzymanych oszacowań wskazuje, że błędy popełnione podczas estymacji natężenia przepływów związanych z zatrudnieniem pomiędzy parami gmin mają pewnego rodzaju charakter systematyczny. Mianowicie przeszacowane zostały wartości dojazdów na dużych odległościach przy jednoczesnym niedoszacowaniu przepływów pomiędzy pobliskimi gminami. Jedną z przyczyn takiego stanu rzeczy może być fakt, że estymację przepływów pomiędzy parami gmin przeprowadzono na wartościach względnych i dopiero po otrzymaniu oszacowań przemnożono je przez wartości mianownika, czyli liczbę osób pracujących w gminach przyjazdu. Ponieważ dojazdy na dużych odległościach skierowane są zazwyczaj do największych ośrodków miejskich, charakteryzujących się dużą liczbą osób pracujących, podczas przechodzenia z wartości względnych przepływów, oszacowanych przy pomocy modelu czteropoziomowego, na bezwzględne, konieczne do konstrukcji macierzy migracji, błędy popełnione przy szacowaniu dojazdów na dużych odległościach uległy powiększeniu.

Rozdział 5. Delimitacja lokalnych rynków pracy w Polsce

Na podstawie macierzy migracji z badania przepływów ludności związanych z zatrudnieniem oraz przeprowadzonych szacunków podjęto próbę delimitacji lokalnych rynków pracy w Polsce. W tym celu zastosowano dwa algorytmy zaczerpnięte z literatury (por. Coombes, Green, Openshaw, 1986; Kristensen, 1998) oraz samodzielnie opracowane podejście taksonomiczne, opisane w rozdziale drugim. W pierwszej kolejności, w podrozdziale 5.1. delimitację przeprowadzono na podstawie oryginalnych danych o dojazdach do pracy dla roku 2006 (por. Dojazdy do pracy w Polsce, red. K. Kruszka, 2010). W dalszej kolejności podjęto próbę delimitacji lokalnych rynków pracy w Polsce, dokonaną w oparciu o oszacowaną w rozdziale czwartym macierz przepływów ludności związanych z zatrudnieniem dla 2011 roku. Rozważania kończy ewaluacja otrzymanych wyników oraz analiza porównawcza wyodrębnionych lokalnych rynków pracy.

5.1. Delimitacja lokalnych rynków pracy w Polsce dla roku 2006

Delimitację lokalnych rynków pracy przeprowadzono dla najmniejszych jednostek podziału administracyjnego kraju, tj. wszystkich gmin polskich. Uwzględnienie podziału na część miejską i wiejską spowodowało, że liczba jednostek bazowych wynosiła 3062. Określenie zakresu czasowego badania jako rok 2006 zdeterminowane było dostępnością danych.

5.1.1. Zastosowanie europejskiego algorytmu regionalizacji

Europejski algorytm regionalizacji, związany jest z ideą konstrukcji lokalnych rynków pracy wokół wyodrębnionych ośrodków centralnych. Charakteryzując tę metodę delimitacji, wskazano na problem definicji ośrodków centralnych oraz dużą liczbę parametrów. Procedura postępowania została szczegółowo przedstawiona w podrozdziale 2.3.

W trakcie działania algorytmu porównuje się wartości dziesięciu parametrów z odpowiednimi wartościami granicznymi. Wielkości parametrów, dla każdej z jednostek bazowych wyznaczone są na podstawie macierzy przepływów związanych z zatrudnieniem. Modyfikacja wartości granicznych daje możliwość dostosowania procedury do warunków danego kraju. Decyzja dotycząca poszczególnych wartości

krytycznych ma charakter ekspercki i zależy od subiektywnej opinii badacza. W zależności od doboru tych wartości możliwe jest otrzymanie sprzecznych ze sobą wyników. Autorzy algorytmu, nie sprecyzowali sposobu ustalania wartości krytycznych poszczególnych parametrów. Tym samym pozostawili, stosującym go badaczom, dużą dowolność. Wartości krytyczne ustala się w zależności od liczby jednostek bazowych, ich wielkości mierzonej przeciętną liczbą ludności je zamieszkującej oraz łatwością migracji pomiędzy jednostkami. Stosując algorytm w warunkach polskich, postanowiono wykorzystać część wartości krytycznych określonych w pierwotnej wersji zastosowanej dla Wielkiej Brytanii (por. Coombes, Green, Openshaw, 1986). Modyfikacji poddano jedynie niektóre z nich (por. tab. 5.1). Podobne podejście zastosowali w swoich pracach również inni badacze (por. Casado-Diaz, 2000; Newell, Papps, 2001).

Aplikację algorytmu europejskiego poprzedzono wstępną oceną liczby i wielkości jednostek bazowych we wcześniejszych zastosowaniach procedury w innych krajach, oraz w Polsce (por. tab. 5.2). Stwierdzono, między innymi, że natężenie migracji pomiędzy gminami w Polsce było znacznie mniejsze niż między jednostkami bazowymi w Wielkiej Brytanii. Jest to między innymi, wynikiem większej powierzchni jednostki bazowej oraz mniejszej łatwości przejazdu. W związku z powyższym, uznano za konieczne dostosowanie wartości granicznych dwóch parametrów stosowanych w procedurze określania ośrodków centralnych³⁴. Jedną z tych wielkości była wartość graniczna parametru α_1 . Decyduje ona o uznaniu jednostki bazowej za ośrodek centralny. Parametr ten zdefiniowany jest jako stosunek liczby pracujących mieszkańców gminy (w dowolnej gminie, w tym również gminie zamieszkania), do liczby ludności pracującej na terenie rozważanej gminy (przy czym miejsce zamieszkania jest dowolne). Jako wartość graniczną w warunkach polskich przyjęto $\alpha_1=1,1$. W pierwotnej wersji algorytmu wartość graniczna parametru α_1 wynosiła aż 1,3. Jednak w warunkach polskich kryterium tego nie spełniało prawie żadne z miast (łącznie z Warszawą). Drugim z parametrów warunkujących uznanie danej jednostki bazowej za ośrodek centralny, jest parametr α_2 . Zdefiniowany jest on jako udział pracujących mieszkańców rozważanej gminy wśród podejmujących zatrudnienie na jej terenie. W tym przypadku, jako wartość graniczną przyjęto $\alpha_2=0,75$. Oznacza to, iż aby dana gmina mogła zostać uznana

³⁴ W literaturze polskiej dotyczącej ośrodków centralnych zwraca się uwagę na fakt, że definicja wielkiego miasta powinna być dostosowana do realiów polskich, ponieważ część kryteriów przyjmowanych w literaturze zagranicznej jest tak restrykcyjna, że nie spełnia ich żadne polskie miasto (por. Jałowicki, 2000).

za ośrodek centralny, liczba jej mieszkańców pracujących na terenie gminy zamieszkania, powinna przekraczać 75%. W oryginalnej wersji algorytmu wartość krytyczna parametru α_2 została ustalona na poziomie $\alpha_2=0,55$ (por. tab. 5.1). W Polsce jednak, w przeważającej mierze, mieszkańcy dużych miast poszukują pracy raczej na terenie swojego miasta.

Tabela 5.1. Wartości krytyczne parametrów europejskiego algorytmu regionalizacji w wersji pierwotnej, opracowanej dla Wielkiej Brytanii oraz zastosowanej w warunkach polskich dla roku 2006

Wartość	Parametr									
	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
pierwotna	1,3	0,55	0,5	0,1	0,01	0,002	2000	0,75	0,625	0,9267
zastosowana	1,1	0,75	0,5	0,1	0,01	0,002	2000	0,75	0,625	0,9267

Źródło: Coombes, Green, Openshaw, (1986) oraz opracowanie własne

Tabela 5.2. Liczba jednostek bazowych oraz wyznaczonych lokalnych rynków pracy w Polsce w roku 2006 na tle innych krajów

Charakterystyka wyników delimitacji	Kraj		
	Wielka Brytania	Nowa Zelandia	Polska
liczba obszarów bazowych	10102	1716	3062
liczba otrzymanych lokalnych rynków pracy	322	140	222

Źródło: Papps, Newell (2002) oraz opracowanie własne

Zastosowanie, europejskiego algorytmu regionalizacji (z uwzględnieniem powyższych zmian), pozwoliło skonstruować 222 lokalne rynki pracy na podstawie danych z 3062 gmin (z podziałem na część miejską i wiejską). Otrzymano mniej lokalnych rynków pracy niż w Wielkiej Brytanii, jednak więcej niż w Nowej Zelandii (por. tab. 5.2). Ich przestrzenne rozmieszczenie wraz z ośrodkami centralnymi przedstawiono na wykresie mapowym (por. wykr. 5.1) oraz w postaci szczegółowej listy załączonej w aneksie tabelarycznym. Zgodnie z ideą europejskiego algorytmu regionalizacji, każdy lokalny rynek pracy skonstruowany został wokół ośrodka centralnego stymulującego rozwój regionu.



Wykres 5.1. Lokalne rynki pracy z ośrodkami centralnymi według europejskiego algorytmu regionalizacji, Polska 2006

Źródło: Opracowanie własne

Oceniając wyniki zastosowania europejskiego algorytmu regionalizacji, analizie poddano wielkość wyznaczonych lokalnych rynków pracy mierzoną zarówno poprzez liczbę gmin wchodzących w ich skład jak i powierzchnię w kilometrach kwadratowych oraz liczbę ludności zatrudnionej na ich terenie i liczbę mieszkańców w wieku produkcyjnym (por. tab. 5.3).

Tabela 5.3. Charakterystyki wielkości lokalnych rynków pracy, europejski algorytm regionalizacji, Polska, 2006

Miara wielkości	Charakterystyka statystyczna					
	min	Q ₁	Q ₂	\bar{X}	Q ₃	max
Liczba gmin	1	6	9	13,77	15,75	130
Powierzchnia w km ²	15	581,5	966	1406	1518	13150
Liczba ludności pracującej	242	5 313	12 930	38 910	27 940	1 090 000
Liczba ludności w wieku produkcyjnym	1 761	23 870	50 350	111 300	96 210	2 037 000

Zródło: Opracowanie własne

Średnia wielkość lokalnego rynku pracy mierzona liczbą gmin wynosiła 14. Jednocześnie wielkość połowy spośród wyodrębnionych rynków pracy liczyła 9 gmin. Dokładnie 75% lokalnych rynków pracy składało się co najwyżej z 16 gmin, zaś tylko 25% z nich liczyło 16 lub więcej jednostek bazowych. Podobnym rozkładem charakteryzowały się również powierzchnia wyznaczonych lokalnych rynków pracy oraz liczba ludności, zarówno zatrudnionej jak i mieszkającej na terenie wyodrębnionych obszarów. Wyniki te świadczą o silnej asymetrii prawostronnej rozkładu lokalnych rynków pracy według wielkości. Otrzymano wiele mniejszych lokalnych rynków pracy, zaś największe liczą dużo więcej gmin niż sugeruje średnia. Wynika to z siły oddziaływania największych ośrodków miejskich, które skłaniają do dojazdu do pracy nawet w przypadku dużych odległości. Fakt ten uzasadnić można większą dostępnością miejsc pracy w dużych miastach, często wyższym wynagrodzeniem oraz znacznie korzystniejszymi warunkami zatrudnienia. Ważną przyczyną jest również, lepsza łatwość dojazdu do dużych miast niż między gminami peryferyjnymi. Z kolei ludność zamieszkująca w lokalizacji utrudniającej zatrudnienie w największych ośrodkach miejskich, często nie znajduje motywacji do podejmowania wysiłku codziennego dojazdu na dużej odległości. Osoby te łatwiej podejmują decyzję o dojazdach do pracy do sąsiednich gmin, pomimo mniej korzystnych warunków zatrudnienia.

Największy, pod względem liczby jednostek, lokalny rynek pracy składał się ze 130 gmin. Był to lokalny rynek pracy, którego rozwój stymulowała gmina miasto Warszawa (por. tab. 5.4). Wielkość tego lokalnego rynku pracy wynika z silnego oddziaływania Warszawy na otaczający region. Drugi pod względem wielkości lokalny rynek pracy skonstruowano wokół Poznania. Składa się on z 97 gmin. Jego duża wielkość, podobnie jak w przypadku Warszawy, wynika z silnego oddziaływania ośrodka centralnego. Trzecim pod względem wielkości mierzonej liczbą gmin, lokalnym rynkiem

pracy był w 2006 r. rynek stymulowany przez Białystok. Liczy on 96 gmin, czyli o jedną mniej niż w przypadku Poznania. W tym przypadku siła oddziaływania ośrodka centralnego jest wyraźnie niższa. Natężenie dojazdów do pracy w województwie podlaskim, z którym w dużym stopniu pokrywa się wskazany lokalny rynek pracy, jest najniższe w całej Polsce (por. Dojazdy do pracy w Polsce, red K. Kruszka, 2010). Duży rozmiar lokalnego rynku pracy Białegostoku, można uzasadniać tym, że w najbliższym sąsiedztwie nie ma innych miast, mogących pełnić rolę ośrodków centralnych. Większa liczba silnych ośrodków gospodarczych, potencjalnie spowodowałaby podział tego rynku. Dwadzieścia największych pod względem liczby gmin lokalnych rynków pracy wymieniono w tabeli 5.4.

Tabela 5.4. Ośrodki centralne stymulujące rozwój 20 największych pod względem liczby jednostek bazowych lokalnych rynków pracy, Polska, 2006

Ośrodek centralny	Wielkość lokalnego rynku pracy			
	liczba gmin	łącna powierzchnia	Liczba osób pracujących	Liczba ludności w wieku produkcyjnym
Warszawa	130	11050	987139	2031437
Poznań	97	8400	423059	1001033
Białystok	96	13146	160170	592072
Gdańsk	66	6669	260501	767426
Bydgoszcz	62	6753	215942	642343
Lublin	60	6603	156293	513744
Katowice	54	3294	563876	1432108
Łódź	52	4362	294686	828967
Kraków	52	3282	340384	874147
Opole	49	5023	128598	409207
Lubawa	47	7466	105029	313013
Wrocław	42	4476	268983	679694
Szczecin	42	5556	165886	495930
Płock	40	4702	71623	239112
Rzeszów	39	2669	120055	342644
Polkowice	36	3237	88533	248727
Dobre Miasto (województwo warmińsko-mazurskie, powiat olsztyński)	32	5506	41757	171180
Kleszczów (województwo łódzkie, powiat bełchatowski)	30	2663	55592	164279
Kielce	29	2724	104105	334156
Bielsko-Biała	28	1210	128475	350788

Źródło: Opracowanie własne

Należy przy tym podkreślić, że wielkość lokalnego rynku pracy, mierzona liczbą gmin, z uwzględnieniem podziału na część miejską i wiejską, nie zawsze pokrywa się z jego wielkością rozumianą, jako powierzchnia obszaru danego rynku. Ze względu na wielkość mierzoną powierzchnią, ranking lokalnych rynków pracy byłby nieco inny. Największy pod względem tego kryterium jest bowiem lokalny rynek pracy skupiony wokół Białegostoku, nieco mniejszy zaś rynek wyodrębniono wokół Warszawy, rynek Poznania plasuje się na trzeciej pozycji (por. tab. 5.4). Istotną zmianę zauważono również w przypadku lokalnych rynków pracy skupionych wokół Opola, Lubawy, Wrocławia, Szczecina, Płocka czy Dobrego Miasta. W skład tych rynków pracy wchodzi stosunkowo mało gmin, jednak większość z nich charakteryzuje się stosunkowo dużą powierzchnią. Dzięki temu rynki te są większe pod względem powierzchni od lokalnych rynków pracy skupionych wokół Katowic, Łodzi czy Krakowa, w których skład wchodzi więcej, jednak mniejszych, gmin (por. tab. 5.4).

Jeszcze inny byłby ranking wyznaczonych lokalnych rynków pracy ze względu na liczbę osób pracujących na ich terenie. Największy również pod tym względem jest lokalny rynek pracy, którego ośrodkiem centralnym jest Warszawa. Na drugiej pozycji znalazłby się jednak rynek skupiony wokół Katowic, który charakteryzuje się stosunkowo małą powierzchnią i jest dopiero siódmy pod względem liczby gmin wchodzących w jego skład. Z kolei lokalny rynek pracy skupiony wokół Białegostoku, choć znajdował się na trzeciej i drugiej pozycji w poprzednich rankingach, jest dopiero dziesiąty ze względu na liczbę osób pracujących na jego terenie. Podobną kolejność lokalnych rynków pracy uzyskałoby porównując liczbę mieszkańców w wieku produkcyjnym (por. tab. 5.4).

W obszarach peryferyjnych, mocno oddalonych od dużych miast wyodrębniono wiele małych lokalnych rynków pracy. Gminy Zakrzewo z województwa kujawsko-pomorskiego i Cisna z województwa podkarpackiego uzyskały status ośrodków centralnych, jednak nie przyłączono do nich żadnych innych gmin, w związku z czym tworzą one dwa lokalne rynki pracy złożone z jednej gminy każdy. Wyodrębniono również 12 lokalnych rynków pracy w skład których wchodziły po dwie gminy oraz 20 lokalnych rynków pracy złożonych z trzech gmin.

Dyskusyjną kwestią może być definiowanie oraz łączenie w trakcie działania algorytmu ośrodków centralnych. Na przykład w trakcie działania zastosowanej procedury, jako ośrodek centralny lokalnego rynku pracy, do którego należy Olsztyn uznane zostało

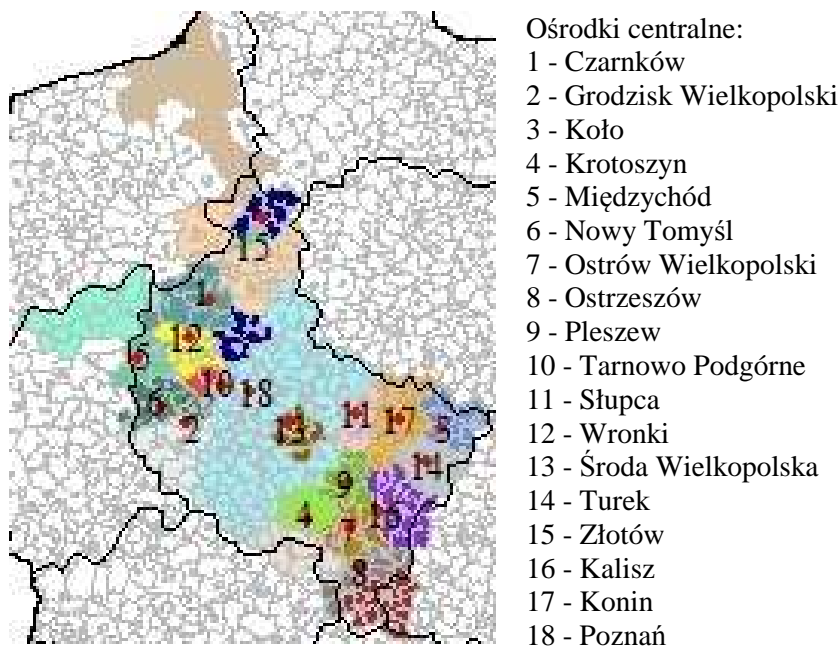
położone w pobliżu Dobre Miasto. Niewątpliwie wpływ na to miało utworzenie w 1997 r. w Dobrym Mieście, podstrefy Warmińsko - Mazurskiej Specjalnej Strefy Ekonomicznej³⁵ (WMSSE). Głównym celem utworzenia strefy było przyspieszenie rozwoju gospodarczego Dobrego Miasta oraz stworzenie inwestorom korzystnych warunków finansowo-podatkowych³⁶. Powyższy przykład wskazuje, że rozwój części otrzymanych lokalnych rynków pracy stymulowany jest przez więcej niż jedno miasto. Zgodnie z zasadami algorytmu w takim przypadku, tylko jedno z tych miast uzyskuje status ośrodka centralnego. Wybór tego miasta może być kontrowersyjny.

W celu bardziej wnikliwej analizy wyników przeprowadzonej delimitacji obserwacje postanowiono ograniczyć do gmin należących do jednego województwa. W tym celu wybrano Wielkopolskę, na terenie której, znajduje się drugi pod względem liczby gmin lokalny rynek pracy. W rozważaniach uwzględniono wszystkie gminy należące do lokalnych rynków pracy, w których chociaż jedna z gmin leżała na terenie Wielkopolski (por. wykr. 5.2).

Na terenie województwa wielkopolskiego wyodrębniono 18 ośrodków centralnych (por. wykr. 5.2). W skład lokalnych rynków pracy stymulowanych przez nie wchodziły głównie gminy z Wielkopolski, ale także niektóre gminy z województw sąsiednich (por. wykr. 5.2). Z drugiej strony część gmin z Wielkopolski przydzielona została do stref wpływów miast spoza tego województwa. Tak więc, granice wyznaczonych lokalnych rynków pracy nie zawsze mieściły się w ramach województwa. Wydaje się naturalnym, że ludność mieszkająca blisko granicy województwa, podejmuje dojazdy do pracy do ośrodków centralnych położonych poza jego terenem. Powyższa obserwacja skłania do sformułowania postulatu, aby delimitację lokalnych rynków pracy przeprowadzać wyłącznie dla całego kraju łącznie. Ograniczenie do jednego województwa czy regionu może spowodować zniekształcenie wyznaczonych lokalnych rynków pracy.

³⁵ Warmińsko-Mazurska Specjalna Strefa Ekonomiczna została ustanowiona na okres 20 lat Rozporządzeniem Rady Ministrów z dnia 9 września 1997 r. w sprawie ustanowienia warmińsko-mazurskiej specjalnej strefy ekonomicznej. W pierwotnym kształcie Strefa obejmowała obszar 334,96 ha gruntów zlokalizowanych w 3 podstrefach – Bartoszyce, Dobre Miasto i Szczytno. Podstrefa Olsztyn została utworzona w październiku 2004 r.

³⁶ Inwestorzy podejmujący działalność gospodarczą na terenie WMSSE, po spełnieniu określonych warunków mogą uzyskać tzw. pomoc regionalną na wspieranie nowych inwestycji (np. 50% całkowitych kosztów inwestycji lub 50% dwuletnich kosztów pracy nowo zatrudnionych pracowników). Na obszarze WMSSE funkcjonuje 7 podmiotów gospodarczych, które zatrudniają około 1000 osób. Na terenie Gminy Dobre Miasto funkcjonuje ponad 900 przedsiębiorców (por. oficjalny serwis internetowy Dobrego Miasta <http://www.dobremiasto.com.pl/index.php?pageId=9>).



Wykres 5.2. Lokalne rynki pracy wraz z ośrodkami centralnymi według europejskiego algorytmu regionalizacji, Wielkopolska 2006
 Źródło: Opracowanie własne

Lokalny rynek pracy stymulowany przez Poznań był zdecydowanie największy w tym regionie. Oddziaływanie Poznania nie miało jednak jednakowej siły we wszystkich kierunkach (por. wyk. 5.3). W niewielkiej odległości na zachód od Poznania, położone są inne ośrodki centralne, takie jak Tarnowo Podgórne, Wronki, Grodzisk Wielkopolski czy Nowy Tomyśl, silnie oddziałujące na gminy z ich otoczenia (por. wyk. 5.2). Z kolei na południowy wschód oddziaływanie Poznania okazało się znacznie silniejsze. Wyrwę w strefie wpływów Poznania z tej strony spowodowało jedynie wyodrębnienie lokalnego rynku pracy skupionego wokół Środy Wielkopolskiej (por. wyk. 5.2).



Wykres 5.3. Lokalne rynki pracy Poznania według europejskiego algorytmu regionalizacji, Wielkopolska 2006
 Źródło: Opracowanie własne

5.1.2. Zastosowanie algorytmu Kristensena

Zdecydowanie mniej złożony od europejskiego algorytmu regionalizacji Coombes'a, jest algorytm indukcyjny Kristensena. Przede wszystkim, konstrukcja lokalnych rynków pracy

nie odbywa się wokół wyodrębnionych ośrodków centralnych. W procedurze Kristensena, agregacja lokalnych rynków pracy odbywa się w sposób hierarchiczny. Szczegółowy opis metody delimitacji Kristensena, zamieszczono w podrozdziale 2.4.

Podobnie, jak w przypadku aplikacji europejskiego algorytmu regionalizacji, również w podejściu Kristensena, w pierwszej kolejności podjęto próbę dopasowania obu parametrów α_1 i α_2 (por. tab. 5.5). Pierwszy z nich jest wartością graniczną ilorazu alokacji, będącego charakterystyką względną przepływu pomiędzy parą gmin. Wielkość tego ilorazu powinna zostać przekroczone, jeżeli analizowany przepływ ten ma zostać uznany za istotny. Kryterium ilorazu alokacji zostało zaostrzone w porównaniu z wersją stosowaną w Danii. Wynika to z faktu, że w warunkach duńskich podziałowi na lokalne rynki pracy poddawano tylko 275 jednostek bazowych (por. tab. 5.6) w związku z czym wartości przepływów pomiędzy nimi były w naturalny sposób większe, niż w przypadku aż 3062 gmin w Polsce. Drugi parametr określa wartość, o którą stopniowo obniża się podobieństwo wymagane przy łączeniu jednostek bazowych. W warunkach polskich wartość α_2 została zwiększona, ponieważ zbyt niski poziom tego parametru prowadził do coraz większego rozdrobnienia. Wynika to ze schematu działania procedury, w ramach której, po obniżeniu podobieństwa wymaganego przy łączeniu gmin nie ma możliwości rozszerzania wcześniej wyodrębnionych lokalnych rynków pracy, a jedynie tworzenie nowych.

Tabela 5.5. Parametry algorytmu w wersji pierwotnej opracowanej dla Danii oraz zastosowanej w warunkach polskich

Wartość	Parametr	
	α_1	α_2
pierwotna	0,75	0,05
zastosowana	1	0,2

Źródło: Opracowanie własne

Tabela 5.6. Charakterystyka wyników delimitacji według algorytmu Kristensena w Danii i Polsce

Charakterystyka wyników delimitacji	Kraj	
	Dania	Polska
Liczba obszarów bazowych	275	3062
Liczba otrzymanych lokalnych rynków pracy	43	868

Źródło: Opracowanie własne

Niestety, pomimo wielu zastosowanych kombinacji wartości parametrów ustalanych na wstępie działania procedury nie udało się uzyskać mniejszej liczby, aniżeli 868 lokalnych rynków pracy (por. tab. 5.6). Ich przestrzenne rozmieszczenie przedstawiono na wykresie mapowym (por. wykr. 5.4) oraz w postaci szczegółowej listy załączonej w aneksie tabelarycznym.



Wykres 5.4. Lokalne rynki pracy według algorytmu Kristensena, Polska 2006

Źródło: Opracowanie własne

Ponad połowa z utworzonych lokalnych rynków, składa się z dwóch jednostek bazowych (por. wyk. 5.4, tab. 5.7). Jednocześnie 75% wyodrębnionych rynków składa się z nie więcej niż trzech gmin i charakteryzuje się powierzchnią o ponad połowę mniejszą od przeciętnego powiatu (por. tab. 5.7). W porównaniu z wcześniej dyskutowaną przeciętną wielkością lokalnego rynku pracy zbliżoną do powiatu wydaje się, że rozdrobnienie takie nie może być uznane za uzasadnione. Największy lokalny rynek skupiony wokół Warszawy składa się z zaledwie 48 gmin i charakteryzuje się powierzchnią 3594 km², mniej niż milionem osób pracujących na jego terenie i około milionem mieszkańców w wieku produkcyjnym (por. tab. 5.7). Obszar ten nie odzwierciedla więc strefy wpływów miasta Warszawy.

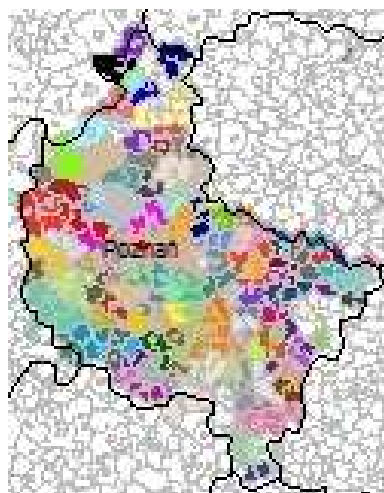
Tabela 5.7. Charakterystyki wielkości lokalnych rynków pracy, algorytm Kristensena, Polska, 2006

Miara wielkości	Charakterystyka statystyczna					
	min	Q ₁	Q ₂	\bar{X}	Q ₃	max
Liczba gmin	2	2	2	3,528	3	48
Powierzchnia w km ²	7	145	215	264,7	317,8	3594
Liczba ludności pracującej	127	888	1693	6708	4549	773200
Liczba ludności w wieku produkcyjnym	1664	6321	9682	20570	17160	1150000

Zródło: Opracowanie własne

Również wokół innych największych miast polskich nie występują widoczne strefy wpływów. Dla przykładu przeanalizowano lokalne rynki pracy wyodrębnione w Wielkopolsce, ze szczególnym uwzględnieniem lokalnego rynku pracy wyznaczonego wokół Poznania. Wyodrębnione lokalne rynki pracy są z reguły spójne, jednak ich rozkład wydaje się nie odpowiadać faktycznemu oddziaływaniu poszczególnych miast, co widoczne jest na przykład w przypadku Poznania, który należy do lokalnego rynku pracy składającego się z zaledwie ośmiu gmin (por. wyk. 5.5).

Wyniki te uznano za niezadowolające stwierdzając, że algorytm Kristensena nie nadaje się do delimitacji lokalnych rynków pracy w Polsce. Jedną z przyczyn takiego rezultatu może być fakt, że liczba i przeciętna wielkość jednostek bazowych w Polsce i w Danii, dla której algorytm został oryginalnie zaproponowany, różniła się w znacznym stopniu (por. tab. 5.6).



Wykres 5.5. Lokalne rynki pracy według algorytmu Kristensena, Wielkopolska 2006

Źródło: Opracowanie własne

Ponadto, zgodnie ze schematem działania algorytmu, na określonym poziomie podobieństwa, może nie być możliwości dołączenia kolejnej jednostki bazowej do już istniejącego lokalnego rynku pracy. Zauważmy, że jednocześnie, na tym samym poziomie podobieństwa, istnieje możliwość zbudowania nowego lokalnego rynku pracy. Wobec powyższego, procedura ta wydaje się być podejściem dyskusyjnym, a wręcz kontrowersyjnym.

5.1.3. Zastosowanie podejścia taksonomicznego

W następnej kolejności podjęto próbę zastosowania autorskiego podejścia taksonomicznego. Zostało ono szczegółowo przedstawione w podrozdziale 2.5. Wyodrębniając lokalne rynki pracy, zastosowano algorytm z mocnym jak i również słabym kryterium łączenia.

W wyniku zastosowania podejścia taksonomicznego z mocnym kryterium łączenia otrzymano 239 lokalnych rynków pracy. Natomiast przyjęcie słabego kryterium, pozwoliło uzyskać 134 lokalne rynki pracy. Ich strukturę przedstawiono na wykresie mapowym (por. odpowiednio wykr. 5.6 i 5.7) oraz w postaci szczegółowej listy załączonej w aneksie tabelarycznym. W przypadku przyjęcia słabego kryterium łączenia zaobserwowano równomierny wzrost wielkości lokalnych rynków pracy, w stosunku do wyników delimitacji z mocnym kryterium łączenia. Mianowicie, o około połowę zwiększyły się wszystkie kwartyle wielkości lokalnych rynków pracy mierzonej zarówno

liczbą gmin wchodzących w ich skład jak i ich łączną powierzchnią oraz liczbą osób pracujących i mieszkańców w wieku produkcyjnym (por. tab. 5.8).



Wykres 5.6. Lokalne rynki pracy, podejście taksonomiczne, mocne kryterium łączenia, Polska 2006

Źródło: Opracowanie własne

W wyniku osłabienia kryterium łączenia, część wcześniej występujących oddzielnie lokalnych rynków pracy została połączona. Częstą sytuacją było „wchłanianie” przez większe lokalne rynki pracy mniejszych, np. w przypadku lokalnego rynku pracy Warszawy. Jednak łączono także lokalne rynki pracy o zbliżonej wielkości (por. wykr. 5.6 i 5.7), np. w przypadku połączenia lokalnych rynków pracy Sanoka i Ustrzyk Dolnych/Leska. Ponadto część lokalnych rynków pracy, otrzymanych w wyniku przyjęcia mocnego kryterium łączenia, nie została w całości połączona w przypadku zastosowania słabego kryterium łączenia. Sytuacja taka mogła mieć miejsce, wówczas,

gdy wstępne lokalne rynki pracy przydzielono według słabszego kryterium łączenia do innych wstępnych lokalnych rynków pracy i w dalszej części działania algorytmu nie było już potrzeby ich łączenia.



Wykres 5.7. Lokalne rynki pracy, podejście taksonomiczne, słabe kryterium łączenia, Polska 2006

Źródło: Opracowanie własne

Rozkład wielkości lokalnych rynków pracy, wyznaczonych przy pomocy obu wariantów podejścia taksonomicznego, podobnie jak w przypadku zastosowania europejskiego algorytmu regionalizacji, jest prawostronnie asymetryczny. Większość z wyodrębnionych rynków jest stosunkowo niewielka, zarówno pod względem liczby gmin jak również ich powierzchni, liczby pracowników czy mieszkańców w wieku produkcyjnym. Z kolei lokalne rynki pracy skonstruowane wokół największych miast, zarówno w przypadku

zastosowania mocnego jak i słabego kryterium łączenia, były znacznie większe od pozostałych (por. tab. 5.8, 5.9 i 5.10). Ponadto średnia wielkość wyznaczonych lokalnych rynków pracy, mierzona liczbą gmin wchodzących w ich skład oraz polem powierzchni, bardziej zbliżona jest do trzeciego kwartyla niż do mediany. Asymetria ta staje się jeszcze wyraźniejsza, jeżeli pod uwagę wziąć wielkość lokalnych rynków pracy mierzoną przy pomocy liczby osób pracujących na ich terenie lub mieszkańców w wieku produkcyjnym, ponieważ dla tak określonej wielkości średnia przewyższa trzeci kwartył.

W obydwóch przypadkach, słabego i mocnego kryterium, największym okazał się lokalny rynek pracy skupiony wokół Warszawy. Jest on znacznie większy od pozostałych, jednak w przypadku zmiany mocnego kryterium łączenia na słabe, wielkość tego rynku zwiększyła się tylko z 209 na 230 gmin (por. tab. 5.9 i 5.10). W przypadku obu kryteriów, kształt lokalnego rynku pracy, którego rozwój stymuluje Warszawa jest zbliżony, ponieważ lokalny rynek pracy wyznaczony w oparciu o mocne kryterium grupowania zawiera się w lokalnym rynku pracy wyodrębnionym zgodnie ze słabym kryterium łączenia (por. wyk. 5.6 i 5.7). W przypadku przyjęcia mocnego kryterium, obszar ten dodatkowo rozdzielony zostaje pomiędzy strefy wpływów Ciechanowa (6 gmin), Przasnysza (7 gmin) i Makowa Mazowieckiego (8 gmin), które przy założeniu słabego kryterium łączenia wchodziły w skład lokalnego rynku pracy Warszawy.

Tabela 5.8. Charakterystyki wielkości lokalnych rynków pracy, podejście taksonomiczne z mocnym i słabym kryterium łączenia, Polska, 2006

Kryterium łączenia	Miara wielkości	min	Q ₁	Q ₂	\bar{X}	Q ₃	max
Mocne	liczba gmin	2	6	8	12,76	13	209
	powierzchnia w km ²	75	585	816	1303	1384	18820
	ludność pracująca	1369	6872	12020	33640	22580	1054000
	ludność w wieku produkcyjnym	6925	27940	43430	102400	83720	2342000
Słabe	liczba gmin	2	10,25	15	22,75	26,25	230
	powierzchnia w km ²	82	922	1582	2323	2660	21280
	ludność pracująca	1778	13340	23940	59990	45940	1079000
	ludność w wieku produkcyjnym	8356	49740	87570	182600	154400	2439000

Źródło: Opracowanie własne.

W procedurze z mocnym kryterium łączenia, wyodrębniono dwa drugie co do wielkości mierzonej liczbą gmin, lokalne rynki pracy. Są to rynek pracy skupiony wokół Lublina i Katowic. W skład każdego z nich wchodzi 75 gmin. Warto zauważyć, że lokalny rynek pracy, do którego należał Lublin, charakteryzuje się przy tym niemal dwukrotnie większą powierzchnią od rynku, do którego należały Katowice. Jednak jednocześnie rynek pracy skupiony wokół Lublina liczył około trzykrotnie mniej osób pracujących i mieszkańców w wieku produkcyjnym. Lokalny rynek pracy wokół Poznania liczył o jedną gminę mniej. Jednocześnie rynek ten jest zbliżony jest powierzchnią do rynku Lublina i charakteryzuje się liczbą osób pracujących i mieszkańców w wieku produkcyjną większą aniżeli rynek Lublina, jednak mniejszą niż rynek Katowic. Lokalny rynek pracy skupiony wokół Krakowa składa się z 69 gmin. Jego powierzchnia zbliżona jest do niewielkiego pod tym względem rynku Katowic, zaś charakterystyki zaludnienia bardziej zbliżone są do rynku Poznania (por. tab. 5.9).

Tabela 5.9. Największe pod względem liczby jednostek bazowych lokalne rynki pracy według podejścia taksonomicznego z mocnym kryterium łączenia, polska, 2006

Miasto	Wielkość lokalnego rynku pracy			
	liczba gmin	łącna powierzchnia	Liczba osób pracujących	Liczba ludności w wieku produkcyjnym
Warszawa	209	18820	1054471	2341594
Lublin	75	7935	172923	592682
Katowice	75	4354	587127	1520542
Poznań	74	7165	372644	821106
Kraków	69	4417	328964	886758
Szczecin	61	7505	174086	541340
Łódź	59	4856	302236	859384
Olsztyn	53	8759	115648	352130
Białystok	52	7220	106806	364780
Wrocław	50	5367	274404	708202
Rzeszów	44	2879	121679	360647
Bydgoszcz	42	5192	166151	460919
Opole	38	3666	100886	314264
Gdańsk	37	3849	200746	556936
Tarnów	32	2279	67052	244552
Częstochowa	32	2871	104707	307467
Kielce	31	3009	88906	295376
Nowy Sącz	27	2364	53609	212823
Włocławek	27	2262	44070	156599
Zielona Góra	24	3245	60190	185515

Uwaga: Podane miasta nie pełniły roli ośrodków centralnych w procedurze delimitacji

Źródło: Opracowanie własne

Drugi, pod względem wielkości mierzonej liczbą gmin, lokalny rynek pracy wyodrębniony w procesie ze słabym kryterium łączenia, składa się z 97 gmin i skupiony jest wokół Poznania. Pod względem powierzchni przewyższają go jednak rynki, do których należy Olsztyn (70 gmin) i Białystok (69 gmin). Biorąc pod uwagę liczbę osób pracujących, za większe należy uznać lokalny rynek pracy skupiony wokół Katowic (91 gmin) (por. tab. 5.10).

Tabela 5.10. Największe pod względem liczby jednostek bazowych lokalne rynki pracy według podejścia taksonomicznego ze słabym kryterium łączenia, polska, 2006

Miasto	Wielkość lokalnego rynku pracy			
	liczba gmin	łącna powierzchnia	Liczba osób pracujących	Liczba ludności w wieku produkcyjnym
Warszawa	230	21276	1078811	2438664
Poznań	97	9118	405340	950365
Katowice	91	5267	684973	1816527
Kraków	83	5304	350764	981030
Lublin	82	8497	178070	618199
Wrocław	74	8272	313371	846885
Olsztyn	70	12170	135484	439877
Białystok	69	9553	120524	430513
Łódź	64	5373	320951	920725
Szczecin	56	6812	170065	524377
Bydgoszcz	55	6398	193767	565146
Opole	50	4694	118280	389107
Gdańsk	47	4668	214682	613303
Kielce	46	4811	105047	364983
Rzeszów	44	2879	121679	360647
Legnica / Lubin	42	3933	116174	328493
Radom	39	3622	78803	317398
Tarnów	38	2841	70851	260948
Częstochowa	38	3293	113529	346774
Piotrków Trybunalski / Kleszczów / Bełchatów	35	3406	74477	222413

Uwaga: Podane miasta nie pełniły roli ośrodków centralnych w procedurze delimitacji

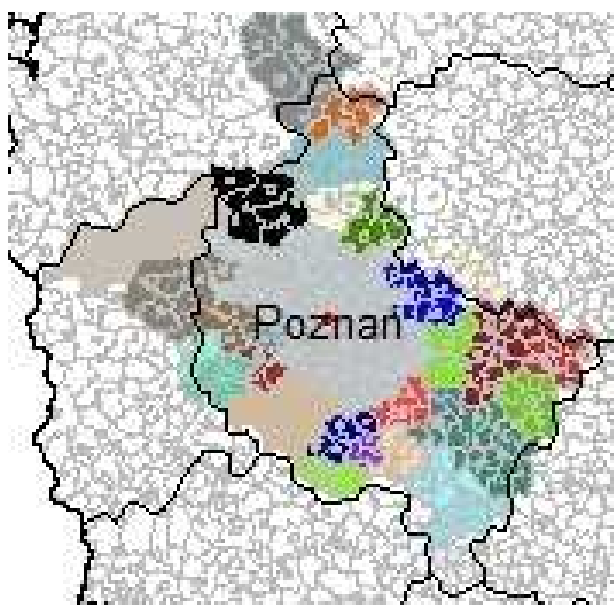
Źródło: Opracowanie własne

Należy podkreślić, że użyte powyżej nazwy lokalnych rynków pracy pochodzące od miast wchodzących w ich skład są umowne. Wybrano mianowicie największe miasto

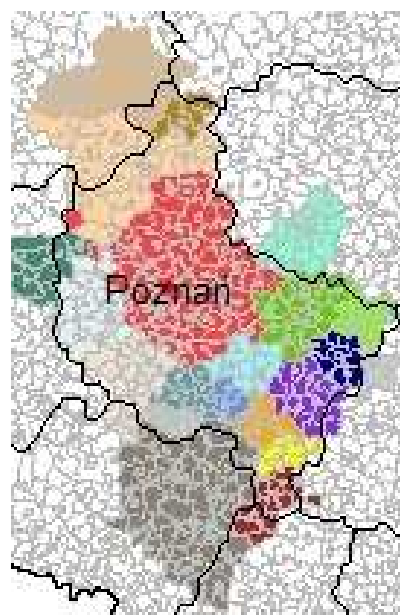
ze względu na liczbę osób pracujących na jego terenie (lub kilka, jeżeli były porównywalnej wielkości). W procesie delimitacji lokalnych rynków pracy zgodnym z podejściem taksonomicznym, wszystkie jednostki traktowane są bowiem jednakowo i żadnym z nich nie przyznaje się statusu ośrodka centralnego.

W procedurze delimitacji opartej na podejściu taksonomicznym wstępny lokalny rynek pracy powstaje w momencie połączenia dwóch gmin, zatem każdy lokalny rynek pracy, wyodrębniony zgodnie z tą metodą, musi składać się, co najmniej z dwóch. Stosując zarówno mocne jak i słabe kryterium łączenia wyznaczono tylko jeden lokalny rynek pracy złożony z dwóch gmin, były to część miejska i wiejska gminy Zawadzkie z województwa opolskiego. W procesie wykorzystującym mocne kryterium łączenia, wyodrębniono ponadto dziewięć lokalnych rynków pracy złożonych z trzech gmin. Przyjmując słabe kryterium łączenia nie wyznaczono żadnego lokalnego rynku pracy złożonego z trzech gmin.

W celu lepszego zilustrowania różnic pomiędzy procedurami wykorzystującymi słabe i mocne kryterium łączenia, przedstawiono uzyskane w wyniku ich zastosowania delimitacje dla Wielkopolski. Lokalne rynki pracy w Wielkopolsce zobrazowano na poniższych wykresach mapowych 5.8 i 5.9.



Wykres 5.8. Lokalne rynki pracy według podejścia taksonomicznego z mocnym kryterium łączenia, Wielkopolska 2006
Źródło: Opracowanie własne



Wykres 5.9. Lokalne rynki pracy według podejścia taksonomicznego ze słabym kryterium łączenia, Wielkopolska 2006
Źródło: Opracowanie własne

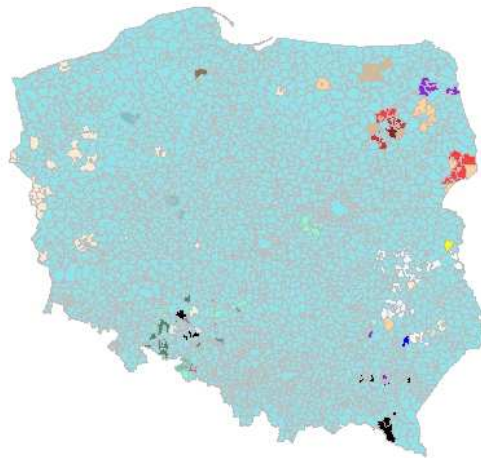
Istotną różnicę zauważyć można porównując lokalny rynek pracy skupiony wokół Poznania.

W przypadku mocnego kryterium, łącznie składa się on z 74 gmin i sięga dość daleko na północny zachód (por. wyk. 5.8). Przy zastosowaniu słabego kryterium łączenia, zgodnie z ogólną tendencją, liczba gmin wchodzących w jego skład wzrosła do 95. Jednak tylko 69 gmin należących do rynku wyodrębnionego według mocnego kryterium, należało także do rynku wyodrębnionego według słabego kryterium łączenia. Powiększenie nastąpiło kosztem wchłonięcia trzech mniejszych lokalnych rynków pracy, skupionych wokół Gniezna, Chodzieży i Wągrowca, które położone są na północny wschód od Poznania (por. wyk. 5.9). Jednocześnie jednak część gmin należąca do strefy wpływów Poznania, przy uwzględnieniu mocnego kryterium łączenia, zgodnie z procedurą uwzględniającą słabe kryterium, przyporządkowana została do innych lokalnych rynków pracy. Mianowicie część miejska i wiejska Wroniek przydzielona została do lokalnego rynku pracy Piły, a Kwilcz oraz część miejska i wiejska Pniew zasiliły lokalny rynek pracy Grodziska Wielkopolskiego/Nowego Tomysła (choć przy zastosowaniu słabego kryterium łączenia, gminy te należały do lokalnego rynku pracy skupionego wokół Poznania) (por. wyk. 5.8 i 5.9).

Ponadto część miejską i wiejską Pyzdr dołączono do lokalnego rynku pracy Słupcy, a część miejską i wiejską Buku do powstałego w ten sposób lokalnego rynku pracy Opalenicy, która z kolei w przypadku słabego kryterium łączenia wchodziła w skład lokalnego rynku pracy Grodziska Wielkopolskiego/Nowego Tomysła.

5.2. Próba delimitacji lokalnych rynków pracy w Polsce dla roku 2011

W podrozdziale 5.2 podjęto próbę wyodrębnienia lokalnych rynków pracy w Polsce, aktualnych dla roku 2011. W tym celu posłużono się macierzą migracji oszacowaną w rozdziale czwartym. W pierwszej kolejności zastosowano europejski algorytm regionalizacji. Otrzymane wyniki przedstawiono na wykresie mapowym 5.10.



Wykres 5.10. Lokalne rynki pracy wraz z ośrodkami centralnymi według europejskiego algorytmu regionalizacji, Polska 2011, na podstawie „surowych” oszacowań
 Źródło: Opracowanie własne

Posługując się „surowymi” oszacowaniami otrzymano podział, w którym większość gmin skupiona była w jednym „lokalnym” rynku pracy, pokrywającym niemal całą Polskę (por. wyk. 5.10). Uzyskane wyniki nie pozwoliły zidentyfikować jednostek terytorialnych, zgodnych z definicją lokalnego rynku pracy, a więc umożliwiających codzienny dojazd do pracy, bez konieczności zmiany miejsca zamieszkania. Otrzymany podział nie spełniał oczekiwań wynikających z przeprowadzanej delimitacji lokalnych rynków pracy w Polsce. Analizując sytuację, stwierdzono, że powstałe przy szacowaniu przepływów pomiędzy parami gmin błędy mają charakter systematyczny. Mianowicie przeszacowane zostały wartości dojazdów na dużych odległościach przy jednoczesnym niedoszacowaniu przepływów pomiędzy pobliskimi gminami.

W celu korekty powstałych błędów przyjęto założenie, że mają one taką samą strukturę jak błędy popełnione w oszacowaniach dla roku 2006 (oszacowania wyznaczone dla roku 2006 nie były wykorzystywane w procesie delimitacji, a jedynie przy ocenie precyzji szacunku w rozdziale 4):

$$\frac{M_{ij}^{e2011}}{M_{ij}^{t2011}} = \frac{M_{ij}^{e2006}}{M_{ij}^{t2006}} , \quad (5.1)$$

gdzie:

M_{ij}^{e2011} - nieznana wartość empiryczna przepływu pomiędzy gminą i -tą a j -tą w roku 2011,

M_{ij}^{e2006} - znana wartość empiryczna przepływu pomiędzy gminą i -tą a j -tą w roku 2006,

M_{ij}^{t2011} - wartość teoretyczna przepływu pomiędzy gminą i -tą a j -tą w roku 2011, oszacowana przy pomocy modelu wielopoziomowego skonstruowanego w rozdziale 4,

M_{ij}^{t2006} - wartość teoretyczna przepływu pomiędzy gminą i -tą a j -tą w roku 2006, oszacowana przy pomocy modelu wielopoziomowego skonstruowanego w rozdziale 4.

Na podstawie powyższego założenia dokonano przeskalowania wartości teoretycznych przepływów związanych z zatrudnieniem w roku 2011 zgodnie ze wzorem:

$$\widehat{M_{ij}^{e2011}} = M_{ij}^{t2011} \times \frac{M_{ij}^{e2006}}{M_{ij}^{t2006}} \quad . \quad (5.2)$$

Uwzględnienie takiej korekty spowodowało zmniejszenie regularności popełnianych błędów względem odległości, na której podejmowane były dojazdy. Przepływy pomiędzy gminami położonymi w najbliższym sąsiedztwie pozostały jednak nadal niedoszacowane, jednocześnie zaś część wpływów największych miast „przejęta” została przez obszary je okalające. Spowodowało to przyznanie statusu ośrodków centralnych wielu gminom nie pełniącym w rzeczywistości takiej funkcji, a jedynie leżącym w sąsiedztwie dużych miast (por. wyk. 5.11). W konsekwencji strefy wpływu tych miast podzielone zostały pomiędzy otaczające je gminy i wyznaczono w tych obszarach wiele bardzo małych lokalnych rynków pracy (por. wyk. 5.11).



Wykres 5.11. Lokalne rynki pracy wraz z ośrodkami centralnymi według europejskiego algorytmu regionalizacji, Polska 2011, oszacowania przeskalowane według wzoru (5.2)

Uwaga Gminy mające status ośrodków centralnych oznaczono czerwonymi punktami

Źródło: Opracowanie własne

Tabela 5.11. Charakterystyki wielkości lokalnych rynków pracy, europejski algorytm regionalizacji, Polska, 2011, oszacowania przeskalowane według wzoru (5.2)

Miara wielkości	Charakterystyka statystyczna					
	min	Q ₁	Q ₂	\bar{X}	Q ₃	max
Liczba gmin	1	3	6	11,02	11	223
Powierzchnia w km ²	3	213,5	503	785,3	933	13 450
Liczba ludności pracującej	133	1 806	5 662	21 830	15 710	1 068 000
Liczba ludności w wieku produkcyjnym	1 122	10 420	243 20	62 390	55 620	1 920 000

Źródło: Opracowanie własne

Ponieważ wprowadzenie omówionej korekty nie pozwoliło na uzyskanie zadowalających wyników podjęto próbę ponownego skorygowania oszacowań przepływów związanych z zatrudnieniem. W tym celu wykorzystano tę własność, że sumy elementów kolumn macierzy migracji równe są liczbom osób pracujących w gminach, które z kolei znane są również dla roku 2011. W celu nadania tej własności również oszacowanej dla 2011 macierzy migracji dokonano następującej korekty:

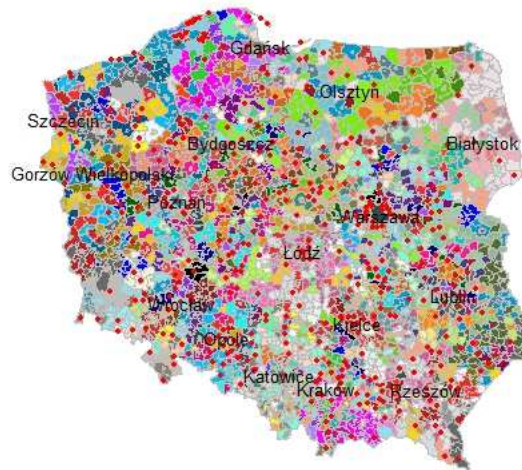
$$\widehat{M}_{ij}^{e2011} = \widehat{M}_{ij}^{2011} \times \frac{p_j^{2011}}{\sum_{i=1}^{3062} \widehat{M}_{ij}^{e2011}} \quad , \quad (5.3)$$

gdzie:

\widehat{M}_{ij}^{e2011} - skorygowana zgodnie ze wzorem (5.2) wartość oszacowania przepływu pomiędzy gminą i-tą a j-tą dla roku 2011,

p_j^{2011} - liczba osób pracujących na terenie j-tej gminy w 2011 roku.

Uwzględnienie takiej korekty spowodowało jednak jeszcze mocniejsze przeszacowanie oddziaływania gmin okalających duże miasta, co szczególnie widoczne jest na przykładzie Warszawy (por. wykr. 5.12).



Wykres 5.12. Lokalne rynki pracy wraz z ośrodkami centralnymi według europejskiego algorytmu regionalizacji, Polska 2011, oszacowania przeskalowane według wzoru (5.3)

Uwaga Gminy mające status ośrodków centralnych oznaczono czerwonymi punktami

Źródło: Opracowanie własne

Tabela 5.12. Charakterystyki wielkości lokalnych rynków pracy, europejski algorytm regionalizacji, Polska, 2011, oszacowania przeskalowane według wzoru (5.3)

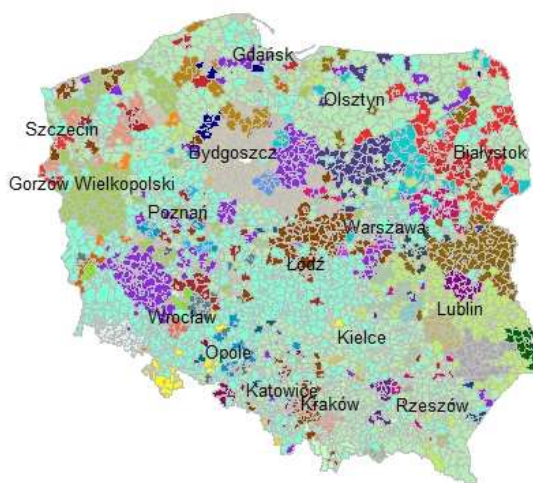
Miara wielkości	Charakterystyka statystyczna					
	min	Q ₁	Q ₂	\bar{X}	Q ₃	max
Liczba gmin	1	2	5	7,689	9	101
Powierzchnia w km ²	3	213,5	503	785,3	933	13 450
Liczba ludności pracującej	133	1 806	5 662	21 830	15 710	1 068 000
Liczba ludności w wieku produkcyjnym	1 122	10 420	24 320	62 390	55 620	1 920 000

Źródło: Opracowanie własne

W sytuacji zbyt dużej liczby ośrodków centralnych, wokół których wyznaczony powinien zostać jeden lokalny rynek pracy, podjęto próbę dostosowania odpowiednich parametrów europejskiego algorytmu regionalizacji. Jednak nawet drastyczne obniżanie wartości progowych, stosowanych w etapie drugim, w procesie łączenia wstępnie wyznaczonych (etap pierwszy) ośrodków centralnych, nie doprowadziło do masowego ich łączenia.

Podobne problemy napotkano przy próbie wyznaczenia lokalnych rynków pracy aktualnych dla roku 2011 przy pomocy podejścia taksonomicznego, zarówno z mocnym jak i słabym kryterium łączenia. Delimitację przy zastosowaniu tej metody przeprowadzono w oparciu o dane po drugiej z opisanych powyżej korekt.

Uwzględniając mocne kryterium łączenia wyznaczono zaledwie 66 lokalnych rynków pracy, przedstawionych na wykresie mapowym 5.13. Wyznaczone obszary są często niespójne i mają zdecydowanie zbyt szeroki zakres terytorialny, aby mogły być uznane za realne lokalne rynki pracy (por. wyk. 5.13).



Wykres 5.13. Lokalne rynki pracy według podejścia taksonomicznego z mocnym kryterium łączenia, Polska 2011, oszacowania przeskalowane według wzoru (5.3)

Uwaga Gminy mające status ośrodków centralnych oznaczono czerwonymi punktami
 Źródło: Opracowanie własne

Tabela 5.13. Charakterystyki wielkości lokalnych rynków pracy, podejście taksonomiczne z mocnym kryterium łączenia, Polska, 2011, oszacowania przeskalowane według wzoru (5.3)

Miara wielkości	Charakterystyka statystyczna					
	min	Q ₁	Q ₂	\bar{X}	Q ₃	max
Liczba gmin	2	3	9,5	46,23	26	812
Powierzchnia w km ²	76	273	894	4 719	2585	85 850
Liczba ludności pracującej	602	2 161	11 800	130 900	42 120	3 026 000
Liczba ludności w wieku produkcyjnym	3 798	10 460	50 600	374 200	156 500	7 895 000

Źródło: Opracowanie własne

Sytuacja, w której jednocześnie Warszawa i Kraków przyporządkowane są do tego samego lokalnego rynku pracy (por. tab. 5.14), dyskwalifikuje przeprowadzoną delimitację. W sposób oczywisty, podział taki nie zgadza się ze stanem faktycznym. W związku z tym zaniechano dalszej analizy wyników przeprowadzonej delimitacji, uznając ją za niezasadną.

Tabela 5.14. Największe pod względem liczby jednostek bazowych lokalne rynki pracy według podejścia taksonomicznego z mocnym kryterium łączenia, Polska 2011, oszacowania przeskalowane według wzoru (5.2)

Miasto	Wielkość lokalnego rynku pracy			
	liczba gmin	łącna powierzchnia	Liczba osób pracujących	Liczba ludności w wieku produkcyjnym
Warszawa / Kraków	812	85847	2393705	6633176
Poznań	686	63188	2827068	7961147
Zielona Góra	189	17038	415108	1304041
Łódź	162	13106	525239	1622750
Lublin	157	16029	287856	1072138
Bydgoszcz	145	16789	320602	1043998
Gorzów Wielkopolski	110	13723	174070	575964
Biała Podlaska	72	9348	70295	304856
Ciechanów	63	7533	63960	262727
Białystok	47	7278	111210	361569

Uwaga: Podane miasta nie pełniły roli ośrodków centralnych w procedurze delimitacji

Źródło: Opracowanie własne

Uwzględniając słabe kryterium łączenia uzyskano podział na 48 obszarów, przedstawiony na wykresie 5.14. Podobnie jak w przypadku mocnego kryterium łączenia, wyodrębnione rynki pracy są niespójne i charakteryzują się bardzo dużym zasięgiem terytorialnym.

Tabela 5.15. Charakterystyki wielkości lokalnych rynków pracy, podejście taksonomiczne ze słabym kryterium łączenia, Polska, 2011, oszacowania przeskalowane według wzoru (5.3)

Miara wielkości	Charakterystyka statystyczna					
	min	Q ₁	Q ₂	\bar{X}	Q ₃	max
Liczba gmin	2	3	7	63,56	27,75	1093
Powierzchnia w km ²	76	269,5	633,5	6 489	2 644	97 040
Liczba ludności pracującej	681	1 906	8 998	179 900	43 970	4 504 000
Liczba ludności w wieku produkcyjnym	3 798	9 977	43 450	514 500	160 200	1 1860 000

Źródło: Opracowanie własne



Wykres 5.14. Lokalne rynki pracy według podejścia taksonomicznego ze słabym kryterium łączenia, Polska 2011, oszacowania przeskalowane według wzoru (5.3)

Uwaga: Gminy mające status ośrodków centralnych oznaczono czerwonymi punktami

Źródło: Opracowanie własne

Tabela 5.16. Największe pod względem liczby jednostek bazowych lokalne rynki pracy, podejście taksonomiczne ze słabym kryterium łączenia, Polska 2011, oszacowania przeskalowane według wzoru (5.2)

Miasto	Wielkość lokalnego rynku pracy			
	liczba gmin	łącna powierzchnia	Liczba osób pracujących	Liczba ludności w wieku produkcyjnym
Kraków / Poznań	1093	97040	4173077	11910576
Warszawa	758	85446	1887272	5422907
Zielona Góra / Gorzów Wielkopolski / Ruda Śląska	322	32667	629144	2006668
Lublin	163	16549	293562	1098028
Bydgoszcz	145	16789	320602	1043998
Ciechanów	63	7533	63960	262727
Białystok	47	7278	111210	361569
Ostrołęka	40	6088	48743	218199
Skawina - miasto	38	3283	47537	198526
Chojnice	37	4811	54230	190641

Uwaga: Podane miasta nie pełniły roli ośrodków centralnych w procedurze delimitacji

Źródło: Opracowanie własne

Największy z wyznaczonych obszarów składa się aż z 1093 gmin (por. tab. 5.15). Ponadto w jego skład wchodzi Kraków i Poznań (por. tab. 5.16). Obszar ten nie jest spójny terytorialnie, i w oczywisty sposób nie nosi cech lokalnego rynku pracy. W związku z tym dalsza analiza otrzymanych wyników nie wydaje się być uzasadniona.

5.3. Podsumowanie otrzymanych wyników

Wyniki delimitacji, przeprowadzonych dla 2006 roku, przy zastosowaniu europejskiego algorytmu regionalizacji oraz podejścia taksonomicznego, są do pod wieloma względami zbliżone. W obu przypadkach otrzymano nieco ponad 200 lokalnych rynków pracy, przy czym około 25% z nich składało się z nie więcej niż 6 gmin (por. tab. 5.17). Również wartości pozostałych kwartyli były w obu przypadkach zbliżone (por. tab. 5.17).

Tabela 5.17. Charakterystyka lokalnych rynków otrzymanych przy zastosowaniu różnych algorytmów, Polska 2006

Metoda delimitacji	Liczba lokalnych rynków pracy	Charakterystyki wielkości lokalnych rynków pracy mierzonej liczbą gmin				
		<i>min</i>	Q_1	Q_2	Q_3	<i>max</i>
Europejski algorytm regionalizacji	222	1	6	9	16	128
Algorytm Kristensen'a	868	2	2	2	3	48
Podejście taksonomiczne – mocne kryterium łączenia	239	2	6	8	13	209
Podejście taksonomiczne – słabe kryterium łączenia	134	2	10,25	15	26,25	230

Źródło: Opracowanie własne

Wobec podobieństwa uzyskanych rezultatów, można sformułować pytanie, dotyczące wyboru metody delimitacji. Mianowicie, czy każdy lokalny rynek pracy powinien być skupiony wokół ośrodka centralnego? W celu rozwiązania tak postawionego dylematu postanowiono przeanalizować najistotniejsze różnice wyników otrzymanych w procedurze delimitacji przeprowadzonej przy pomocy europejskiego algorytmu regionalizacji oraz podejścia taksonomicznego. W pierwszym przypadku lokalne rynki pracy określane są wokół wyodrębnionych ośrodków centralnych, a w drugim, wszystkie gminy traktowane są jednakowo.

Największe rozbieżności odnotowano w przypadku regionów charakteryzujących się niskim natężeniem ośrodków centralnych (np. Podlasie). W algorytmie europejskim wyodrębniono w tamtym regionie jeden lokalny rynek pracy skupiony wokół

Białegostoku. Z kolei zastosowanie podejścia taksonomicznego umożliwiło wskazanie w zbliżonym obszarze około pięciu lokalnych rynków pracy (zarówno w przypadku mocnego jak i słabego kryterium łączenia). Jednocześnie wielkość lokalnego rynku pracy, do którego należał Białystok zmniejszyła się z 96 gmin wyodrębnionych w trakcie europejskiego algorytmu regionalizacji, do 52 przy zastosowaniu podejścia taksonomicznego z mocnym kryterium łączenia, oraz 69 dla słabego kryterium łączenia. Podział Podlasia na kilka mniejszych lokalnych rynków pracy zdaje się lepiej odzwierciedlać sytuację rzeczywistą.

Odwrotną sytuację zauważono na Śląsku, gdzie natężenie potencjalnych ośrodków centralnych jest bardzo duże. W przypadku zastosowania europejskiego algorytmu regionalizacji wyodrębniono w tym regionie dużo małych lokalnych rynków pracy. Natomiast w wyniku podejścia taksonomicznego, które nie uwzględnia ośrodków centralnych, otrzymano mniej większych lokalnych rynków pracy. Wynik taki uzyskano stosując zarówno mocne jak i słabe kryterium łączenia. W przypadku delimitacji przeprowadzonej przy pomocy europejskiego algorytmu regionalizacji, największy lokalny rynek pracy wokół Katowic, składał się zaledwie z 47 gmin. Natomiast przy wykorzystaniu podejścia taksonomicznego lokalny rynek pracy, do którego należą Katowice liczył 75 oraz 91 gmin, odpowiednio z mocnym i słabym kryterium łączenia. R. Krzysztofik i J. Runge określili na tym obszarze aglomerację policentryczną, obejmującą łącznie 88 gmin (por. Krzysztofik, Runge, 2011). Również w tym wypadku, wyniki delimitacji przeprowadzonych przy pomocy podejścia taksonomicznego lepiej odzwierciedlają faktyczną strukturę lokalnych rynków pracy w tym specyficznym obszarze.

Kolejną ważną kwestią jest odpowiedź na pytanie, jakiej wielkości powinien być przeciętny lokalny rynek pracy. Zagadnienie to można sformułować inaczej, mianowicie ile lokalnych rynków pracy powinien uwzględniać optymalny podział. W prowadzonych w tej dziedzinie badaniach wielu autorów utożsamia lokalny rynek pracy z powiatem (por. Gołata, 2004). Również Szumlicz (1993) w swojej pracy podaje, że lokalny rynek pracy powinien odpowiadać wielkością terenowi działania lokalnej agencji pośrednictwa pracy. Można zatem przyjąć, że lokalny rynek pracy winien składać się przeciętnie z kilku (kilkunastu) gmin, a jego wielkość powinna być zbliżona do powiatu. Na tej podstawie można przyjąć, że Polska składa się z prawie 400 lokalnych rynków pracy.

Lub, jeśli połączyć miasta na prawach powiatu (powiaty grodzkie) z otaczającymi je powiatami ziemskimi, otrzymalibyśmy około 300 lokalnych rynków pracy.

Porównując wyniki delimitacji przeprowadzonych w innych krajach, zauważyć jednak można, że przeciętna wielkość wyznaczanych lokalnych rynków pracy jest bardzo zróżnicowana. Na przykład w Grecji, w procesie delimitacji przeprowadzonej przy pomocy procedury iteracyjnej, obszar kraju podzielono na aż 667 lokalnych rynków pracy

(por. Prodromidis, 2010). W badaniu tym wykorzystano dane dotyczące dojazdów do pracy ze Spisu Ludności w 2001 roku dla 1034 jednostek bazowych, co oznacza, że przeciętny lokalny rynek pracy składał się mniej niż dwóch jednostek bazowych i miał powierzchnię około 200 km². Z kolei w Słowenii, na podstawie dostępnych danych o przepływach, związanych z zatrudnieniem, między 192 jednostkami bazowymi, wyodrębniono 31 lokalnych rynków pracy (por. Drobne, Konjar, Lisec, Pichler Milanović, Zavodnik Lamovšek, 2010). Przeciętny lokalny rynek pracy w Słowenii składał się więc z nieco ponad 6 jednostek bazowych i charakteryzował się powierzchnią bliską 650 km².

Optymalna liczba lokalnych rynków pracy, powinna wynikać z celu przeprowadzanej delimitacji. Sprecyzowanie potrzeb gospodarczych, społecznych i administracyjnych w tym względzie, pozwoli na wyodrębnienie czynników determinujących przyjęcie określonych rozwiązań. Uwzględniając złożoność zagadnienia oraz jego dyskusyjny charakter, pozostawiono je poza zakresem rozważań jako otwarty problem badawczy.

Podsumowując, zarówno w przypadku zastosowania europejskiego algorytmu regionalizacji jak i podejścia taksonomicznego z mocnym kryterium łączenia, dla danych aktualnych dla roku 2006, otrzymano zbliżone pod wieloma względami wyniki. Największe rozbieżności odnotowano w regionach o nietypowym natężeniu ośrodków centralnych. W przypadku zastosowania każdego z algorytmów, liczba wyznaczonych lokalnych rynków pracy była zbliżona do liczby powiatów w Polsce. Podział administracyjny kraju na powiaty, ze względu na brak danych, często utożsamiano z lokalnym rynkiem pracy. W przypadku przyjęcia słabego kryterium łączenia, uniknięto z kolei tworzenia wielu małych rynków pracy złożonych z gmin peryferyjnych. Wydaje się więc, że wybór metody delimitacji powinien wynikać z odpowiedzi na pytania,

czy każdy lokalny rynek pracy należy konstruować wokół ośrodka centralnego oraz pytania o przeciętną wielkość lokalnego rynku pracy.

Zastosowanie algorytmu Kristensena do danych polskich, aktualnych dla roku 2006, dostarczyło rezultatów wskazujących blisko 900 lokalnych rynków pracy, które składały się średnio z 3,5 gminy. Tak duże rozdrobnienie nie umożliwia charakterystyki powiązań gospodarczych i oceniono je jako niezadowalające. Wydaje się, że podejście zaproponowane przez Kristensena dobrze sprawdza się wyłącznie w sytuacji podziału niewielu dużych jednostek bazowych o powierzchni rzędu 150 km², jakimi dysponowano w Danii, dla której metoda ta została opracowana. W warunkach polskich dane o dojazdach do pracy dostępne są dla aż 3062 gmin (z podziałem na część miejską i wiejską), zatem przeciętna jednostka bazowa charakteryzuje się powierzchnią zaledwie 100 km².

Struktura oszacowań przepływów związanych z zatrudnieniem uzyskanych przy pomocy modelu wielopoziomowego, choć znacznie poprawiła rozwiązanie uzyskane z klasycznego modelu regresji, nie pozwoliła na przeprowadzenie delimitacji lokalnych rynków pracy, której wyniki mogłyby być aktualne dla roku 2011. Staraniem Autorki było jak najrzetelniesze odzwierciedlenie w modelu, struktury migracji związanych z zatrudnieniem. Dlatego skonstruowano odrębne modele dla oszacowania przyjazdów do pracy do każdej gminy oraz migracji pomiędzy każdą parą gmin. Wydawało się, że regresja wielopoziomowa spełnia wymogi takiej konstrukcji. Przypuszczenie to potwierdziły szacunki przyjeżdżających do pracy. Należy również podkreślić, że oszacowania diagonalii macierzy przepływów charakteryzowały się bardzo wysoką precyzją szacunku. Natomiast uwzględnienie kierunków migracji pomiędzy parami gmin z zastosowaniem modelu krzyżowego, nie dostarczyło szacunków o zadowalającej precyzji. Skomplikowana struktura migracji dla par gmin nie została właściwie odzwierciedlona w modelu. Z drugiej strony nie można wykluczyć, że przy szerszym zestawie zmiennych objaśniających, albo po prostu silniejszej ich korelacji ze zmienną objaśnianą lub innej definicji tej zmiennej, model wielopoziomowy dałby zadowalające rezultaty.

W związku z powyższym można stwierdzić, że stosowanie modeli wielopoziomowych do szacowania charakterystyk lokalnego rynku pracy, takich jak na przykład natężenie dojazdów do pracy jest zasadne w przekroju jednostek podziału terytorialnego. Nie udało

się jednak wykazać właściwości tej metody do szacowania przepływów związanych z zatrudnieniem określonych dla par jednostek podziału administracyjnego. W tej sytuacji najodpowiedniejszym podejściem wydaje się przeprowadzanie delimitacji lokalnych rynków pracy wyłącznie w oparciu o dane empiryczne na temat dojazdów do pracy. Podejście to jest tym bardziej uzasadnione, że w Urzędzie Statystycznym w Poznaniu trwają prace nad przeprowadzeniem kolejnego badania mobilności pracowniczey obejmującego kierunki dojazdów. Ponadto planowane jest, aby badanie to miało charakter okresowy.

ZAKOŃCZENIE

Głównym celem rozprawy była delimitacja lokalnych rynków pracy w Polsce. Podjęty problem badawczy jest bardzo aktualny, gdyż lokalny rynek pracy w warunkach gospodarki wolnorynkowej, jest przedmiotem szczególnego zainteresowania władz samorządowych, inwestorów, przedsiębiorców, jak również indywidualnych osób, aktualnych i przyszłych mieszkańców. Wynika to bezpośrednio z priorytetów polityki zrównoważonego rozwoju regionalnego, realizowanej w Polsce, jak również w krajach Unii Europejskiej. W trakcie realizacji tak sformułowanego celu pracy weryfikacji poddane zostały następujące hipotezy badawcze głoszące, że:

1. dojazdy do pracy dobrze określają lokalne rynki pracy i dlatego mogą być podstawą ich delimitacji,
2. modelowanie wielopoziomowe umożliwia oszacowanie macierzy migracji,
3. delimitacja lokalnych rynków pracy niewymagająca wyodrębnienia ośrodków centralnych, jest w Polsce bardziej uzasadniona,
4. duża liczba parametrów w stosowanych algorytmach umożliwia kształtowanie wyników,
5. delimitacja lokalnych rynków pracy powinna obejmować obszar całego kraju.

Realizację głównego celu rozprawy oraz postawionych hipotez badawczych, przeprowadzono poprzez wypełnienie zadań sprecyzowanych w celach szczegółowych. Wśród nich wyszczególniono przegląd i ocenę metod delimitacji lokalnych rynków pracy, adaptację omówionych rozwiązań do warunków polskich oraz sformułowanie własnej propozycji, a także przedstawienie idei modelowania wielopoziomowego i jego wykorzystanie do oszacowania macierzy migracji w roku 2011.

Wnikliwe studia literaturowe pozwoliły rozpoznać oraz ocenić rozwiązania stosowane w innych państwach. W ten sposób, realizując pierwszy cel szczegółowy pracy, zwrócono uwagę na powszechność modeli delimitacji działających w oparciu o dane dotyczące dojazdów do pracy. Na tej podstawie, dokonano wstępnej, pozytywnej weryfikacji pierwszej hipotezy badawczej, głoszącej, że dojazdy do pracy dobrze określają lokalne rynki pracy i mogą być podstawą ich delimitacji.

Cel pracy osiągnięty został poprzez praktyczne przeprowadzenie delimitacji lokalnych rynków pracy w Polsce, na podstawie dostępnych danych dla roku 2006. Jego realizacja łączy się bezpośrednio z weryfikacją hipotez trzeciej, czwartej i piątej. Otrzymano cztery różne, choć pod wieloma względami podobne, klasyfikacje gmin na obszary pełniące rolę lokalnych rynków pracy. Zastosowane algorytmy, zarówno propozycje innych autorów, jak i własna, działają w oparciu o znajomość macierzy przepływów związanych z zatrudnieniem. Analiza i ocena i otrzymanych wyników delimitacji, jako odpowiadających realnej strukturze lokalnych rynków pracy w Polsce, stanowiła empiryczną weryfikację pierwszej hipotezy badawczej. Dojazdy do pracy dobrze określają lokalne rynki pracy i mogą być podstawą ich delimitacji.

Pierwszą z przedstawionych delimitacji przeprowadzono przy pomocy europejskiego algorytmu regionalizacji, rekomendowanego przez Eurostat, jako standardowa metoda delimitacji lokalnych rynków pracy w krajach europejskich. Procedura ta zaliczana jest do metod opartych na zasadach i składa się z pięciu dość skomplikowanych etapów. Ponadto

jej aplikacja wymaga zdefiniowania przez badacza wielu parametrów wstępnych, które pozwalają dostosować ją do realiów danego kraju. Drugą delimitację otrzymano stosując procedurę Kristensena. W metodzie tej wybór lokalnych rynków pracy nie odbywa się wokół wyodrębnionych ośrodków centralnych, co zmniejsza liczbę parametrów. Jednak ograniczenia wprowadza hierarchiczny sposób agregacji. W obu metodach zaistniała konieczność korekty części parametrów, w stosunku do wartości sugerowanych przez autorów. Zastosowane podejście pozwoliło uwzględnić specyfikę warunków polskich. Wybór konkretnych wartości parametrów poprzedzono, potwierdzającą prawdziwość hipotezy czwartej, dyskusją ich wpływu na otrzymane wyniki.

Trzecią i czwartą delimitację wyznaczono w oparciu o dwa warianty autorskiego algorytmu, nazywanego w pracy podejściem taksonomicznym. Procedury te można zaliczyć do grupy metod indukcyjnych, które ze swojej natury są mniej skomplikowane od podejść dedukcyjnych. Ponadto aplikacja tych metod nie wymaga definiowania przez badacza żadnych parametrów wejściowych, przez co otrzymane wyniki delimitacji są obiektywne i nie zależą w żaden sposób od subiektywnych wyborów badacza.

Istotną różnicą w podejściu zaczerpniętym z literatury oraz samodzielnie opracowanej metodzie jest fakt, że działanie europejskiego algorytmu regionalizacji oparte jest na założeniu, że każdy lokalny rynek pracy stymulowany jest przez ośrodek centralny. Założenie to, oparte na teorii Thunena, choć w większości przypadków jest słuszne, nie sprawdza się w niektórych obszarach Polski o szczególnej specyfice. Na przykład na Śląsku, gdzie wiele rozwiniętych miast położonych jest w niewielkiej odległości, trudno przyjmować, że tylko jedno z nich pełni rolę ośrodka centralnego. Z drugiej strony, przykładowo w województwie podlaskim natężenie miast, które można by uznać za ośrodki centralne jest tak małe, że zasadne wydaje się raczej wyodrębnienie lokalnych rynków pracy bez wyraźnego ośrodka centralnego, niż przypisywanie do miasta Białystok wpływów sięgających niemal całego województwa. W związku z powyższym, właściwym rozwiązaniem wydaje się stosowanie podejścia taksonomicznego, umożliwiającego konstrukcję lokalnego rynku pracy, bez wskazywania stymulującego jego rozwój dokładnie jednego ośrodka centralnego.

Pozytywna weryfikacja trzeciej hipotezy badawczej, głoszącej, że w warunkach polskich bardziej uzasadniona jest delimitacja lokalnych rynków pracy niewymagająca wyodrębnienia ośrodków centralnych, wykazana została w wyniku oceny rezultatów europejskiego algorytmu regionalizacji i podejścia autorskiego. Wynika ona ze specyfiki silnego zróżnicowania terytorialnego Polski pod względem poziomu rozwoju gospodarczego kraju, a także stopnia urbanizacji. Prawdziwość hipotezy trzeciej potwierdziła szczegółowa analiza porównawcza wyników delimitacji w specyficznych regionach o dużej koncentracji ośrodków centralnych oraz ich braku.

Słuszność czwartej hipotezy badawczej potwierdziły rozważania poświęcone dostosowaniu algorytmu europejskiego oraz metody Kristensena do warunków polskich. Każdorazowa zmiana parametrów, wpływała na ostateczne rezultaty. Brak precyzyjnie określonych zasad ustalania parametrów, powodował duże rozbieżności wyników. Zatem brak konieczności definiowania a priori parametrów wejściowych uznać można za przewagę autorskiego podejścia taksonomicznego, ponieważ gwarantuje jednoznaczność otrzymanych wyników.

Weryfikację piątej hipotezy badawczej, głoszącej, że delimitacja lokalnych rynków pracy powinna obejmować obszar całego kraju, a nie tylko poszczególnych województw czy regionów, przeprowadzono także w trakcie analiz wyników badania empirycznego.

Niezależnie od zastosowanej metody, obszar części z wyznaczonych lokalnych rynków pracy nie zawierał się w tylko jednym województwie. Mianowicie gminy położone blisko granic województw były ze sobą łączone, bez względu na fakt, że należały do różnych jednostek podziału administracyjnego wyższego stopnia. Sytuację tą przeanalizowano szczegółowo na przykładzie lokalnych rynków pracy, do których należały gminy z obszaru województwa wielkopolskiego. Podział taki odzwierciedla rzeczywistą sytuację na rynku pracy. W związku z tym, ograniczenie rozważań do gmin należących do tylko jednego województwa (czy dowolnego regionu mniejszego niż cały kraj), spowodowałoby zniekształcenie lokalnych rynków pracy położonych w pobliżu jego granic. Również ograniczanie analiz do terytorium Polski wydaje się być podejściem dyskusyjnym. Zwłaszcza struktura lokalnych rynków pracy położonych w obszarach przygranicznych może być zniekształcona. Podejście ograniczające zasięg terytorialny lokalnych rynków pracy do granic kraju w sytuacji jego przynależności do zintegrowanego ekonomicznie systemu, jakim jest Unia Europejska, wydaje się być dość sztuczne, zwłaszcza w odniesieniu do ściany zachodniej. Jednym z kierunków dalszych prac będzie więc zbadanie wpływu warunków zatrudnienia oraz wymagań potencjalnej siły roboczej w państwach sąsiednich na kształt przygranicznych lokalnych rynków pracy w Polsce.

W trakcie jak najlepszego dostosowywania metod delimitacji lokalnych rynków pracy, do realiów panujących w Polsce osiągnięto drugi cel szczegółowy pracy, którym była adaptacja europejskiego algorytmu regionalizacji oraz przedstawienie autorskiego podejścia. Z kolei aplikacja omówionych i zaproponowanych procedur pozwoliła zrealizować piąty cel szczegółowy, który polegał na zastosowaniu wybranych podejść zaczerpniętych z literatury oraz autorskiego modelu do delimitacji lokalnych rynków pracy w Polsce.

Należy jednak zaznaczyć, że pomimo wskazanych różnic, wyniki delimitacji lokalnych rynków pracy w Polsce, dokonanej przy pomocy europejskiego algorytmu regionalizacji oraz pierwszego z wariantów autorskiej metody, nazywanego w treści pracy podejściem taksonomicznym z mocnym kryterium łączenia, są do siebie bardzo zbliżone. Z kolei w przypadku zastosowania drugiego z wariantów autorskiej metody, nazywanej w treści pracy podejściem taksonomicznym ze słabym kryterium łączenia uzyskano prawie o połowę mniej lokalnych rynków pracy. Optymalna liczba i wielkość wyznaczanych lokalnych rynków pracy jest kwestią dyskusyjną. Porównując wyniki analiz

przeprowadzonych w innych krajach zauważyć można dużą różnorodność w tej dziedzinie. Co więcej, podobne badania nie zostały w Polsce przeprowadzone w warunkach gospodarki wolnorynkowej, dlatego też nie ma możliwości porównania otrzymanych wyników z wcześniejszymi analizami. Wydaje się, że pożądana liczba i rozmiar lokalnych rynków pracy zdeterminowana może być również celem ich delimitacji.

Obok delimitacji lokalnych rynków pracy w Polsce, której wyniki byłyby aktualne dla roku 2006, podjęto próbę estymacji macierzy migracji oraz jej predykcji, umożliwiającą przeprowadzenie podobnej analizy dla roku 2011. Ze względu na strukturę badanej populacji oraz zmiennych postanowiono w tym celu zastosować metodologię modelowania wielopoziomowego. W pracy omówiono ideę oraz zasady konstrukcji modelu wielopoziomowego. Uznano, że modele te, uwzględniając relacje pomiędzy czynnikami określonymi dla jednostek indywidualnych oraz na poziomach zagregowanych, stwarzają możliwość odzwierciedlenia skomplikowanej struktury dojazdów do pracy. Dlatego zaproponowano dwa schematy szacowania wielkości macierzy migracji: na diagonalu i poza diagonalą, konstruując dwa odrębne modele.

Na przekątnej macierzy znaleźć powinny się liczby osób pracujących w gminie zamieszkania. Oszacowania należy, więc, przeprowadzić dla gmin, które 'zagnieżdżone' są w powiatach, które z kolei 'zagnieżdżone' są w województwach, dzięki temu skonstruowano model trzy-poziomowy. Poza przekątną macierzy migracji znaleźć powinny się z kolei wartości przepływów związanych z zatrudnieniem pomiędzy parami gmin. Prowadząc oszacowania dla par gmin uwzględniono, że każda z nich określona jest przez gminę, z której przepływ ten pochodzi oraz gminę, do której jest on skierowany. Ponadto obie gminy 'zagnieżdżone' są w powiatach i województwach. W związku z tym podjęto konstrukcję modelu czteropoziomowego z dwoma kryteriami grupowania na drugim poziomie (gmin). W ten sposób zrealizowano trzeci cel szczegółowy pracy. W przypadku obu modeli uzyskano znaczną poprawę precyzji szacunku, w porównaniu z klasyczną regresją liniową. Ponadto przy pomocy modelu skonstruowanego w celu oszacowania wartości na diagonalu udało się wyjaśnić aż 99% zmienności szacowanej zmiennej, zaś w przypadku modelu służącego do estymacji przepływów spoza diagonalu wyjaśniono ponad 70% zmienności szacowanej zmiennej. Przy pomocy obydwóch modeli możliwe było oszacowanie macierzy migracji aktualnej dla roku 2011, a tym samym zrealizowanie czwartego celu szczegółowego.

Na podstawie oszacowanej macierzy migracji podjęto próbę przeprowadzenia delimitacji lokalnych rynków pracy w Polsce, aktualnej dla roku 2011. Otrzymane wyniki okazały się jednak dyskusyjne. Po analizie przyczyn takiej sytuacji stwierdzono, że oszacowania przepływów związanych z zatrudnieniem pomiędzy parami gmin, czyli wartości spoza diagonalii macierzy migracji, choć charakteryzują się akceptowalną precyzją szacunku, obciążone są błędem systematycznym. Mianowicie przeszacowane zostały dojazdy podejmowane na dużych odległościach a niedoszacowane przepływy pomiędzy pobliskimi gminami. W celu korekty tego błędu podjęto próbę przeskalowania oszacowań w oparciu o znane błędy popełnione w roku 2006 oraz informację o liczbie osób pracujących na terenie gmin w roku 2011. Wprowadzone korekty nie poprawiły jednak, w zadowalającym stopniu, jakości oszacowanej macierzy migracji. W związku z tym nie udało się uzyskać wiarygodnej klasyfikacji gmin polskich na lokalne rynki pracy aktualnego dla roku 2011. Tym samym nie udało się dokonać pozytywnej weryfikacji drugiej hipotezy badawczej, głoszącej, że zastosowanie modelowania wielopoziomowego umożliwia oszacowanie macierzy przepływów związanych z zatrudnieniem. Z drugiej strony, otrzymane wyniki nie dowodzą również fałszywości tej tezy. Mimo że zmienne objaśniające wprowadzono do modelu w procesie regresji krokowej z szerokiego wachlarza charakterystyk rynku pracy, łatwości dojazdu oraz odległości, istnieje możliwość, że przy zastosowaniu szerszego zestawu zmiennych uzyskano by zadowalające wyniki.

Należy także podkreślić, że oszacowania liczby osób pracujących w gminie zamieszkania, charakteryzowały się bardzo wysoką precyzją szacunku i nie były obciążone błędem systematycznym. Zatem zastosowanie modelu wielopoziomowego sprawdziło się podczas szacowania wielkości na diagonalii macierzy przepływów związanych z zatrudnieniem.

Podsumowując warto podkreślić, że najistotniejszym rezultatem niniejszej pracy jest przeprowadzenie delimitacji lokalnych rynków pracy w Polsce. Znajomość zakresu terytorialnego poszczególnych lokalnych rynków pracy powinna leżeć u podstaw wielu strategicznych decyzji, a także stanowić podstawę innych badań i analiz. Badanie takie zostało przeprowadzone w Polsce po raz pierwszy, w warunkach gospodarki wolnorynkowej. W trakcie realizacji celu głównego pracy, osiągnięto wszystkie pięć celów szczegółowych określonych we wstępie pracy. Ponadto udało się pozytywnie zweryfikować cztery z pięciu postawionych hipotez badawczych. Otwartą kwestią została

jedynie weryfikacja drugiej z nich, która dotyczyła zasadności zastosowania modelowania wielopoziomowego do estymacji charakterystyk dojazdów do pracy. Badania w tym zakresie będą kontynuowane przez autorkę.

Jako dalsze kierunki prac należy wskazać również pogłębienie przeprowadzonej analizy wyznaczonych lokalnych rynków pracy pod kątem odpowiedzi na pytanie, która z zastosowanych metod dała wyniki najlepiej odzwierciedlające sytuację rzeczywistą na polskim rynku pracy. Ważne cele dalszych badań określono także przeprowadzając analizę porównawczą otrzymanych wyników. Dotyczą one wskazania optymalnej metody konstrukcji oraz liczby lokalnych rynków pracy. Ponadto podjęte zostaną dalsze próby estymacji macierzy migracji. W tym celu przeprowadzone zostaną kolejne oszacowania macierzy przepływów związanych z zatrudnieniem, w oparciu o zmienne objaśniające uwzględniające charakterystyki rynku mieszkaniowego. Z drugiej strony, w Urzędzie Statystycznym w Poznaniu trwają prace nad przeprowadzeniem kolejnego badania mobilności pracowniczej obejmującego kierunki dojazdów. W sytuacji publikacji wyników z tego zakresu możliwe byłoby przeprowadzenie delimitacji aktualnej na moment tego badania na podstawie danych empirycznych.

LITERATURA

- Andersen A.K. (2002), *Are Commuting Areas Relevant for the Delimitation of Administrative Regions in Denmark?*, *Regional Studies*, 36 (8), 833-844.
- Bamber A., Walter R. (2009), *Implementing the Australian statistical geography standard (ASGS)*. In: Ostendorf B., Baldock P., Bruce D., Burdett M., Corcoran P.(eds.), *Proceedings of the Surveying & Spatial Sciences Institute Biennial* cran.r-project.org/web/packages/lme4/vignettes/Implementation.pdf. International Conference, Adelaide 2009, Surveying & Spatial Sciences Institute, pp. 605-616.
- Bates D. (2010), *lme4: Mixed-effects Modeling with R*, Springer
- Bates D. (2013a), *Computational methods for mixed models*, <http://cran.r-project.org/web/packages/lme4/vignettes/Theory.pdf>.
- Bates D. (2013b), *Linear mixed model implementation in lme4*, <http://cran.r-project.org/web/packages/lme4/vignettes/Implementation.pdf>.
- Bates D. (2013c), *Penalized least squares versus generalized least squares representations of linear mixed models*, <http://cran.r-project.org/web/packages/lme4/vignettes/PLSvGLS.pdf>.
- Beguín H. (1992), *Region a ośrodki centralne*, w: *Ekonomiczna analiza przestrzenna*, red. C. Ponsard, Wyd. AE w Poznaniu, Poznań.
- Biecek P. (2011), *Analiza danych z programem R - Modele liniowe z efektami stałymi, losowymi i mieszanymi*, PWN, Warszawa.
- Bliese P. (2012), *Multilevel Modeling in R (2.4). A Brief Introduction to R, the multilevel package and the nlme package*, Paul Bliese, April 10, http://cran.r-project.org/doc/contrib/Bliese_Multilevel.pdf
- Bołt T., Krauze K., Kulawczuk T. (1985), *Agregacja modeli ekonometrycznych*, Państwowe Wydawnictwo Ekonomiczne, Warszawa.
- Brown L., Holmes J. (1971), *The delimitation of functional regions, nodal regions, and hierarchies by functional distance approaches*, *Journal of Regional Science*, 1, 57-72.
- Bureau of Transport and Regional Economics (BTRE) (2003), *Focus on Regions. No. 1 Industry Structure*, Information Paper, No. 49, Canberra (Australia).
- Carlsson F., Johansson M., Persson L.O. and Tegsjö B. (1993), *Creating Labour Market Areas and Employment Zones. New Regional Divisions in Sweden Based on*

Commuting Statistics, CERUM Report. Umeå (Sweden) The Centre for Regional Science (CERUM), Umeå University.

- Casado-Diaz J.M., (2000), *Local Labour Market Areas in Spain: A case study*, *Regional Studies*, 34 (9), 843-856,
- Casado-Diaz J.M., Coombes M. (2011), *The delineation of 21st century local labour market areas: a critical review and a research agenda*, *Boletín de la Asociación de Geógrafos Españoles*, 57, 7-32.
- Cegielski J., Kluszewski S., 1952, *Dojazdy ludności do Warszawy w latach 1950–1951*, IBM, Materiały i Dokumentacja, ser. E, z. 4.
- Cervero R. (1995), *Planned Communities, Self-containment and Commuting: A Crossnational Perspective*. *Urban Studies*, vol. 32, No. 7, 1135-1161.
- Chojnicki Z. (1966), *Zastosowanie modeli grawitacji i potencjału w badaniach przestrzenno-ekonomicznych*, Komitet Przestrzennego Zagospodarowania Kraju Polskiej Akademii Nauk, Warszawa.
- Coombes M.G. (1998), *1991-based Travel-to-Work Areas*. Londres, Office for National Statistics.
- Coombes M.G. (2002), *Travel to Work Areas and the 2001 Census*. Report to the Office for National Statistics, Newcastle. CURDS, March 2002, 1-16.
- Coombes M.G., Bond S. (2008), *Travel-to-Work Areas: the 2007 review*. London, Office of National Statistics, 2008, 1-16.
- Coombes M.G., Dixon J.S., Goddard J.B., Openshaw S. and Taylor P.J. (1982), *Functional Regions for the Population Census of Great Britain*, in: *Geography and the Urban Environment, Progress in Research and Applications V* (Herbert, D.T. and Johnston, R.J., eds.). Chichester, Wiley, 63-112.
- Coombes M.G., Green A.E. and Openshaw S. (1986), *An Efficient Algorithm to Generate Official Statistical Reporting Areas*, *Journal of the Operational Research Society*, 37 (10), 943-53.
- Coombes M.G., Wymer C., Charlton M.E., Bailey S., Stonehouse A. and Openshaw S. (1997), *Review of Travel-to-Work Areas and Small Area Unemployment Rates*, *Labour Market Trends*, 105, 9-12.
- Cronbach L.J., Webb N. (1975), *Between-Class and Within-Class Effects in a Reported Aptitude X Treatment Interaction: Reanalysis of a Study by G. L. Anderson*, *Journal of Educational Psychology*, vol. 67, 717-724.

- Czyż T., (1996), *Zastosowanie modelu potencjału w regionalizacji strukturalnej Polski*, red. T. Czyż, Wyd. Bogucki Wydawnictwo Naukowe, Poznań.
- De Leeuw, J., Kreft, I. G. G. (1986), *Random coefficient models for multilevel analysis*. Journal of Educational Statistics, 11, 57–85.
- Dempster A. P., Rubin, D. B., Tsutakawa, R. K. (1981), *Estimation in covariance component models*. Journal of the American Statistical Society, 76, 341-353.
- Dempster, A.P.; Laird, N.M.; Rubin, D.B. (1977), *Maximum Likelihood from Incomplete Data via the EM Algorithm*, Journal of the Royal Statistical Society, Series B, 39 (1), 1–38.
- DiPrete T. A., Forristal, J. D. (1994), *Multilevel models: Methods and substance*. Annual Review of Sociology, 20, 331–357.
- Dojazdy do pracy w Polsce. (2010) *Terytorialna identyfikacja przepływów ludności związanych z zatrudnieniem*, red. K. Kruszka, Główny Urząd Statystyczny i Urząd Statystyczny w Poznaniu, Poznań 2010.
- Dojazdy do pracy w 2010 roku na podstawie BAEL, (2010), Departament Pracy przy Głównym Urzędzie Statystycznym, Materiał na konferencję prasową w dniu 22 grudnia 2011 r.
- Drobne, S., Konjar, M., Lisec, A., Pichler Milanović, N. and Zavodnik Lamovšek, A. (2010), *Functional regions defined by urban centres of international importance: the case of Slovenia*, REAL CORP Proceedings, Vienna. Available online at: http://www.corp.at/archive/CORP2010_153.pdf;
- Enders C. K., Tofighi, D. (2007), *Centering predictor variables in cross-sectional multilevel models: A new look at an old issue*. Psychological Methods, 12, 121–138.
- Eurostat (1992), *Study on Employment Zones*, Eurostat (E/LOC/20), Luxembourg.
- Filas-Przybył S., Klimanek T., Kowalewski J. (2012), *Analiza dojazdów do pracy za pomocą modelu grawitacji*, Taksonomia 19, Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu 242, 135-143.
- Filas-Przybył S., Stachowiak D. (2010), *Badanie przepływów ludności związanych z zatrudnieniem w: Procesy metropolizacyjne w teorii naukowej i praktyce* konferencja naukowa Łódź, 12-14 października 2009, GUS, Warszawa.
- Florez-Revuelta F.; Casado-Diaz J.M., Martinez-Bernabeau L. (2008), *An evolutionary approach to the delineation of functional areas based on travel-to-work flows*, International Journal of Automation and Computing, vol.5, No. 1, 10-21.

- Frątczak E., Gołata E., Klimanek T., Ptak-Chmielewska A., Pęczkowski M. (2009), *Wielowymiarowa analiza statystyczna. Teoria – przykłady zastosowań z systemem SAS*, red. E. Frątczak, Oficyna Wydawnicza SGH w Warszawie.
- Frątczak E., Mianowska K. (2012), *Modele mieszane*, w: *Zaawansowane metody analiz statystycznych*, red. Frątczak E., Oficyna Wydawnicza SGH w Warszawie.
- Fusco G., Cagliani M. (2011), *Hierarchical Clustering through Spatial Interaction Data. The Case of Commuting Flows in South-Eastern France*, Computer Science, Vol. 6782/2011, 135-151.
- Gawryszewski A. (1989), *Przestrzenna ruchliwość ludności Polski, 1952-1985*. Ossolineum, Warszawa, ss. 370. Prace Habilitacyjne IGiPZ PAN.
- Gawryszewski A. (1977), *Dojazdy do pracy*, w: *Rozmieszczenie i migracje ludności a system osadniczy Polski Ludowej*, Prace Geograficzne, nr 117, IGiPZ PAN.
- Gawryszewski A. (2005), *Ludność Polski w XX wieku*, Warszawa: Instytut Geografii i Przestrzennego Zagospodarowania PAN, Monografie, nr 5.
- Gawryszewski A., Korcelli P., Nowosielska E. (1998), *Funkcje metropolitalne Warszawy*, Zeszyty IGiPZ PAN, 53.
- Gawryszewski A., Potrykowska A. (1980), *Rozkłady odległości dojazdów do pracy do wybranych miast w latach 1959-1973*. Przegląd Geograficzny 52, 4, 789-807.
- Gazińska M. (1995), *Modele ciężenia i potencjału w badaniu bezrobocia w województwie szczecińskim*, Przegląd Statystyczny, R. XLII, z. 3-4.
- Gocał T., Rakowski W. (1991), *Delimitacja regionów i subregionów migracyjnych w zakresie dojazdów do pracy*, Monografie i Opracowania / Szkoła Główna Planowania i Statystyki; 332, Podseria: Uwarunkowania Demograficzne Rozwoju Społeczno-Gospodarczego Polski; 17, Warszawa.
- Goldstein H. (2003), *Multilevel Statistical Models*, 3rd edition, London: Edward Arnold.
- Gołata E. (2004), *Estymacja bezpośrednia bezrobocia na lokalnym rynku pracy*, Wydawnictwo Akademii Ekonomicznej w Poznaniu.
- Gołata E. (2006), *System informacji o rynku pracy a potrzeby analizy w ujęciu regionalnym i lokalnym*, [w:] *Praca w perspektywie ekonomicznej* (red.) H. Januszek, Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań, 2006, s. 89-110, ISSN 1641-2168, ISBN 978-83-7417-165-6
- Gołata E., Dehnel G., Gruchociak H. (2011), *Analiza przestrzenna w badaniu dojazdów do pracy w Polsce*, Taksonomia 18.

- Gontarski Z. (1980), *Obszary metropolitalne w Polsce. Delimitacja i charakterystyka struktury przestrzennej*, Biuletyn KPZK PAN, nr 109, Warszawa.
- Goodman J.F.B. (1970), *The definition and analysis of local labour markets: some empirical problems*, British Journal of Industrial Relations, 8, 179-186.
- Góra M., Sztanderska U. (2006), *Wprowadzenie do analizy lokalnego rynku pracy. Przewodnik*, Ministerstwo Pracy i Polityki Społecznej, Departament Analiz Ekonomicznych i Prognoz, Warszawa 2006.
- Gruchociak H. (2010), *Dojazdy do pracy w woj. wielkopolskim*, Wiadomości Statystyczne, r. LV, No 9, 49-60.
- Gruchociak H. (2012a), *Konstrukcja estymatora regresyjnego dla danych o strukturze dwupoziomowej*, Taksonomia 19, 590-600.
- Gruchociak H. (2012b), *Możliwości zastosowania modelowania wielopoziomowego w badaniach ekonomicznych*, Przegląd Statystyczny, 2012, 4, 409-433.
- Gruchociak H. (2012c), *Odległość w wyodrębnianiu stref wpływu wielkich miast w Polsce*, Zeszyty Naukowe UEP, nr 227, Poznań.
- Gruchociak H. (2012d), *Delimitacja lokalnych rynków pracy w Polsce*, Przegląd Statystyczny, Numer Specjalny 2 z Okazji Kongresu Statystyki Polskiej, 277-297.
- Gumuła W., Socha J., Wojciechowski W. (2007), *Presja płacowa oraz niedopasowanie strukturalne na rynku pracy w świetle badań NBP*, Materiały i Studia NBP, Departament Analiz Makroekonomicznych i Strukturalnych, Zeszyt nr 219, Warszawa.
- GUS (1973), *Strefy wpływów dużych miast w świetle dojazdów do pracy*, GUS, Warszawa.
- GUS (1986), *Delimitacja regionów intensywnych dojazdów do pracy*, GUS, Warszawa.
- Hoffmann-Martinot V. (2004), *Towards an Americanization of French Metropolitan Areas?*, Referat przygotowany na spotkanie programu International Metropolitan Observatory w Bordeaux, 9-10.01.2004.
- Hofmann D. A., Gavin, M. B. (1998), *Centering decisions in hierarchical linear models: implications for research in organizations*. Journal of Management, 24 (5), 623-641.
- Hox J. (2002), *Multilevel Analysis. Techniques and Applications*, Lawrence Erlbaum Associates, Publishers, London.

- Hox J. (2010), *Multilevel Analysis. Techniques and Applications*, Second Edition, Quantitative Methodology Series, Routledge, New York.
- INSEE (1998), *Les zonages: enjeux et méthodes*, INSEE Méthodes 83. Paris. INSEE.
- ISTAT (1997), *I sistemi locali del lavoro 1991*, edited by F. Sforzi, Rome, ISTAT.
- ISTAT (2005), *I Sistemi Locali del Lavoro, Censimento 2001. Dati definitivi*, Rome, ISTAT.
- Jałowiecki B., 2000, *Spółeczna przestrzeń metropolii*, Scholar, Warszawa.
- Jarmołowicz W.(red.) (2008), *Podstawy makroekonomii*, Wydawnictwo AE Poznań
- Jarmołowicz W., Kalinowska-Sufinowicz B. (2012), *Funkcjonowanie regionalnego rynku pracy (na przykładzie województwa wielkopolskiego w latach 2000-2010)*, Acta Universitatis Lodzensis, Folia Oeconomica 268, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Kallioras, D., Kandyliis, Y., Kromyidakis, N. & Pantazis, P. (2011), *Delineation of local labour markets in Greece on the basis of travel-to-work flows*, European Regional Science Association (ERSA) Conference Papers, ersa11p75.
- Klassen L.H., Paelinck J.H.P., Wagenaar S. (1982), *Systemy przestrzenne*, PWN, Warszawa.
- Klimanek T. (2003), *Wielopoziomowa analiza struktury agrarnej gminy w systemie Geo-Info*, Praca doktorska napisana na Akademii Ekonomicznej w Poznaniu na Wydziale Zarządzania w Katedrze Statystyki, Poznań,
- Komornicki T., Śleszyński P. (2009), *Typologia obszarów wiejskich pod względem powiązań funkcjonalnych i relacji miasto-wieś*, [w:] J. Bański (red.), *Analiza zróżnicowania i perspektyw rozwoju obszarów wiejskich w Polsce do 2015 roku*, Studia Obszarów Wiejskich, 16, Warszawa, 9-37.
- Konjar, M., Lisec, A. & Drobne, S. (2010). *Methods for delineation of functional regions using data on commuters*, 13th AGILE International Conference on Geographic Information Science, Guimarães, Available online at: http://agile2010.dsi.uminho.pt/pen/ShortPapers_PDF/93_DOC.pdf;
- Kopczewska K. (2006), *Ekonometria i statystyka przestrzenna z wykorzystaniem pakietu R CRAN*, CeDeWu, Warszawa.
- Kopczewska K. (2010), *Modele zmian stopy bezrobocia w ujęciu przestrzennym*, Wiadomości Statystyczne, nr 5, GUS, PTS, Warszawa.

- Korcelli P. (1981), *Regiony miejskie w systemie osadniczym Polski*, w: Studia nad migracjami i przemianami systemu osadniczego w Polsce, Prace Geograficzne, nr 140, IGiPZ PAN.
- Korcelli P., Gawryszewski A., Potrykowska A. (1992), *Przestrzenna struktura ludności Polski. Tendencje i perspektywy*, Studia KPZK PAN, t. XCVIII.
- Korcelli P., Potrykowska A., Bodzak D. (1981), *Układ przestrzenny i współzależności ośrodków dojazdów do pracy*, w: Studia nad migracjami i przemianami systemu osadniczego w Polsce, Prace Geograficzne, nr 140, IGiPZ PAN.
- Korcelli P., Śleszyński P., 2006, *Delimitacja Funkcjonalnych Obszarów Miejskich 2006 na potrzeby programu ESPON 1.4.3. Study on Urban Function* (mapa opublikowana m.in. w: P. Korcelli, M. Degórski, D. Drzazga, T. Komornicki, T. Markowski, J. Szlachta, G. Węclawowicz, J. Zaleski, J. Zaucha, 2010, Ekspertycki projekt Koncepcji Przestrzennego Zagospodarowania Kraju, Studia, 78, KPZK PAN, Warszawa, s. 27).
- Kosztowniak A. (2011), *Kapitał zagraniczny jako determinant rozwoju gospodarczego w różnych systemach gospodarczych - przypadek Chin*, w: Wzrost gospodarczy - rynek pracy - polityka państwa, red. Kwiatkowski E., Kucharski L., Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Kotowska I., red. (2012), *Rynek pracy i wykluczenie społeczne w kontekście percepcji Polaków*. Diagnoza społeczna 2011, Ministerstwo Pracy i Polityki Społecznej, Warszawa.
- Kreft, I, de Leeuw J.(1998), *Introducing multilevel modeling*, ISM Introducing Statistical Methods, Sage Publications, London.
- Kreft I., de Leeuw, J., Aiken, L. (1995), *The effect of different forms of centering in hierarchical linear models*. Multivariate Behavioral Research, 30, 1–22.
- Kristensen K. (1998), *Functional Economic Areas of Denmark: Applying Input-Output Techniques to Commuting*, Bornholms Forskningscenter.
- Kryńska E. (2010), *Podręcznik użytkownika metod, narzędzi i procedur diagnozowania zapotrzebowania na kwalifikacje i umiejętności na regionalnym i lokalnym rynku pracy*, Ministerstwo Pracy i Polityki Społecznej, Warszawa, 2010.
- Krzysztofik R., Runge. J (2011), *Delimitacja regionu górnośląsko-zagłębiowskiego metropolii „Silesia”*, Uniwersytet Śląski, Wydział Nauk o Ziemi, Sosnowiec.
- Krzyśko M. (1996), *Statystyka matematyczna*, Wydawnictwo Naukowe UAM, Poznań.
- Krzyśko M. (2009), *Podstawy wielowymiarowego wnioskowania statystycznego*, Wydawnictwo Naukowe UAM, Poznań.

- Kwiatkowski E., Włodarczyk P. (2012), *Determinanty zmian zatrudnienia w krajach Unii Europejskiej w latach 2005-2010*, Acta Universitatis Lodziensis, Folia Oeconomica 268, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Lagnerö M. (2003), *Local Labour Markets*. Working Paper. Statistics, Stockholm.
- Laird N.M., Ware J.H. (1982), Random Effects Models for Longitudinal Data, *Biometrics*, vol.38, No.4, 963-974.
- Landre M. (2012), *Journey-to-work Data: delineating Commuting Regions in Dalarna, Sweden*, ISPRS International Journal of Geo_Information, 2012, 1, 294-314.
- Landre M., Håkansson J.(2013), *Rule versus Interaction Function Evaluating Regional Aggregations of Commuting Flows in Sweden*. European Journal of Transport and Infrastructure Research, TU Delft, 13 (1) 1-19.
- Lijewski T.(1967), *Dojazdy do pracy w Polsce*, Studia KPZK PAN, t. 14, Warszawa.
- Lin X., (1997), *Variance component testing in generalized linear models with random effects*, *Biometrika*, 84, 309-25,
- Lindley D.V, Smith. A.F.M. (1972), *Bayes estimates for the Linear Model*, *Journal of the Royal Statistical Society, Series B (Methodological)*, vol.34, Issue 1, 1-41
- Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996), *SAS system for mixed models*. Cary; NC: SAS Institute, Inc.
- Longford N. T. (1987), *A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects*. *Biometrika*, 74, 817-827.
- Longford N.T.(1993), *Random coefficient models*, Oxford, Clarendon Press.
- Maik W. (1976), *Założenia i zastosowanie koncepcji zespołów osadniczych*, *Przegląd Geograficzny*, 47, 2.
- METREX (The Network of European Metropolitan Regions and Areas) (2010), *Intra-metropolitan polycentricity in practice - Reflections, challenges and conclusions from 12 European metropolitan areas*, Final report of the METREX - Expert Group on Intra-Metropolitan Polycentricity.
- Misiak T., Tokarski T., (2012), *Wewnątrzregionalne zróżnicowanie rynku pracy w Polsce*, *Wiadomości Statystyczne*, nr 12/2012.
- Mitchell, W., Bill, A. & Watts, M. (2007). *Identifying functional regions in Australia using hierachical aggregation techniques*, Working Paper no 07-06, Centre of Full Employment and Equity, University of Newcastle, Calaghan, NSW.

- Nel, J.H., Krygsman, S.C., De Jong, T. (2008), *The identification of possible future provincial boundaries for South Africa based on an intramax analysis of journey-to-work data*, Orion, vol. 24, no. 2, pp. 131-156.
- Newell J., Papps K., (2001), *Identifying Functional Labour Market Areas in New Zealand: A Reconnaissance Study using Travel-to-Work Data*, New Zealand Department of Labour Occasional Paper Series, 2001/6.
- Newell, J., Perry, M. (2005), *Explaining continuity in New Zealand's local labour market areas 1991 to 2001*. Australasian Journal of Regional Studies, vol. 11, No. 2, 155-174.
- Nystuen J. D., Dacey M.F., (1961), *A Graph Theory Interpretation of Nodal Regions*, Papers and Proceedings of the Regional Science Association , 7, 29–42.
- OECD (2002): *Redefining Territories. The Functional Regions*. Paris. OECD Publications.
- Ogólna charakterystyka badania przepływów ludności związanych z zatrudnieniem w 2006 r., opracowanie: Ośrodek Statystyki Miast, Urząd Statystyczny w Poznaniu,
http://www.stat.gov.pl/cps/rde/xbcr/poznan/ASSETS_Dojazdy_do_pracy_metodologia_do_internetu.pdf.
- Ohme J. (1988), *Problemy i metody badania zasięgu i struktury dojazdów do pracy*, w: Problemy geografii osadnictwa i ludności w Polsce, Wydawnictwo Naukowe UAM, Poznań - Zielona Góra, Seria Geografia, nr 42.
- Openshaw W. S., Rao, L. (1995), *Algorithms for re-engineering 1991 Census geography*, Environment & Planning A, vol.27, 425-446.
- Paccagnella O. (2006), *Centering or not centering in multilevel models?* Evaluation Review, 30, 66–85.
- Papps K.L., Newell J.O. (2002), *Identifying Functional Labour Market Areas in New Zealand: A Reconnaissance Study Using Travel-to-Work Data*, IZA Discussion Paper 443, Institute for the Study of Labor, Bonn.
- Paradysz J., Paradysz K. (2012), *Obszary bezrobocia w Polsce. Problem benchmarkowy*, Taksonomia 19, Prace naukowe UE we Wrocławiu nr 242, 106-115.
- Parysek J. (1996), *Rozwój lokalny i lokalna gospodarka przestrzenna*, praca zbiorowa pod redakcją J. Paryska, Bogucki Wydawnictwo Naukowe, Poznań.
- Parysek J. (2001), *Podstawy gospodarki lokalnej*, Wydawnictwo Naukowe UAM, Poznań.

- Pinheiro J. C., Bates D. M. (2000), *Mixed-effects models in S and S-PLUS*, New York: Springer-Verlag.
- Potrykowska A. (1983), *Współzależności między dojazdami do pracy a strukturą społeczną i demograficzną regionu miejskiego Warszawy w latach 1950-1973*, Dokumentacja Geograficzna, z. 5, IGiPZ PAN.
- Potrykowski M., Taylor Z. (1984), *Geografia transportu*, PWN, Warszawa.
- Pracujący w gospodarce narodowej*, (2011), GUS, Warszawa.
- Prodromidis P.-I. K. (2008), *Deriving Labour Market Areas in Greece from Commuting Flows*, Center of Planning and Economic Research, 99, Athens, located at http://www.kepe.gr/pdf/D.P/dp_99.pdf.
- Prodromidis P.-I. K. (2010), *Identifying Spatial Labour Markets in Greece from the 2001 Travel-to-Work Patterns*, South-Eastern Europe Journal of Economics, 8 (1), 111-128.
- Rao J. N. K., (2003), *Small Area Estimation*, Wiley & Sons, New York.
- Raudenbush S.W., Bryk A.S. (2002), *Hierarchical Linear Models. Applications and Data Analysis Methods*, Second Edition, Sage Publications, London_Thousand Oaks_New Delhi.
- Rocznik Statystyczny Pracy, (2006), GUS, Warszawa
- Rosik P., Stępnia M., Wiśniewski R. (2010), *Dojazdy do pracy do Warszawy i Białegostoku – alternatywne podejścia metodologiczne*, Studia Regionalne i Lokalne Nr 2(40)/2010.
- Roszka W. (2013), *Statystyczna integracja danych w badaniach społeczno-ekonomicznych*. Praca doktorska, UE w Poznaniu, Wydział Informatyki i Gospodarki Elektronicznej.
- Runge J., (1993), *Lokalny rynek pracy na przykładzie wybranych miast Zagłębia Dąbrowskiego*, w: Problematyka lokalnych systemów osadniczych (red. W.Maik), Uniwersytet Mikołaja Kopernika, Toruń, s. 155-161.
- Rydlewski J. P., (2009), *Estymatory największej wiarygodności w uogólnionych modelach regresji nieliniowej*, Praca doktorska napisana na Uniwersytecie Jagiellońskim na Wydziale Matematyki i Informatyki w Instytucie Matematyki, Kraków.
- Sakamoto Y., Ishiguro, M., and Kitagawa G. (1986). *Akaike Information Criterion Statistics*, D. Reidel Publishing Company.

- SCB (Statistics Sweden) (2010), *Construction and use of local labour areas in Sweden*. Örebro. Statistics Sweden.
- Schwarz G. (1978), *Estimating the Dimension of a Model*, *Annals of Statistics*, 6, 461-464,
- Sforzi F. (1987), *L'identificazione spaziale*, in: *Becatini G. (Ed) Mercato e Forze Locali: Il Distretto Industriale*. Il Mulino, Bologna.
- Singer J. D. (1998), *Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models*. *Journal of Educational and Behavioral Statistics*, 23, 323–355.
- Smart M. W. (1974), *Labour Market Areas: Uses and Definition*, vol.2, No 4, 238-353, Pergamon Press, Oxford.
- Smart M. W. (1981), *Labour Market Areas in Great Britain: development since 1961*, *Geoforum*, 12, 4, s. 301-318.
- Smętkowski M. (2001), *Nowe relacje między metropolią i regionem w gospodarce informacyjnej*. *Studia Regionalne i Lokalne*, 4.
- Smętkowski M., (2007), *Delimitacja obszarów metropolitalnych – nowe spojrzenie*, (w:) G. Gorzelak, A. Tucholska (red.), *Rozwój, region, przestrzeń*, Ministerstwo Rozwoju Regionalnego, EUROREG, Warszawa:215-233.
- Smętkowski M., Jałowiecki B., Gorzelak G. (2008), *Obszary metropolitalne w Polsce: problemy rozwojowe i delimitacja*, *Raporty i Analizy EUROREG*, 1, Warszawa.
- Snijders T., Bosker R. (1999), *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling*, Sage Publications, Newbury Park, CA.
- Sokołowski A. (1992), *Empiryczne testy istotności w taksonomii*, *Zeszyty Naukowe AE w Krakowie*, Seria Specjalna Monografie nr 108.
- Sokołowski A. (2012), *Q uniwersalna miara odległości*, *Taksonomia 19*, *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*.
- Spis Kadrowy (1983), *Dojazdy do pracy*, GUS, Warszawa.
- Spotila J.T. (2000), *Standards for Defining Metropolitan and Micropolitan Statistical Areas*. *Federal Register*, vol. 65, No. 249, December 27, 82228-82238.
- Stachowski J. (1993), *Problemy delimitacji lokalnych rynków pracy*, w: *Problematyka lokalnych systemów osadniczych* (red. W. Maik), Uniwersytet Mikołaja Kopernika, Toruń, 155-161.

- Statistics Norway (2001), *Classification of Economic Regions*. Oslo-Kongsvinger (Norway). Statistics Norway.
- Strahl D. (2010), *Innowacyjność europejskiej przestrzeni regionalnej a dynamika rozwoju gospodarczego*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław.
- Strahl D. (ed) (1998), *Taksonomia struktur w badaniach regionalnych*, Wydawnictwo AE we Wrocławiu.
- Swianiewicz P., Klimska U. (2005), *Społeczne i polityczne zróżnicowanie aglomeracji w Polsce – waniliowe przedmieście, mozaika przedmieść*, Prace i Studia Geograficzne, tom 35, 45-70, Warszawa.
- Szumlicz T. (1993), *Lokalne rynki pracy jako podstawa delimitacji powiatów*, w: Regionalne uwarunkowania polityki społecznej, (red. L. Frąckiewicz), AE Katowice, 88-93
- Śleszyński P, Czapiewski K (2011), *Znaczenie ośrodków miejskich oraz ich hierarchicznych powiązań dla regionalnego i lokalnego rozwoju ekonomicznego i społecznego Polski wschodniej*. Ekspertyza wykonana na zlecenie Ministerstwa Rozwoju Regionalnego, ss.149.
- Śleszyński P., 2012, *Kierunki dojazdów do pracy*, *Wiadomości Statystyczne*, 2012/11, 59-74.
- Tolbert C.M., Sizer M.(1996), *US Commuting Zones and Labor Market Areas: 1990 Update*, ERS Staff Paper AGES-9614. Rural Economy Division, Economic Research Service, US Department of Agriculture.
- Twisk J. W. R. (2010), *Analiza wielopoziomowa – przykłady zastosowań Praktyczny podręcznik biostatyki i epidemiologii*, Oficyna Wydawnicza SGH, Warszawa,
- Unolt J.(1999), *Ekonomiczne problemy rynku pracy*, Wyd. BPS Śląsk, Katowice.
- US Department of Agriculture (2004), *Measuring Rurality: Commuting Zones and Labor Market Areas*. Briefing Room at Economic Research Service (ERS) website. Available at <http://www.ers.usda.gov/Briefing/Rurality/LMACZ/>.
- US Department of Labor (2003), *Labor Market Areas, 2003*. Washington, US Department of Labor. Bureau of Labor Statistics.
- US Department of Labor (2013), *Labor Market Areas, 2013*. Washington, US Department of Labor. Bureau of Labor Statistics.

- Van der Haegen H., Pattyn M. (1980), *An operationalization of the concept of city region in West- European perspective: the Belgian city regions*, Tijdschrift voor Economische en Sociale Geografie, 61, 70–77.
- Van der Laan L, Schalke R. (2001), *Reality versus Policy: The Delineation and Testing of Local Labour Market and Spatial Policy Areas*. European Planning Studies, vol.9, No. 2, 201-221.
- Verbeke G., Molenberghs, G. (2000), *Linear mixed models for longitudinal data*. Springer Verlag, Berlin, New York.
- Vliegen M. (2003), *Territorial referencing, regionalisation and statistical description* Proceedings of the Potsdam SCORUS Satellite Conference 21.-23.08.2003. Workshop Three: Territorial Units. Potsdam Conference Proceedings, 93-100.
- Watts M. (2004), *Local Labour Markets in New South Wales: Fact or Fiction?*, Centre of Full Employment and Equity, WP 04-12, The University of Newcastle, Australia.
- West B., Welch K., Gałeccki A. (2007), *Linear Mixed Models. A Practical Guide Using Statistical Software*, Chapman & Hall/CRC, New York.
- Węziak D. (2007), *Wielopoziomowe modelowanie regresyjne w analizie danych*, Wiadomości Statystyczne, 9 (556), 1-12.
- Wolfinger, R.D., Tobias, R.D., Sall, J. (1994), *Computing Gaussian likelihoods and their derivatives for general linear mixed models*, SIAM Journal on Scientific Computing, 15(6),1294-1310
- Zagórska A., Rauziński R. (2011), *Rynek pracy kadr z wyższym wykształceniem na Śląsku Opolskim 1960-2025*, w: Wzrost gospodarczy - rynek pracy - polityka państwa, red. Kwiatkowski E., Kucharski L., Wydawnictwo Uniwersytetu Łódzkiego, Łódź.

SPIS TABEL

Tabela 1.1. Zasady wyznaczania lokalnych rynków pracy

Tabela 1.2. Wyjazdy i przyjazdy do pracy w przekroju miasto - wieś, Polska 2006 r.

Tabela 2.1. Klasyfikacja przykładowych metod delimitacji lokalnych rynków pracy, stosowanych w różnych krajach

Tabela 2.2. Porównanie europejskiego algorytmu regionalizacji, algorytmu Kristensena oraz podejścia taksonomicznego

Tabela 4.1. Współczynniki korelacji liniowej Pearsona między udziałem przyjeżdżających do pracy a potencjalnymi zmiennymi objaśniającymi, poziom gmin, 2006

Tabela 4.2. Współczynniki korelacji liniowej Pearsona między udziałem przyjeżdżających do pracy a potencjalnymi zmiennymi objaśniającymi, poziom powiatów, 2006

Tabela 4.3. Współczynniki korelacji liniowej Pearsona między udziałem przyjeżdżających do pracy a potencjalnymi zmiennymi objaśniającymi, poziom województw, 2006

Tabela 4.4. Ocena parametru strukturalnego oraz błędu standardowego szacunku, klasyczny model regresji liniowej natężenia przyjazdów do pracy bez zmiennych objaśniających, Polska, 2006 (wzór 4.5)

Tabela 4.5. Wybrane kryteria oceny jakości dopasowania klasycznego modelu regresji liniowej natężenia przyjazdów do pracy bez zmiennych objaśniających, Polska, 2006 (wzór 4.5)

Tabela 4.6. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, klasyczny model regresji liniowej natężenia przyjazdów do pracy, Polska, 2006 (wzór 4.6)

Tabela 4.7. Wybrane kryteria oceny jakości dopasowania klasycznego modelu regresji liniowej natężenia przyjazdów do pracy bez zmiennych objaśniających (wzór 4.6) oraz jego porównanie z etapem 0.0 (wzór 4.5)

Tabela 4.8. Ocena parametru strukturalnego oraz błędu standardowego szacunku, dwupoziomowy model natężenia przyjazdów do pracy, etap 1.1, Polska, 2006 (wzór 4.7)

Tabela 4.9. Wybrane kryteria oceny jakości dopasowania modelu dwupoziomowego oszacowanego w etapie 1.1 (wzór 4.7) oraz jego porównanie z etapem 0.0 (wzór 4.5)

Tabela 4.10. Ocena parametru strukturalnego oraz błędu standardowego szacunku, trzypoziomowy model natężenia przyjazdów do pracy, etap 1.2, Polska, 2006 (wzór 4.8)

Tabela 4.11. Wybrane kryteria oceny jakości dopasowania modelu trzypoziomowego oszacowanego w etapie 1.2 (wzór 4.8) oraz jego porównanie z etapem 1.1 (wzór 4.7)

Tabela 4.12. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 2, Polska, 2006 (wzór 4.9)

Tabela 4.13. Wybrane kryteria oceny jakości dopasowania modelu trzypoziomowego oszacowanego w etapie 2 (wzór 4.9) oraz jego porównanie z etapem 1.2 (wzór 4.8)

Tabela 4.14. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 3.1, poziom gmin, Polska, 2006 (wzór 4.10)

Tabela 4.15. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 3.1, poziom powiatów, Polska, 2006 (wzór 4.10)

Tabela 4.16. Wybrane kryteria oceny jakości dopasowania modelu trzypoziomowego oszacowanego w etapie 3.1 (wzór 4.10) oraz jego porównanie z etapem 2 (wzór 4.9)

Tabela 4.17. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 4.1, poziom gmin, Polska, 2006 (wzór 4.11)

Tabela 4.18. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 4.1, poziom powiatów, Polska, 2006 (wzór 4.11)

Tabela 4.19. Wybrane kryteria oceny jakości dopasowania modelu trzypoziomowego oszacowanego w etapie 4.1 (wzór 4.11) oraz jego porównanie z etapem 3.1 (wzór 4.10)

Tabela 4.20. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 4.2, poziom gmin, Polska, 2006 (wzór 4.12)

Tabela 4.21. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 4.1, poziom powiatów, Polska, 2006 (wzór 4.12)

Tabela 4.22. Wybrane kryteria oceny jakości dopasowania modelu trzypoziomowego oszacowanego w etapie 4.2 (wzór 4.12) oraz jego porównanie z etapem 4.1 (wzór 4.11)

Tabela 4.23. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.1, poziom gmin, Polska, 2006 (wzór 4.13)

Tabela 4.24. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.1, poziom powiatów, Polska, 2006 (wzór 4.13)

Tabela 4.25. Oceny współczynników kierunkowych interakcji między zmiennymi objaśniającymi z różnych poziomów oraz błędy standardowe szacunku, model natężenia przyjazdów do pracy, etap 5.1, Polska, 2006 (wzór 4.13)

Tabela 4.26. Wybrane kryteria oceny jakości dopasowania modelu trypoziomowego oszacowanego w etapie 5.1 (wzór 4.13) oraz jego porównanie z etapem 4.2 (wzór 4.12)

Tabela 4.27. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.2, poziom gmin, Polska, 2006 (wzór 4.14)

Tabela 4.28. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.2, poziom powiatów, Polska, 2006 (wzór 4.14)

Tabela 4.29. Oceny współczynników kierunkowych interakcji między zmiennymi objaśniającymi z różnych poziomów oraz błędy standardowe szacunku, model natężenia przyjazdów do pracy, etap 5.2, Polska, 2006 (wzór 4.14)

Tabela 4.30. Wybrane kryteria oceny jakości dopasowania modelu trypoziomowego oszacowanego w etapie 5.2 (wzór 4.14) oraz jego porównanie z etapem 5.1 (wzór 4.13)

Tabela 4.31. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.3, poziom gmin, Polska, 2006 (wzór 4.15)

Tabela 4.32. Oceny parametrów strukturalnych oraz błędów standardowych szacunku, model natężenia przyjazdów do pracy, etap 5.3, poziom powiatów, Polska, 2006 (wzór 4.15)

Tabela 4.33. Oceny współczynników kierunkowych interakcji między zmiennymi objaśniającymi z różnych poziomów oraz błędy standardowe szacunku, model natężenia przyjazdów do pracy, etap 5.3, Polska, 2006 (wzór 4.15)

Tabela 4.34. Wybrane kryteria oceny jakości dopasowania modelu trypoziomowego oszacowanego w etapie 5(2/3) (wzór (4.15) oraz jego porównanie z etapem 5(1/3) (wzór 4.14)

Tabela 4.35. Współczynniki korelacji liniowej Pearsona między potencjalnymi charakterystykami natężenia dojazdów do pracy a zmiennymi objaśniającymi, poziom par gmin, Polska, 2006

Tabela 4.36. Współczynniki korelacji liniowej Pearsona między natężeniem dojazdów do pracy a potencjalnymi zmiennymi objaśniającymi, poziom gmin wyjazdu oraz gmin przyjazdu, 2006

Tabela 4.37. Współczynniki korelacji liniowej Pearsona między natężeniem dojazdów do pracy a potencjalnymi zmiennymi objaśniającymi, poziom powiatów wyjazdu oraz powiatów przyjazdu, 2006

Tabela 4.38. Współczynniki korelacji liniowej Pearsona między natężeniem dojazdów do pracy a potencjalnymi zmiennymi objaśniającymi, poziom województw wyjazdu i województw przyjazdu, 2006

Tabela 4.39. Parametry testów analizy wariancji weryfikujących zróżnicowanie natężenia dojazdów do pracy we wszystkich rozważanych przekrojach terytorialnych, Polska, 2006

Tabela 4.40. Oceny parametru strukturalnego oraz błędu standardowego szacunku, klasyczny model regresji liniowej natężenia dojazdów do pracy pomiędzy parami gmin bez zmiennych objaśniających, Polska, 2006 (wzór 4.20)

Tabela 4.41. Wybrane kryteria oceny jakości dopasowania klasycznego modelu regresji liniowej natężenia dojazdów do pracy bez zmiennych objaśniających (wzór 4.5)

Tabela 4.42. Oceny parametrów strukturalnych oraz błędów standardowych szacunku klasycznego modelu regresji liniowej natężenia dojazdów do pracy pomiędzy parami gmin ze zmiennymi objaśniającymi, Polska, 2006 (wzór 4.21)

Tabela 4.43. Wybrane kryteria oceny jakości dopasowania klasycznego modelu regresji liniowej natężenia przyjazdów do pracy ze zmiennymi objaśniającymi (wzór 4.6) oraz jego porównanie z etapem 0.0 (wzór 4.5)

Tabela 4.44. Oceny parametru strukturalnego oraz błędu standardowego szacunku modeli natężenia przyjazdów do pracy, podetapy etapu pierwszego, Polska, 2006 (wzór 4.22)

Tabela 4.45. Oceny wariancji składników losowych modeli natężenia przyjazdów do pracy, podetapy etapu pierwszego, Polska, 2006 (wzór 4.22)

Tabela 4.46 Wybrane kryteria oceny jakości dopasowania modeli natężenia przyjazdów do pracy oszacowanych w etapie 1 (wzór 4.22) oraz ich porównanie z wcześniejszymi podetapami

Tabela 4.47. Oceny parametrów strukturalnych oraz błędów standardowych szacunku modelu natężenia przyjazdów do pracy, etap drugi, Polska, 2006 (wzór 4.23)

Tabela 4.48. Oceny wariancji składników losowych natężenia przyjazdów do pracy, etap drugi, Polska, 2006 (wzór 4.23)

Tabela 4.49. Wybrane kryteria oceny jakości dopasowania modelu natężenia przyjazdów do pracy oszacowanego w etapie 2 (wzór 4.23) oraz ich porównanie z etapem 1.6

Tabela 4.50. Oceny parametrów strukturalnych oraz błędów standardowych szacunku modelu natężenia przyjazdów do pracy, podetapy etapu trzeciego, Polska, 2006 (wzór 4.24)

Tabela 4.51. Oceny wariancji składników losowych modelu natężenia przyjazdów do pracy, podetapy etapu trzeciego, Polska, 2006 (wzór 4.24)

Tabela 4.52. Wybrane kryteria oceny jakości dopasowania modelu natężenia przyjazdów do pracy oszacowanego w etapie 3 (wzór 4.24) oraz ich porównanie z etapem 2

Tabela 4.53. Oceny parametrów strukturalnych oraz błędów standardowych szacunku modelu natężenia przyjazdów do pracy, podetapy etapu czwartego, Polska, 2006 (wzór 4.25)

Tabela 4.54. Oceny wariancji składników losowych modelu natężenia przyjazdów do pracy, podetapy etapu czwartego, Polska, 2006 (wzór 4.25)

Tabela 4.55. Wybrane kryteria oceny jakości dopasowania modelu natężenia przyjazdów do pracy oszacowanego w etapie 4 (wzór 4.25) oraz ich porównanie z etapem 3.5

Tabela 4.56. Oceny parametrów strukturalnych oraz błędów standardowych szacunku modelu natężenia przyjazdów do pracy, podetapy etapu piątego, Polska, 2006 (wzór 4.26)

Tabela 4.57. Oceny wariancji składników losowych modelu natężenia przyjazdów do pracy, podetapy etapu piątego, Polska, 2006 (wzór 4.26)

Tabela 4.58. Wybrane kryteria oceny jakości dopasowania modelu natężenia przyjazdów do pracy oszacowanego w etapie 5 (wzór 4.26) oraz ich porównanie z etapem 4.1

Tabela 4.59. Wartości średnie ocen błędów standardowych szacunku natężenia przyjazdów do pracy w gminach, wybrane modele, Polska, 2006

Tabela 4.60. Charakterystyki rozkładu błędów standardowych szacunku natężenia przyjazdów do pracy w gminach według modelu trypoziomowego w porównaniu z klasycznym modelem regresji liniowej, Polska, 2006

Tabela 4.61. Wartości współczynników determinacji modeli natężenia przyjazdów do pracy dla elementów na diagonalu macierzy migracji

Tabela 4.62. Wartości współczynników determinacji modeli natężenia dojazdów do pracy dla elementów poza diagonalą macierzy migracji

Tabela 4.62. Wartości współczynników determinacji modeli natężenia dojazdów do pracy dla elementów poza diagonalą macierzy migracji

Tabela 5.1. Parametry europejskiego algorytmu regionalizacji w wersji pierwotnej opracowanej dla Wielkiej Brytanii oraz zastosowanej w warunkach polskich dla roku 2006

Tabela 5.2. Liczba jednostek bazowych oraz wyznaczonych lokalnych rynków pracy w Polsce w roku 2006 na tle innych krajów

Tabela 5.3. Charakterystyki wielkości lokalnych rynków pracy wyznaczonych przy pomocy europejskiego algorytmu regionalizacji, Polska, 2006

Tabela 5.4. Ośrodki centralne stymulujące rozwój 20 największych pod względem liczby jednostek bazowych lokalnych rynków pracy, Polska, 2006

Tabela 5.5. Parametry algorytmu Kristensena w wersji pierwotnej opracowanej dla Danii oraz zastosowanej w warunkach polskich

Tabela 5.6. Charakterystyka wyników delimitacji według algorytmu Kristensena w Danii i w Polsce

Tabela 5.7. Charakterystyki wielkości lokalnych rynków pracy wyznaczonych przy pomocy podejścia taksonomicznego z mocnym i słabym kryterium łączenia, Polska, 2006

Tabela 5.8. Liczba oraz wielkość (mierzona liczbą gmin wchodzących w ich skład) lokalnych rynków otrzymanych przy zastosowaniu różnych algorytmów, Polska 2006

Tabela 5.9. Największe pod względem liczby jednostek bazowych lokalne rynki pracy według podejścia taksonomicznego z mocnym kryterium łączenia, polska, 2006

Tabela 5.10. Największe pod względem liczby jednostek bazowych lokalne rynki pracy według podejścia taksonomicznego ze słabym kryterium łączenia, polska, 2006

Tabela 5.11. Charakterystyki wielkości lokalnych rynków pracy wyznaczonych przy pomocy europejskiego algorytmu regionalizacji, Polska, 2011, oszacowania przeskalowane według wzoru (5.2)

Tabela 5.12. Charakterystyki wielkości lokalnych rynków pracy wyznaczonych przy pomocy europejskiego algorytmu regionalizacji, Polska, 2011, oszacowania przeskalowane według wzoru (5.3)

Tabela 5.13. Charakterystyki wielkości lokalnych rynków pracy wyznaczonych przy pomocy podejścia taksonomicznego z mocnym kryterium łączenia, Polska, 2011, oszacowania przeskalowane według wzoru (5.3)

Tabela 5.14. Największe pod względem liczby jednostek bazowych lokalne rynki pracy według podejścia taksonomicznego z mocnym kryterium łączenia, Polska 2011, oszacowania przeskalowane według wzoru (5.2)

Tabela 5.15. Charakterystyki wielkości lokalnych rynków pracy wyznaczonych przy pomocy podejścia taksonomicznego ze słabym kryterium łączenia, Polska, 2011, oszacowania przeskalowane według wzoru (5.3)

Tabela 5.16. Największe pod względem liczby jednostek bazowych lokalne rynki pracy według podejścia taksonomicznego ze słabym kryterium łączenia, Polska 2011, oszacowania przeskalowane według wzoru (5.2)

Tabela 5.17. Liczba oraz wielkość (mierzona liczbą gmin wchodzących w ich skład) lokalnych rynków otrzymanych przy zastosowaniu różnych algorytmów, Polska 2006

Tabela A.2.1. Szczegółowy skład lokalnych rynków pracy wyznaczonych przy pomocy europejskiego algorytmu regionalizacji oraz podejścia taksonomicznego z mocnym i słabym kryterium łączenia, Polska, 2006

SPIS SCHEMATÓW

Schemat. 2.1. Etapy Europejskiego Algorytmu Regionalizacji

Schemat. 2.2. Etapy algorytmu Kristensena

Schemat 3.1. Schemat jednostek klasycznej populacji o strukturze P-poziomowej

Schemat 3.2. Schemat jednostek populacji o strukturze P oraz Q poziomowej z dwoma kryteriami grupowania

SPIS WYKRESÓW

Wykres 1.1. Stosunek przyjeżdżających do wyjeżdżających według województw, Polska 2006 r.

Wykres 1.2. Wyjeżdżający do pracy w stosunku do ludności w wieku produkcyjnym według miejsca zamieszkania, Polska, 2006 r.

Wykres 1.3. Przyjeżdżający do pracy według obszaru zatrudnienia, Polska, 2006

Wykres 1.4. Stosunek liczby przyjeżdżających do pracy do ludności w wieku produkcyjnym, przekrój powiatów, Polska, 2006

Wykres 1.5. Stosunek liczby wyjeżdżających do pracy do ludności w wieku produkcyjnym, przekrój powiatów, Polska, 2006

Wykres 1.6. Stosunek liczby bezrobotnych do ludności w wieku produkcyjnym, przekrój powiatów, Polska, 2006

Wykres 4.1. Rozkład średniego natężenia przyjazdów do pracy w gminach, przekrój powiatów, Polska, 2006

Wykres 4.2. Średnie natężenie przyjazdów do pracy w gminach, przekrój województw, Polska, 2006

Wykres 4.3. Zależność pomiędzy zmienną objaśnianą i zmiennymi objaśniającymi w przekroju gmin, dla każdego z województw rozważana indywidualnie

Wykres 4.4. Średnie natężenie przyjazdów do pracy pomiędzy parami gmin, przekrój gmin wyjazdu, Polska, 2006

Wykres 4.5. Średnie natężenie wyjazdów do pracy pomiędzy parami gmin, przekrój gmin przyjazdu, Polska, 2006

Wykres 4.6. Średnie natężenie przyjazdów do pracy pomiędzy parami gmin, przekrój powiatów wyjazdu, Polska, 2006

Wykres 4.7. Średnie natężenie wyjazdów do pracy pomiędzy parami gmin, przekrój powiatów przyjazdu, Polska, 2006

Wykres 4.8. Średnie natężenie przyjazdów do pracy pomiędzy parami gmin, przekrój województw wyjazdu, Polska, 2006

Wykres 4.9. Średnie natężenie wyjazdów do pracy pomiędzy parami gmin, przekrój województw przyjazdu, Polska, 2006

Wykres 4.10. Rozkład ocen błędów standardowych szacunku natężenia przyjazdów dla różnych modeli, Polska, 2006

Wykres 5.1. Lokalne rynki pracy wraz z ośrodkami centralnymi według europejskiego algorytmu regionalizacji, Polska 2006

Wykres 5.2. Lokalne rynki pracy wraz z ośrodkami centralnymi według europejskiego algorytmu regionalizacji, Wielkopolska 2006

Wykres 5.3. Lokalne rynki pracy Poznania według europejskiego algorytmu regionalizacji, Wielkopolska 2006

Wykres 5.4. Lokalne rynki pracy według Algorytmu Kristensena, Polska 2006

Wykres 5.5. Lokalne rynki pracy według algorytmu Kristensena, Wielkopolska 2006

Wykres 5.6. Lokalne rynki pracy, podejście taksonomiczne, mocne kryterium łączenia, Polska 2006

Wykres 5.7. Lokalne rynki pracy, podejście taksonomiczne, słabe kryterium łączenia, Polska 2006

Wykres 5.8. Lokalne rynki pracy według podejścia taksonomicznego z mocnym kryterium łączenia, Wielkopolska 2006

Wykres 5.9. Lokalne rynki pracy według podejścia taksonomicznego ze słabym kryterium łączenia, Wielkopolska 2006

Wykres 5.10. Lokalne rynki pracy wraz z ośrodkami centralnymi według europejskiego algorytmu regionalizacji, Polska 2011, na podstawie „surowych” oszacowań

Wykres 5.11. Lokalne rynki pracy wraz z ośrodkami centralnymi według europejskiego algorytmu regionalizacji, Polska 2011, oszacowania przeskalowane według wzoru (5.2)

Wykres 5.12. Lokalne rynki pracy wraz z ośrodkami centralnymi według europejskiego algorytmu regionalizacji, Polska 2011, oszacowania przeskalowane według wzoru (5.3)

Wykres 5.13. Lokalne rynki pracy według podejścia taksonomicznego z mocnym kryterium łączenia, Polska 2011, oszacowania przeskalowane według wzoru (5.3)

Wykres 5.14. Lokalne rynki pracy według podejścia taksonomicznego ze słabym kryterium łączenia, Polska 2011, oszacowania przeskalowane

Aneks

Aneks A.1. Opis trudności obliczeniowych i kody źródłowe

OPIS TRUDNOŚCI OBLICZENIOWYCH PRZY KONSTRUKCJI MODELU WIELOPOZIOMOWEGO

W podrozdziale 4.2, w którym oszacowane zostały wartości spoza diagonalii macierzy migracji, konieczne było dokonanie oszacowań dla zbiorowości złożonej z aż 25 375 par gmin. Ponadto w modelu uwzględniono jedną zmienną objaśniającą z pierwszego poziomu (par gmin), po 18 zmiennych objaśniających z poziomów drugich ze względu na oba kryteria grupowania (gmina wyjazdu i gmina przyjazdu), po 10 zmiennych z poziomów trzecich ze względu na oba kryteria grupowania (powiat wyjazdu i powiat przyjazdu) oraz po 11 zmiennych objaśniających z poziomów czwartych ze względu na oba kryteria grupowania (województwo wyjazdu i województwo przyjazdu). W sumie zaproponowano więc aż 79 potencjalnych zmiennych objaśniających. Ponadto podczas konstrukcji modelu rozważono, zgodnie z procedurą opisaną w rozdziale 3.3, wprowadzenie wielu efektów losowych związanych z przynależnością par gmin do wspomnianych wyżej jednostek podziału administracyjnego oraz interakcji pomiędzy zmiennymi z różnych poziomów.

Parametry modelu oszacowane zostały zgodnie z metodą największej wiarygodności. Kolejne zmienne objaśniające oraz pozostałe rozszerzenia modelu wprowadzono w pięciu etapach, podzielonych na podetapy, zgodnie ze schematem postępowania opisanym w rozdziale trzecim. W ramach pod etapów rozszerzenia wprowadzono zgodnie z procedurą regresji krokowej ‘w przód’, jako kryterium poprawy dopasowania modelu przyjmując wynik testu ilorazu wiarygodności.

Obliczenia przeprowadzone zostały w programie R 2.15.2 z wykorzystaniem biblioteki lme4³⁷ (por. Bates, 2010; Bates, 2013a; Bates, 2013b; Biecek, 2011). Zgodnie z procedurą oprogramowaną w tej bibliotece parametry modeli wielopoziomowych szacowane są z wykorzystaniem macierzy rzadkich, co w istotnym stopniu skraca czas obliczeń w porównaniu z innymi algorytmami z tego zakresu (por. Bates, 2010). Wielkość zbiorowości, dla której przeprowadzono oszacowania oraz liczba potencjalnych

³⁷ Linear mixed-effects models using Eigen and S4 classes

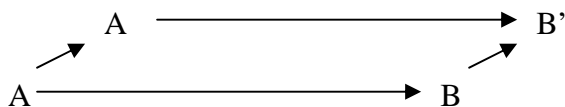
zmiennych objaśniających spowodowały jednak pomimo tego wyjątkowo długi czas obliczeń. Po przygotowaniu pierwszej wersji kodu czas obliczeń oszacowano na kilka miesięcy, pomimo wykorzystania platformy PL-GRID w Poznańskim Centrum Superkomputerowo - Sieciowym.

Ponieważ tak długi czas obliczeń uznano za nieakceptowany, podjęto próbę modyfikacji stosowanej procedury regresji krokowej 'w przód'. W tym celu wykorzystano własność opisaną w lemacie A.1.

Lemat A.1

Niech model B będzie rozwinięciem modelu A, a modele A' i B' rozwinięciami modeli A i B, odpowiednio o ten sam składnik.

Jeżeli test ilorazu wiarygodności nie wskazuje na istotną poprawę modelu A' w stosunku do modelu A, to również model B' nie jest według tego kryterium istotnie lepszy od modelu B



Zamiast szacować w każdym kroku regresji krokowej 'w przód' wszystkie możliwe rozszerzenia modeli o skomplikowanej już strukturze, w pierwszej kolejności wykluczano te z nich, o których na podstawie lematu A.1 wiadomo było, że nie poprawią jakości dopasowania modelu. W tym celu przed każdym krokiem tej procedury z aktualnego modelu wielopoziomowego usuwano komplikacje najbardziej wydłużające czas jego estymacji i do tak uproszczonego modelu dodawano rozważane rozszerzenia rozważane w danym kroku. Jeżeli nie poprawiały one w istotnym stopniu jakości dopasowania modelu prostszego, nie rozważano później dodania ich do modelu pełnego. Dzięki takiemu podejściu udało się znacznie skrócić czas obliczeń. Na przykład w pierwszym kroku regresji krokowej 'w przód,' przeprowadzonej w jedenastym podetapie etapu piątego konstrukcji omawianego w podrozdziale 4.3 modelu czteropoziomowego z dwoma kryteriami grupowania, zgodnie z podejściem tradycyjnym konieczne byłoby szacowanie 198 potencjalnych rozwinięć modelu oszacowanego w podetapie dziesiątym etapu piątego konstrukcji tego modelu. Ponieważ model na tym etapie był już mocno

skomplikowany, estymacja każdego z tych 198 rozwinięć trwałaby ponad pół godziny, co łącznie daje ponad 4 dni obliczeń w celu przeprowadzenia tylko jednego kroku procedury regresji krokowej ‘w przód’ w jednym z podetapów. W celu wykorzystania własności przedstawionej w lemacie A.1, przed rozpoczęciem właściwych obliczeń z modelu oszacowanego w podetapie dziesiątym etapu piątego usunięto najbardziej spowalniające obliczenia składniki losowe, związane ze zmiennymi objaśniającymi z wyższych poziomów. Do tak uproszczonego modelu dodano kolejno każde ze 198 rozszerzeń. Ponieważ estymowane modele były znacznie uproszczone, obliczenia te trwały łącznie mniej niż pół godziny. Ponadto tylko kilkanaście z wprowadzonych rozszerzeń spowodowało statystycznie istotną poprawę jakości dopasowania modelu prostszego. W związku z tym we właściwym kroku procedury regresji krokowej ‘w przód’ model oszacowany w podetapie dziesiątym etapu piątego rozwinięto tylko o te kilkanaście rozszerzeń (z których oszacowanie każdego trwało nieco ponad pół godziny).

W skali konstrukcji całego modelu czteropoziomowego z dwoma kryteriami grupowania opisanego w podrozdziale 4.3, czas obliczeń skrócił się z przynajmniej kilku miesięcy do kilku tygodni. Zaproponowanie procedury regresji krokowej, dostosowanej do modeli wielopoziomowych, która pozwoli jeszcze bardziej skrócić czas obliczeń będzie przedmiotem dalszych prac.

KOD ŹRÓDŁOWY

Kod oprogramowanych funkcji do regresji krokowej ‘w przód’, dostosowany do modeli wielopoziomowych oraz pakietu lme4, w którym wykorzystano przedstawiony wyżej lemat A.1:

```
#### funkcja dostosowana do wprowadzania efektów losowych związanych ze zmiennymi objaśniającymi
krok.lmer.4 <- function(
  start.model_0, xks, nut, alfa, max.iter=1){
  ##objaśnienia wartości wejściowych:
  # start.model_0 - model wejściowy
  # xks - macierz zmiennych dodawanych w danym podetapie do modelu
  # nut - identyfikator poziomu, dla którego wprowadzone mają być zmienne z macierzy xks
  # alfa - poziom istotności
  # max.iter - maksymalna liczba kroków

  xksy<-c(1:ncol(xks)) # wektor z numerami poencjalnych zmiennych
  model.basis <- start.model_0 # model obowiązujący w danym momencie procedury
  w<-0 # ile jest grup
  ww<-numeric(length(xksy)) # w tym wektorze zapisuje ile jest zmiennych w danej grupie, suma tego
  wektora daje kol?
  xksy2<-matrix(0,length(xksy)+1,length(xksy)) # w iwerszach tej macierzy zapisuje, ze do danej grupy
  trafiły następujące xksy
```

```

# maksymalna liczba powtórzeń nie powinna przekroczyć liczby potencjalnych zmiennych
if (max.iter > length(xksy)) max.iter <- length(xksy)
for(i in 1:max.iter){
  if(length(xksy)>0){
    ## deklaracja wektora blocks
    blocks<-matrix(0,length(xksy),w+1)
    blocks2<-numeric(w)
    for(j in 1:length(xksy)){
      blocks[j,1]<-paste("(0+",xks,"[,",xksy[j],"]",nut,")",sep=")
    }
    if(w>0){
      for(k in 1:w){
        baza<-paste("(0")
        for(j in 1:ww[k]){
          baza<-paste(baza,"+",xks,"[,",xksy2[k,j],"]")
        }
        blocks2[k]<-paste(baza,"|",nut,")")
        for(j in 1:length(xksy)){
          blocks[j,k+1]<-paste(baza,"+",xks,"[,",xksy[j],"]|",nut,")",sep=")
        }
      }
    }
  }
  models <- list()
  for(j in 1:length(xksy)){
    models[[j]] <- update(start.model_0, as.formula(paste(". ~ . + ", blocks[j,1])))
    if(w>0){
      for(l in 1:w){
        models[[j]] <- update(models[[j]], as.formula(paste(". ~ . + ", blocks2[l])))
      }
    }
  }
  LL <- unlist(lapply(models, logLik))
  modele<-cbind(c(1:length(xksy)),xksy,LL)
  modele<-as.matrix(modele)
  if(length(xksy)>1){
    modele2<-modele[order(modele[,3],decreasing=TRUE),]
  }else{
    modele2<-modele
  }
  dl<-length(xksy)
  model.basis0<-model.basis
  for(j in 1:dl){
    if((dl-j+1)==modele2[1,1]){
      xksy <- xksy[-modele2[1,1]]
      if(anova(model.basis, models[[modele2[1,1]]])[2,7] < alfa){
        if(w>0){
          models3 <- list()
          for(k in 1:w){
            models3[[k]] <- update(start.model_0, as.formula(paste(". ~ . + ", blocks[modele2[1,1],k+1])))
            for(l in 1:w){
              if(l!=k){
                models3[[k]] <- update(models3[[k]], as.formula(paste(". ~ . + ", blocks2[l])))
              }
            }
          }
        }
      }
    }
  }
  LL2 <-numeric(w)
  for(k in 1:w){
    mod1<-models3[[k]]
    mod2<-models[[modele2[1,1]]]
  }
}

```

```

    LL2[k]<-anova(mod1, mod2)[2,7]
  }
  modele3<-cbind(c(1:length(w)),LL2)
  modele3<-as.matrix(modele3)
  if(w>1){
    modele3<-modele3[order(modele3[,2],decreasing=FALSE),]
  }
  if(modele3[1,2]< alfa){
    model.basis <- models3[[modele3[1,1]]]
    ww[modele3[1,1]]<-ww[modele3[1,1]]+1
    xksy2[modele3[1,1],ww[modele3[1,1]]]<-modele2[1,2]
  }else{
    model.basis <-models[[modele2[1,1]]]
    w<-w+1
    ww[w]<-1
    xksy2[w,1]<-modele2[1,2]
  }
  }else{
    model.basis <-models[[modele2[1,1]]]
    w<-w+1
    ww[w]<-1
    xksy2[w,1]<-modele2[1,2]
  }
  }
  }else{
    if(anova(model.basis0, models[[[(dl-j+1)]]][2,7] > alfa){
      xksy <- xksy[-(dl-j+1)]
    }
  }
  }
  }
  }
  }
  if(w>0){
    blocks_baz<-numeric(w)
    for(k in 1:w){
      blocks_baz[k]<-paste("0")
      for(j in 1:ww[k]){
        blocks_baz[k]<-paste(blocks_baz[k],"+",xks,"[",xksy2[k,j],"")
      }
    }
  }else{
    blocks_baz<-0
  }
  }
  ## zwrot parametrów najlepszego modelu:
  wynik <- list()
  wynik[[1]]<-model.basis
  wynik[[2]]<-w
  wynik[[3]]<-ww
  wynik[[4]]<-blocks_baz
  return(wynik)
}

```

###funkcja dostosowana do wprowadzania interakcji pomiędzy zmiennymi z różnych poziomów w etapie 5
 krok.lmer.5 <- function(

```

  start.model,start.model.0, xks1, xks2, nut1, nut2, alfa, max.iter=1){
  ##objaśnienia wartości wejściowych:
  # start.model - model wejściowy (pełna wersja)
  # start.model.0 - uproszczony model wejściowy
  # xks1 - macierz zmiennych z niższego poziomu
  # xks2 - macierz zmiennych z wyższego poziomu

```

```

# nut1 - identyfikator niższego poziomu
# nut2 - identyfikator wyższego poziomu
# alfa - poziom istotności
# max.iter - maksymalna liczba kroków
model.basis <- start.model
model.basis.0 <- start.model.0
xks_1<-c(1:ncol(xks1))
xks_2<-c(1:ncol(xks2))
blocks<-matrix(0,length(xks_1),length(xks_2))
for(i in 1:length(xks_1)){
  for(j in 1:length(xks_2)){
    blocks[i,j]<-paste("nut1,[",xks_1[i],"]:",nut2,[",xks_2[j],"]",sep=")
  }
}
blocks<-as.vector(blocks)
if (max.iter > length(blocks)) max.iter <- length(blocks)
for(i in 1:max.iter){
  if(length(blocks)>0){
    dl<-length(blocks)
    for(d in 1:dl){
      modelik <- update(model.basis.0, as.formula(paste(". ~ . + ", blocks[(dl-d+1)])))
      if(anova(model.basis.0,modelik)[2,7]>alfa){
        blocks <- blocks[-(dl-d+1)]
      }
      rm("modelik")
    }
    if(length(blocks)>0){
      dl<-length(blocks)
      modele<-matrix(0,length(blocks),2)
      for(j in 1:length(blocks)){
        modele[j,1]<-j
        modelik<-update(model.basis, as.formula(paste(". ~ . + ", blocks[j])))
        modele[j,2]<-anova(model.basis, modelik)[2,7]
        rm("modelik")
      }
      if(dl>1){
        modele2<-modele[order(modele[,2],decreasing=FALSE),]
      }else{
        modele2<-modele
      }
      modele2<-as.matrix(modele2)
      dl<-length(blocks)
      for(j in 1:dl){
        if((dl-j+1)==modele2[1,1]){
          if(modele2[1,2] < alfa){
            model.basis <- update(model.basis, as.formula(paste(". ~ . + ", blocks[(modele2[1,1])]))
            model.basis.0<-update(model.basis.0, as.formula(paste(". ~ . + ", blocks[(modele2[1,1])]))
          }
          blocks <- blocks[-(modele2[1,1])]
        }else{
          if(modele[(dl-j+1),2] > alfa){
            blocks <- blocks[-(dl-j+1)]
          }
        }
      }
    }
  }
}
return(model.basis)
}

```


Oprócz powyżej przedstawionych funkcji znajdujących zastosowanie przy konstrukcji modeli wielopoziomowych samodzielnie oprogramowano wszystkie zastosowane w pracy algorytmy delimitacji lokalnych rynków pracy. Przy tym kody europejskiego algorytmu regionalizacji oraz algorytmu Kristensena przygotowane zostały według szczegółowego opisu słownego zawartego w pracach autorów tych metod (por. Coombes, Green, Openshaw, 1986; Kristensen, 1998). Podejście taksonomiczne w całości zostało opracowane przez autorkę. Kody te nie zostały zamieszone w pracy ze względu na ich dużą objętość, jednak autorka może je w każdej chwili udostępnić osobom zainteresowanym.

Aneks A.2. Skład wyznaczonych lokalnych rynków pracy

Tabela A.2.1. Szczegółowy skład lokalnych rynków pracy wyznaczonych przy pomocy europejskiego algorytmu regionalizacji oraz podejścia taksonomicznego z mocnym i słabym kryterium grupowania, Polska, 2006

nazwa gminy	gmina				numer lokalnego rynku pracy		
	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Bolesławiec (1)	201011	24	10516	26660	1	21	22
Bolesławiec (2)	201022	288	1231	8240	1	21	22
Gromadka (2)	201032	268	591	3541	1	21	22
Nowogrodziec - miasto (4)	201044	16	850	2731	1	21	22
Nowogrodziec - obszar wiejski (5)	201045	160	959	6706	1	21	22
Osiecznica (2)	201052	438	958	4497	1	21	22
Warta Bolesławiecka (2)	201062	110	640	5195	1	21	22
Bielawa (1)	202011	36	3880	20171	2	93	139
Dzierżonów (1)	202021	20	7919	22359	2	93	139
Pieszycy (1)	202031	64	997	6033	2	93	139
Piława Górna (1)	202041	21	602	4227	2	93	139
Dzierżonów (2)	202052	141	471	6162	2	93	139
Łagiewniki (2)	202062	125	450	4688	2	93	139
Niemcza - miasto (4)	202074	20	346	1950	2	93	139
Niemcza - obszar wiejski (5)	202075	52	227	1876	2	93	139
Głogów (1)	203011	35	16043	48176	9	58	74
Głogów (2)	203022	85	460	3612	9	58	74
Jerzmanowa (2)	203032	63	360	2087	9	58	74
Kotła (2)	203042	128	285	2606	9	58	74
Pęczaw (2)	203052	64	124	1456	9	58	74
Żukowice (2)	203062	68	431	2265	9	58	74
Góra - miasto (4)	204014	14	2665	8219	221	22	229
Góra - obszar wiejski (5)	204015	252	421	5043	221	22	229
Jemielno (2)	204022	124	135	1898	221	22	229
Niechlów (2)	204032	151	375	3119	221	22	229
Wąsosz - miasto (4)	204044	3	394	1804	221	49	62
Wąsosz - obszar wiejski (5)	204045	194	236	2808	221	49	62
Jawor (1)	205011	19	5333	16463	19	59	137
Bolków - miasto (4)	205024	8	990	3437	4	59	137
Bolków - obszar wiejski (5)	205025	145	202	3573	4	59	137
Męcinka (2)	205032	148	460	3007	19	59	137
Mściwojów (2)	205042	72	371	2521	19	59	137
Paszowice (2)	205052	101	293	2513	19	59	137
Wądroże Wielkie (2)	205062	89	271	2541	19	58	58
Karpacz (1)	206011	38	984	3334	3	71	92
Kowary (1)	206021	37	2057	7730	3	71	92
Piechowice (1)	206031	43	1233	4378	18	71	92
Szklarska Poręba (1)	206041	75	1297	4708	18	71	92
Janowice Wielkie (2)	206052	58	675	2686	18	71	92
Jeżów Sudecki (2)	206062	94	484	4253	18	71	92
Mysłakowice (2)	206072	88	1525	6510	3	71	92
Podgórzyn (2)	206082	83	759	5216	3	71	92
Stara Kamienica (2)	206092	111	697	3439	18	71	92
Kamienna Góra (1)	207011	18	5161	13596	4	59	104
Kamienna Góra (2)	207022	158	1094	5574	4	59	104
Lubawka - miasto (4)	207034	22	964	4294	4	59	104
Lubawka - obszar wiejski (5)	207035	116	217	3255	4	59	104
Marciszów (2)	207042	82	404	3019	4	59	104
Duszniki-Zdrój (1)	208011	22	1102	3299	6	67	88

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Kłodzko (1)	208021	25	9174	18184	5	67	88
Kudowa-Zdrój (1)	208031	34	1422	6527	6	67	88
Nowa Ruda (1)	208041	37	2990	15519	5	67	88
Polanica-Zdrój (1)	208051	17	1834	4333	6	67	88
Bystrzyca Kłodzka - miasto (4)	208064	11	1999	7033	5	67	88
Bystrzyca Kłodzka - obszar wiejski (5)	208065	327	566	5865	5	67	88
Kłodzko (2)	208072	253	1020	11072	5	67	88
Lądek-Zdrój - miasto (4)	208084	20	1194	3986	5	67	173
Lądek-Zdrój - obszar wiejski (5)	208085	97	83	1664	5	67	173
Lewin Kłodzki (2)	208092	52	133	1181	6	67	88
Międzylesie - miasto (4)	208104	14	474	1840	5	67	88
Międzylesie - obszar wiejski (5)	208105	175	272	3067	5	67	88
Nowa Ruda (2)	208112	140	1236	7854	5	67	88
Radków - miasto (4)	208124	15	260	1560	5	67	88
Radków - obszar wiejski (5)	208125	125	610	4483	5	67	88
Stronie Śląskie - miasto (4)	208134	2	1221	4166	5	67	173
Stronie Śląskie - obszar wiejski (5)	208135	144	136	1123	5	67	173
Szczytna - miasto (4)	208144	80	380	3645	6	67	88
Szczytna - obszar wiejski (5)	208145	52	266	1408	6	67	88
Chojnów (1)	209011	5	2324	9682	19	58	58
Chojnów (2)	209022	231	630	5889	19	58	58
Krotoszyce (2)	209032	67	190	1861	19	58	58
Kunice (2)	209042	93	776	3300	19	58	58
Legnickie Pole (2)	209052	85	888	3240	19	58	58
Miłkowice (2)	209062	87	394	3948	19	58	58
Prochowice - miasto (4)	209074	10	452	2557	19	58	58
Prochowice - obszar wiejski (5)	209075	93	202	2485	19	58	58
Ruja (2)	209082	73	222	1604	19	58	58
Lubań (1)	210011	16	4691	14565	17	45	135
Świeradów-Zdrój (1)	210021	21	927	2977	7	108	237
Leśna - miasto (4)	210034	9	798	2971	17	45	135
Leśna - obszar wiejski (5)	210035	96	274	3793	17	45	135
Lubań (2)	210042	142	401	4297	17	45	135
Olszyna - miasto (4) *	210054	20	812	3047	17	45	135
Olszyna - obszar wiejski (5) *	210055	27	107	1310	7	45	135
Platerówka (2)	210062	48	85	1082	17	45	49
Siekierczyn (2)	210072	49	486	2894	17	45	135
Lubin (1)	211011	41	20795	52736	9	58	74
Lubin (2)	211022	290	704	7367	9	58	74
Rudna (2)	211032	217	2216	4535	9	58	74
Ścinawa - miasto (4)	211044	14	557	4044	9	58	74
Ścinawa - obszar wiejski (5)	211045	150	104	2864	9	58	74
Gryfów Śląski - miasto (4)	212014	7	1126	4642	7	108	178
Gryfów Śląski - obszar wiejski (5)	212015	60	718	1960	7	108	178
Lubomierz - miasto (4)	212024	8	283	1206	7	108	178
Lubomierz - obszar wiejski (5)	212025	123	48	2649	7	71	92
Lwówek Śląski - miasto (4)	212034	17	2264	6344	1	21	22
Lwówek Śląski - obszar wiejski (5)	212035	222	918	5270	1	21	22
Mirsk - miasto (4)	212044	15	381	2697	7	108	237
Mirsk - obszar wiejski (5)	212045	172	642	3119	7	108	237
Wleń - miasto (4)	212054	7	230	1288	18	71	92
Wleń - obszar wiejski (5)	212055	79	134	1744	18	71	92
Cieszków (2)	213012	101	326	2897	8	49	172
Krośnice (2)	213022	179	1410	5117	8	49	172
Milicz - miasto (4)	213034	14	3266	7909	8	49	172
Milicz - obszar wiejski (5)	213035	421	1298	7718	8	49	172
Oleśnica (1)	214011	21	7558	24081	20	49	62
Bierutów - miasto (4)	214024	8	938	3375	20	49	62
Bierutów - obszar wiejski (5)	214025	139	213	3162	20	49	62

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Dobroszyce (2)	214032	132	687	3924	20	49	62
Dziadowa Kłoda (2)	214042	106	475	2945	211	49	38
Międzybórz - miasto (4)	214054	6	1140	1563	8	49	38
Międzybórz - obszar wiejski (5)	214055	82	681	1637	8	49	38
Oleśnica (2)	214062	242	1389	7392	20	49	62
Syców - miasto (4)	214074	17	2212	7184	211	49	38
Syców - obszar wiejski (5)	214075	128	646	3542	211	49	38
Twardogóra - miasto (4)	214084	8	1622	4621	8	49	38
Twardogóra - obszar wiejski (5)	214085	160	1923	3883	20	49	38
Oława (1)	215011	27	7589	20422	20	49	62
Domaniów (2)	215022	94	269	3342	20	49	62
Jelcz-Laskowice - miasto (4)	215034	17	5044	11167	20	49	62
Jelcz-Laskowice - obszar wiejski (5)	215035	151	1776	4020	20	49	62
Oława (2)	215042	235	1425	8617	20	49	62
Chocianów - miasto (4)	216014	7	1450	5487	9	58	74
Chocianów - obszar wiejski (5)	216015	214	326	2821	9	58	74
Gaworzycy (2)	216022	77	187	2462	9	58	74
Grębcice (2)	216032	122	248	3326	9	58	74
Polkowice - miasto (4)	216044	24	17434	14856	9	58	74
Polkowice - obszar wiejski (5)	216045	144	4812	2354	9	58	74
Przemków - miasto (4)	216054	6	709	4263	9	58	74
Przemków - obszar wiejski (5)	216055	101	120	1364	9	58	74
Radwanice (2)	216062	84	449	2927	9	58	74
Borów (2)	217012	99	471	3266	20	49	62
Kondratowice (2)	217022	98	317	2973	20	49	62
Przeworno (2)	217032	112	470	3254	16	49	62
Strzelin - miasto (4)	217044	10	3281	8279	20	49	62
Strzelin - obszar wiejski (5)	217045	161	1435	5957	20	49	62
Wiązów - miasto (4)	217054	9	440	1465	20	49	62
Wiązów - obszar wiejski (5)	217055	133	89	3194	20	49	62
Kostomłoty (2)	218012	145	433	4435	14	49	62
Małczyce (2)	218022	53	419	3821	20	49	62
Miękinia (2)	218032	179	2446	7667	20	49	62
Środa Śląska - miasto (4)	218044	15	3343	5868	20	49	62
Środa Śląska - obszar wiejski (5)	218045	201	1852	6689	20	49	62
Udanin (2)	218052	111	319	3570	10	49	62
Świdnica (1)	219011	22	14828	39643	10	59	51
Świebodzice (1)	219021	30	3826	15210	12	59	75
Dobromierz (2)	219032	86	611	3536	10	59	75
Jaworzyna Śląska - miasto (4)	219044	4	1744	3360	10	59	51
Jaworzyna Śląska - obszar wiejski (5)	219045	63	198	3378	10	59	51
Marcinowice (2)	219052	95	291	4158	10	59	51
Strzegom - miasto (4)	219064	20	3421	11143	10	59	51
Strzegom - obszar wiejski (5)	219065	125	1725	6550	10	59	51
Świdnica (2)	219072	208	1278	9933	10	59	51
Żarów - miasto (4)	219084	6	2049	4542	10	59	51
Żarów - obszar wiejski (5)	219085	82	148	3583	10	59	51
Oborniki Śląskie - miasto (4)	220014	14	1580	5572	11	49	62
Oborniki Śląskie - obszar wiejski (5)	220015	140	822	6183	11	49	62
Prusice - miasto (4) *	220024	11	315	1487	11	49	62
Prusice - obszar wiejski (5) *	220025	147	301	4502	11	49	62
Trzebnica - miasto (4)	220034	8	3840	8259	11	49	62
Trzebnica - obszar wiejski (5)	220035	192	444	6215	11	49	62
Wisznia Mała (2)	220042	103	814	5410	11	49	62
Zawonia (2)	220052	117	330	3550	11	49	62
Żmigród - miasto (4)	220064	10	1105	4374	11	49	62
Żmigród - obszar wiejski (5)	220065	283	454	5219	11	49	62
Boguszów-Gorce (1)	221011	27	793	10516	12	59	75
Jedlina-Zdrój (1)	221021	17	493	3216	12	59	75

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Szczawno-Zdrój (1)	221031	15	1290	3426	12	59	75
Czarny Bór (2)	221042	66	374	3073	4	59	75
Głuszyca - miasto (4)	221054	16	670	4524	12	59	75
Głuszyca - obszar wiejski (5)	221055	46	219	1452	12	59	75
Mioszów - miasto (4)	221064	10	668	2916	12	59	75
Mioszów - obszar wiejski (5)	221065	66	196	1853	12	59	75
Stare Bogaczowice (2)	221072	88	343	2726	12	59	75
Walim (2)	221082	79	416	3654	12	59	75
Wałbrzych (1) *	221091	85	31630	81863	12	59	75
Brzeg Dolny - miasto (4)	222014	17	3505	8462	13	49	194
Brzeg Dolny - obszar wiejski (5)	222015	77	449	2182	13	49	194
Wińsko (2)	222022	249	466	5285	13	49	194
Wołów - miasto (4)	222034	19	4044	8047	13	49	194
Wołów - obszar wiejski (5)	222035	313	1129	6515	13	49	194
Czernica (2)	223012	84	693	6281	20	49	62
Długołęka (2)	223022	213	2407	14049	20	49	62
Jordanów Śląski (2)	223032	57	194	1967	20	49	62
Kąty Wrocławskie - miasto (4)	223044	6	2529	3584	14	49	62
Kąty Wrocławskie - obszar wiejski (5)	223045	171	1412	8170	14	49	62
Kobierzyce (2)	223052	149	7361	9054	20	49	62
Mietków (2)	223062	83	262	2455	14	49	62
Sobótka - miasto (4)	223074	32	1056	4543	20	49	62
Sobótka - obszar wiejski (5)	223075	104	456	3569	20	49	62
Siechnice	223084	16	2549	9284	20	49	62
Święta Katarzyna	223085	87	1923	3883	20	49	62
Żórawina (2)	223092	120	703	5267	20	49	62
Bardo - miasto (4)	224014	5	279	1899	15	67	84
Bardo - obszar wiejski (5)	224015	68	371	1782	15	67	84
Cieplowody (2)	224022	77	183	2028	15	67	84
Kamieniec Żąbkowski (2)	224032	97	1207	5750	15	67	84
Stoszowice (2)	224042	111	441	3616	15	67	84
Żąbkowice Śląskie - miasto (4)	224054	14	4293	10497	15	67	84
Żąbkowice Śląskie - obszar wiejski (5)	224055	132	311	4442	15	67	84
Ziębice - miasto (4)	224064	15	1833	5985	16	67	84
Ziębice - obszar wiejski (5)	224065	208	295	5646	16	67	84
Złoty Stok - miasto (4)	224074	8	576	1902	15	67	84
Złoty Stok - obszar wiejski (5)	224075	67	153	1157	15	67	84
Zawidów (1)	225011	6	720	2873	17	45	49
Zgorzelec (1)	225021	16	7429	21358	17	45	49
Bogatynia - miasto (4)	225034	60	10101	12828	17	45	49
Bogatynia - obszar wiejski (5)	225035	76	506	3902	17	45	49
Pieńsk - miasto (4)	225044	10	1074	3760	17	45	49
Pieńsk - obszar wiejski (5)	225045	101	200	2244	17	45	49
Sulików (2)	225052	95	476	3855	17	45	49
Węgliniec - miasto (4)	225064	9	1272	2024	1	49	62
Węgliniec - obszar wiejski (5)	225065	329	487	3541	1	49	62
Zgorzelec (2)	225072	137	486	5392	17	45	49
Wojcieszów (1)	226011	32	243	2631	18	71	92
Złotoryja (1)	226021	12	3523	11345	19	58	58
Pielgrzymka (2)	226032	105	169	3042	19	58	58
Świerzawa - miasto (4)	226044	2	351	1676	19	71	92
Świerzawa - obszar wiejski (5)	226045	158	304	3451	19	71	92
Zagrodno (2)	226052	122	387	3602	19	58	58
Złotoryja (2)	226062	145	1038	4489	19	58	58
Jelenia Góra (1)	261011	109	23201	56876	18	71	92
Legnica (1)	262011	56	29628	70234	19	58	58
Wrocław (1)	264011	293	198262	425208	20	49	62
Aleksandrów Kujawski (1)	401011	7	2798	8175	21	85	119
Ciechocinek (1)	401021	15	2915	6749	21	85	119

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Nieszawa (1)	401031	10	267	1303	21	85	119
Aleksandrów Kujawski (2)	401042	131	551	6879	21	85	119
Bądkowo (2)	401052	80	195	2801	34	76	99
Koneck (2)	401062	68	188	1997	21	85	119
Raciążek (2)	401072	33	183	1978	21	85	119
Wąganiec (2)	401082	55	528	2851	21	76	99
Zakrzewo (2)	401092	76	352	2188	22	85	119
Brodnica (1)	402011	23	9709	17721	23	14	15
Bobrowo (2)	402022	145	296	3871	23	14	15
Brodnica (2)	402032	127	491	4122	23	14	15
Brzozie (2)	402042	94	244	2148	23	14	15
Górzno - miasto (4)	402054	3	137	830	23	14	15
Górzno - obszar wiejski (5)	402055	117	353	1499	23	14	15
Bartniczka (2) *	402062	83	546	2785	23	14	15
Jabłonowo Pomorskie - miasto (4)	402074	3	815	2282	23	14	15
Jabłonowo Pomorskie - obszar wiejski (5)	402075	132	178	3226	23	14	15
Osiek (2)	402082	75	186	2351	23	14	15
Świedziebna (2)	402092	104	212	2985	23	14	15
Zbiczno (2)	402102	134	160	2841	23	14	15
Białe Błota (2)	403012	122	4265	9426	36	50	63
Dąbrowa Chełmińska (2)	403022	125	730	4545	36	50	63
Dobrecz (2)	403032	130	874	5855	36	50	63
Koronowo - miasto (4)	403044	29	2231	7182	36	50	63
Koronowo - obszar wiejski (5)	403045	384	622	7876	36	50	63
Nowa Wieś Wielka (2)	403052	148	1554	5389	36	50	63
Osielsko (2)	403062	103	1765	6152	36	50	63
Sicienko (2)	403072	180	554	5804	36	50	63
Solec Kujawski - miasto (4)	403084	19	3472	9652	36	50	63
Solec Kujawski - obszar wiejski (5)	403085	156	250	667	36	50	63
Chełmno (1)	404011	14	4377	13308	31	129	230
Chełmno (2)	404022	113	337	3193	31	129	230
Kijewo Królewskie (2)	404032	72	1172	2756	31	129	230
Lisewo (2)	404042	86	415	3294	31	129	230
Papowo Biskupie (2)	404052	70	287	2760	31	72	95
Stolno (2)	404062	99	799	3195	31	129	230
Unisław (2)	404072	73	450	4331	36	50	63
Golub-Dobrzyń (1)	405011	8	2700	8344	37	72	95
Ciechocin (2)	405022	101	247	2303	37	72	95
Golub-Dobrzyń (2)	405032	198	422	5073	37	72	95
Kowalewo Pomorskie - miasto (4)	405044	4	809	2570	37	72	95
Kowalewo Pomorskie - obszar wiejski (5)	405045	137	467	4463	37	72	95
Radomin (2)	405052	81	230	2500	37	72	95
Zbójno (2)	405062	84	378	2578	37	72	95
Grudziądz (2)	406012	165	1215	6649	36	50	138
Gruta (2)	406022	124	361	4100	36	50	138
Łasin - miasto (4)	406034	5	691	2149	36	50	138
Łasin - obszar wiejski (5)	406035	132	318	3105	36	50	138
Radzyń Chełmiński - miasto (4)	406044	2	288	1212	36	50	138
Radzyń Chełmiński - obszar wiejski (5)	406045	89	81	1780	36	50	138
Rogóźno (2)	406052	116	218	2544	36	50	138
Świecie nad Osą (2)	406062	95	342	2710	36	50	138
Inowrocław (1)	407011	30	18528	51179	25	109	180
Dąbrowa Biskupia (2)	407022	147	658	3286	25	109	180
Gniewkowo - miasto (4)	407034	9	1854	4726	25	109	180
Gniewkowo - obszar wiejski (5)	407035	171	1406	4767	25	109	180
Inowrocław (2)	407042	172	1223	6964	25	109	180
Janikowo - miasto (4)	407054	10	2981	5822	25	109	180
Janikowo - obszar wiejski (5)	407055	82	219	2806	25	109	180
Kruszwica - miasto (4)	407064	7	2040	6119	25	109	180

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Kruszwica - obszar wiejski (5)	407065	255	1046	6681	25	109	180
Pakość - miasto (4)	407074	3	984	3711	25	109	180
Pakość - obszar wiejski (5)	407075	83	356	2657	25	109	180
Rojewo (2)	407082	120	489	2905	25	109	180
Złotniki Kujawskie (2)	407092	136	721	5819	25	50	63
Lipno (1)	408011	11	3841	9560	24	72	216
Bobrowniki (2)	408022	95	132	1849	34	76	99
Chrostkowo (2)	408032	74	155	1816	24	72	216
Dobrzyń nad Wisłą - miasto (4)	408044	5	321	1428	34	76	99
Dobrzyń nad Wisłą - obszar wiejski (5)	408045	110	209	3538	34	76	99
Kikót (2)	408052	98	359	4405	24	72	216
Lipno (2)	408062	211	412	6786	24	72	216
Skępe - miasto (4) *	408074	7	689	2201	24	72	216
Skępe - obszar wiejski (5) *	408075	172	169	2332	24	72	216
Tłuchowo (2)	408082	99	210	2732	24	23	24
Wielgie (2)	408092	134	353	3951	24	76	99
Dąbrowa (2)	409012	110	310	2982	25	109	201
Jeziora Wielkie (2)	409022	124	239	3089	25	109	201
Mogilno - miasto (4)	409034	8	3404	8199	25	109	201
Mogilno - obszar wiejski (5)	409035	248	947	7754	25	109	201
Strzelno - miasto (4)	409044	4	1401	3912	25	109	201
Strzelno - obszar wiejski (5)	409045	181	850	3902	25	109	201
Kcynia - miasto (4)	410014	7	701	3016	36	50	63
Kcynia - obszar wiejski (5)	410015	290	624	5582	36	50	63
Mrocza - miasto (4)	410024	4	462	2708	36	50	63
Mrocza - obszar wiejski (5)	410025	147	280	3068	36	50	63
Nakło nad Notecią - miasto (4)	410034	11	4513	12456	36	50	63
Nakło nad Notecią - obszar wiejski (5)	410035	176	2235	8356	36	50	63
Sadki (2)	410042	154	535	4479	36	50	63
Szubin - miasto (4)	410054	7	1807	6045	36	50	63
Szubin - obszar wiejski (5)	410055	324	837	8615	36	50	63
Radziejów (1)	411011	6	1745	3805	26	76	149
Bytoń (2)	411022	73	191	2253	26	76	149
Dobre (2)	411032	71	349	3393	26	76	149
Osiężciny (2)	411042	123	607	4992	26	76	99
Piotrków Kujawski - miasto (4) *	411054	10	767	2860	26	76	149
Piotrków Kujawski - obszar wiejski (5) *	411055	128	63	3065	26	76	149
Radziejów (2)	411062	93	182	2598	26	76	149
Topólka (2)	411072	103	266	3059	33	76	99
Rypin (1)	412011	11	4591	10803	27	101	162
Brzuze (2)	412022	86	373	3169	27	101	162
Rogowo (2)	412032	141	288	2735	27	101	162
Rypin (2)	412042	132	767	4439	27	101	162
Skrwilno (2)	412052	123	266	3602	27	101	162
Wąpielsk (2)	412062	93	157	2448	27	101	162
Kamień Krajeński - miasto (4)	413014	4	467	1501	28	52	65
Kamień Krajeński - obszar wiejski (5)	413015	159	277	2824	28	52	93
Sępólno Krajeńskie - miasto (4)	413024	6	2357	5983	28	52	65
Sępólno Krajeńskie - obszar wiejski (5)	413025	223	354	4118	28	52	65
Sośno (2)	413032	163	383	3182	28	52	65
Więcbork - miasto (4)	413044	4	1295	3639	28	52	65
Więcbork - obszar wiejski (5)	413045	232	324	4496	28	52	65
Bukowiec (2)	414012	111	624	3293	31	50	63
Dragacz (2)	414022	112	564	4553	36	50	138
Drzycim (2)	414032	108	520	3175	30	50	63
Jeżewo (2)	414042	156	1004	4970	30	50	63
Lniano (2)	414052	88	602	2542	30	50	63
Nowe - miasto (4)	414064	4	1817	4066	29	50	138
Nowe - obszar wiejski (5)	414065	103	502	2789	29	50	138

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Osie (2)	414072	210	1325	3300	30	50	63
Pruszcz (2)	414082	142	783	5833	36	50	63
Świecie - miasto (4)	414094	12	8111	17297	31	50	63
Świecie - obszar wiejski (5)	414095	163	751	4562	31	50	63
Świekatowo (2)	414102	64	179	2144	36	50	63
Warlubie (2)	414112	201	811	4160	29	50	138
Chełmża (1)	415011	8	2045	9749	37	72	95
Chełmża (2)	415022	179	654	5941	37	72	95
Czernikowo (2)	415032	170	556	5162	37	72	95
Lubicz (2)	415042	106	2549	11375	37	72	95
Łubianka (2)	415052	84	519	3557	37	72	95
Łysomice (2)	415062	127	904	5460	37	72	95
Obrowo (2)	415072	162	668	6535	37	72	95
Wielka Nieszawka (2)	415082	216	1152	2642	37	72	95
Zławieś Wielka (2)	415092	178	889	7428	37	72	95
Cekcyn (2)	416012	253	693	4110	32	50	63
Gostycyn (2)	416022	136	449	3218	32	52	163
Kęsowo (2)	416032	109	270	2768	32	52	163
Lubiewo (2)	416042	163	652	3508	32	52	163
Śliwice (2)	416052	175	890	3335	32	52	93
Tuchola - miasto (4)	416064	18	3594	8997	32	52	163
Tuchola - obszar wiejski (5)	416065	222	294	3752	32	52	163
Wąbrzeźno (1)	417011	9	3280	8867	37	72	95
Dębowa Łąka (2)	417022	86	216	1909	37	72	95
Książki (2)	417032	86	208	2647	37	72	95
Pluźnica (2)	417042	120	385	3170	37	72	95
Wąbrzeźno (2)	417052	201	608	5284	37	72	95
Kowal (1)	418011	5	458	2220	34	76	99
Baruchowo (2)	418022	107	191	2217	34	76	99
Boniewo (2)	418032	78	134	2120	34	76	99
Brześć Kujawski - miasto (4)	418044	7	1113	2913	34	76	99
Brześć Kujawski - obszar wiejski (5)	418045	143	827	4029	34	76	99
Choceń (2)	418052	100	672	4982	34	76	99
Chodecz - miasto (4)	418064	1	271	1286	34	76	99
Chodecz - obszar wiejski (5)	418065	121	198	2673	34	76	99
Fabianki (2)	418072	76	904	5693	34	76	99
Izbica Kujawska - miasto (4)	418084	2	385	1772	33	76	99
Izbica Kujawska - obszar wiejski (5)	418085	130	116	3028	33	76	99
Kowal (2)	418092	115	216	2474	34	76	99
Lubanie (2)	418102	69	261	2882	34	76	99
Lubień Kujawski - miasto (4)	418114	2	261	858	34	76	99
Lubień Kujawski - obszar wiejski (5)	418115	149	336	3663	34	76	99
Lubraniec - miasto (4)	418124	2	454	2046	34	76	99
Lubraniec - obszar wiejski (5)	418125	146	173	4030	34	76	99
Włocławek (2)	418132	220	842	3930	36	76	99
Barcin - miasto (4)	419014	4	1145	5321	35	50	63
Barcin - obszar wiejski (5)	419015	117	1189	4393	35	50	63
Gąsawa (2)	419022	136	423	3198	35	50	63
Janowiec Wielkopolski - miasto (4)	419034	3	731	2724	35	50	63
Janowiec Wielkopolski - obszar wiejski (5)	419035	128	279	3298	35	50	63
Łabiszyn - miasto (4)	419044	3	674	3005	36	50	63
Łabiszyn - obszar wiejski (5)	419045	164	345	3100	36	50	63
Rogowo (2)	419052	179	552	4324	35	50	63
Żnin - miasto (4)	419064	8	3860	9059	35	50	63
Żnin - obszar wiejski (5)	419065	243	557	6392	35	50	63
Bydgoszcz (1)	461011	175	112322	237856	36	50	63
Grudziądz (1)	462011	58	20408	64410	36	50	138
Toruń (1)	463011	116	63754	139917	37	72	95

nazwa gminy	gmina				numer lokalnego rynku pracy		
	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Włocławek (1)	464011	84	33647	79314	34	76	99
Międzyrzec Podlaski (1)	601011	20	3895	10812	50	106	221
Terespol (1)	601021	10	1252	3735	38	106	174
Biała Podlaska (2)	601032	326	828	7273	38	106	174
Drelów (2)	601042	228	297	3129	50	106	221
Janów Podlaski (2)	601052	136	554	3269	38	106	174
Kodeń (2)	601062	151	291	2286	38	106	174
Konstantynów (2)	601072	87	316	2426	38	106	174
Leśna Podlaska (2)	601082	98	237	2618	38	106	174
Łomazy (2)	601092	199	260	3018	38	106	174
Międzyrzec Podlaski (2)	601102	262	779	5864	50	106	221
Piszczac (2)	601112	170	554	4365	38	106	174
Rokitno (2)	601122	141	167	1786	38	106	174
Rossosz (2)	601132	76	416	1369	38	106	174
Sławatycze (2)	601142	72	194	1439	38	106	174
Sosnówka (2)	601152	148	107	1489	38	5	5
Terespol (2)	601162	141	2292	4107	38	106	174
Tuczna (2)	601172	170	120	1849	38	106	174
Wisznice (2)	601182	173	385	2986	38	106	174
Zalesie (2)	601192	147	601	2534	38	106	174
Biłgoraj (1)	602011	21	8849	18105	39	46	50
Aleksandrów (2)	602022	54	127	1828	39	46	50
Biłgoraj (2)	602032	264	965	7685	39	46	50
Biszczka (2)	602042	106	173	2408	39	46	50
Frampol - miasto (4)	602054	5	298	898	39	46	50
Frampol - obszar wiejski (5)	602055	103	107	2835	39	46	50
Goraj (2)	602062	68	167	2527	39	46	50
Józefów - miasto (4)	602074	5	593	1532	40	15	198
Józefów - obszar wiejski (5)	602075	121	134	2806	40	15	198
Księżpol (2)	602082	141	619	4030	39	46	50
Łukowa (2)	602092	149	834	2625	39	46	50
Obsza (2)	602102	113	271	2633	39	46	50
Potok Górny (2)	602112	111	208	3195	139	46	50
Tarnogród - miasto (4)	602124	11	499	2196	39	46	50
Tarnogród - obszar wiejski (5)	602125	103	122	1969	39	46	50
Tereszpol (2)	602132	144	218	2360	39	46	50
Turobin (2)	602142	162	245	3782	39	5	5
Rejowiec Fabryczny (1)	603011	14	644	2864	45	5	5
Białopole (2)	603022	104	233	2014	45	5	5
Chełm (2)	603032	223	758	7900	56	5	5
Dorohusk (2)	603042	193	929	4154	45	5	5
Dubienka (2)	603052	96	159	1539	45	5	5
Kamień (2)	603062	97	140	2449	45	5	5
Leśniowice (2)	603072	118	183	2325	45	5	5
Rejowiec Fabryczny (2)	603082	88	218	2726	45	5	5
Ruda-Huta (2)	603092	112	158	2902	45	5	5
Sawin (2)	603102	190	251	3343	45	5	5
Siedliszcze (2)	603112	154	424	4167	45	5	5
Wierzbica (2)	603122	146	243	3237	45	5	5
Wojślawice (2)	603132	110	179	2441	45	5	5
Żmudź (2)	603142	136	147	1983	45	5	5
Rejowiec (2) *	603152	106	707	4092	45	5	5
Hrubieszów (1)	604011	33	4149	12056	56	89	188
Dołhobyczów (2)	604022	213	307	3668	56	5	5
Horodło (2)	604032	130	535	3209	56	89	188
Hrubieszów (2)	604042	259	399	6297	56	89	188
Mircze (2)	604052	235	470	4518	56	89	188
Trzeszczany (2)	604062	90	327	2725	56	89	188
Uchanie (2)	604072	121	294	2880	56	89	188

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Werbkowice (2)	604082	187	1103	5999	56	5	5
Batorz (2)	605012	71	257	1919	41	5	5
Chrzanów (2)	605022	72	90	1729	41	12	13
Dzwola (2)	605032	199	272	3815	41	12	13
Godziszów (2)	605042	102	318	3512	41	12	13
Janów Lubelski - miasto (4)	605054	15	3766	7931	41	12	13
Janów Lubelski - obszar wiejski (5)	605055	164	125	2470	41	12	13
Modliborzycze (2)	605062	153	655	4228	41	12	13
Potok Wielki (2)	605072	99	227	2785	41	28	29
Krasnystaw (1)	606011	42	5821	12656	42	81	109
Fajslawice (2)	606022	71	257	2987	56	5	5
Gorzków (2)	606032	96	104	2274	42	81	109
Izbica (2)	606042	138	378	5226	42	81	109
Krasnystaw (2)	606052	151	1193	5530	42	81	109
Kraśniczyn (2)	606062	110	234	2395	42	81	109
Łopiennik Górny (2)	606072	106	177	2369	42	81	109
Rudnik (2)	606092	89	130	1861	42	81	109
Siennica Różana (2)	606102	98	357	2534	42	81	109
Żółkiewka (2)	606112	130	378	3337	42	5	5
Kraśnik (1)	607011	25	10136	23219	43	19	20
Annopol - miasto (4) *	607024	8	511	1794	43	19	20
Annopol - obszar wiejski (5) *	607025	143	197	3749	43	19	20
Dzierzkowice (2)	607032	87	192	3269	43	19	20
Gościeradów (2)	607042	160	478	4345	43	19	20
Kraśnik (2)	607052	105	324	4266	43	19	20
Szastarka (2)	607062	73	221	3536	43	19	20
Trzydnik Duży (2)	607072	104	338	4113	43	19	20
Urzędów (2)	607082	119	578	5256	43	19	20
Wilkołaz (2)	607092	82	218	3337	43	5	5
Zakrzówek (2)	607102	99	366	3984	43	19	20
Lubartów (1)	608011	14	5755	15872	44	5	5
Abramów (2)	608022	85	184	2455	48	5	5
Firlej (2)	608032	126	323	3599	44	5	5
Jeziorzany (2)	608042	66	250	1660	44	5	5
Kamionka (2)	608052	112	449	3752	44	5	5
Kock - miasto (4)	608064	17	384	2126	44	5	212
Kock - obszar wiejski (5)	608065	84	81	1869	44	5	212
Lubartów (2)	608072	158	1020	6213	44	5	5
Michów (2)	608082	136	536	3609	44	5	5
Niedźwiada (2)	608092	96	251	3718	44	5	5
Ostrów Lubelski - miasto (4)	608104	30	452	1415	44	5	5
Ostrów Lubelski - obszar wiejski (5)	608105	92	80	1884	44	5	5
Ostrówek (2)	608112	90	181	2358	44	5	5
Serniki (2)	608122	76	133	2810	44	5	5
Uścimów (2)	608132	107	174	1994	45	5	5
Bełżyce - miasto (4)	609014	23	1265	4611	56	5	5
Bełżyce - obszar wiejski (5)	609015	111	389	4031	56	5	5
Borzechów (2)	609022	67	135	2182	56	5	5
Bychawa - miasto (4)	609034	7	1119	3283	56	5	5
Bychawa - obszar wiejski (5)	609035	140	215	3994	56	5	5
Garbów (2)	609042	103	743	5414	56	5	5
Głusk (2)	609052	64	538	4794	56	5	5
Jabłonna (2)	609062	131	307	4492	56	5	5
Jastków (2)	609072	113	923	7697	56	5	5
Konopnica (2)	609082	93	853	6979	56	5	5
Krzczonów (2)	609092	128	263	2775	56	5	5
Niedzwica Duża (2)	609102	107	1057	6776	56	5	5
Niemce (2)	609112	141	2327	10536	56	5	5
Strzyżewice (2)	609122	109	790	4703	56	5	5

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Wojciechów (2)	609132	81	210	3499	49	5	5
Wólka (2)	609142	73	495	5864	56	5	5
Wysokie (2)	609152	114	225	2834	56	5	5
Zakrzew (2)	609162	75	113	1786	56	5	5
Cyców (2)	610012	148	430	4490	45	5	5
Ludwin (2)	610022	122	429	2976	45	5	5
Łęczna - miasto (4)	610034	19	2746	15641	45	5	5
Łęczna - obszar wiejski (5)	610035	56	127	2200	45	5	5
Milejów (2)	610042	117	863	5710	45	5	5
Puchaczów (2)	610052	92	4584	2972	45	5	5
Spiczyn (2)	610062	83	316	3325	56	5	5
Łuków (1)	611011	36	8444	19992	46	70	147
Stoczek Łukowski (1) *	611021	9	632	1728	46	6	6
Adamów (2)	611032	99	423	3290	46	70	147
Krzywdą (2)	611042	161	1224	6016	46	35	37
Łuków (2)	611052	307	841	9831	46	70	147
Serokomla (2)	611062	77	196	2300	46	70	147
Stanin (2)	611072	161	375	5482	46	70	147
Stoczek Łukowski (2) *	611082	173	468	4738	46	6	6
Trzebieszów (2)	611092	141	763	4221	46	70	147
Wojcieszków (2)	611102	109	241	4012	46	70	147
Wola Mysłowska (2)	611112	121	203	2892	97	6	6
Chodel (2)	612012	108	412	3984	47	5	5
Józefów nad Wisłą (2) *	612022	142	278	4086	47	19	20
Karczmiska (2)	612032	95	446	3739	47	57	182
Łaziska (2)	612042	109	204	3044	47	57	182
Opole Lubelskie - miasto (4)	612054	15	2226	5742	47	57	182
Opole Lubelskie - obszar wiejski (5)	612055	178	338	5288	47	57	182
Poniatowa - miasto (4)	612064	15	1580	6608	47	57	182
Poniatowa - obszar wiejski (5)	612065	70	124	3142	47	57	182
Wilków (2)	612072	78	253	2807	48	57	182
Dębowa Kłoda (2)	613012	188	180	2333	56	35	233
Jabłoń (2)	613022	110	313	2375	56	35	233
Milanów (2)	613032	117	207	2389	56	35	233
Parczew - miasto (4)	613044	8	2452	6619	56	35	233
Parczew - obszar wiejski (5)	613045	139	647	2680	56	35	233
Podewórze (2)	613052	107	82	1044	56	5	5
Siemień (2)	613062	111	264	2816	56	35	233
Sosnowica (2)	613072	172	233	1605	45	5	5
Puławy (1)	614011	50	16413	31366	48	57	73
Baranów (2)	614022	85	370	2393	48	57	73
Janowiec (2)	614032	79	276	2221	48	57	73
Kazimierz Dolny - miasto (4)	614044	30	581	2088	48	57	73
Kazimierz Dolny - obszar wiejski (5)	614045	42	61	2079	48	57	73
Końskowola (2)	614052	90	908	5462	48	57	73
Kurów (2)	614062	101	603	4889	48	57	73
Markuszów (2)	614072	40	151	1807	48	57	73
Nałęczów - miasto (4)	614084	14	1379	2614	49	57	220
Nałęczów - obszar wiejski (5)	614085	49	231	3151	49	57	220
Puławy (2)	614092	162	609	6729	48	57	73
Wąwolnica (2)	614102	62	591	2959	49	57	220
Żyrzyn (2)	614112	129	377	3870	48	57	73
Radzyń Podlaski (1)	615011	19	4964	10634	50	106	212
Borki (2)	615022	112	281	3605	50	106	212
Czemierniki (2)	615032	107	162	2711	50	106	212
Kąkolewnica (2) *	615042	148	432	4775	50	106	221
Komarówka Podlaska (2)	615052	138	224	2607	50	106	212
Radzyń Podlaski (2)	615062	155	639	4749	50	106	212
Ulan-Majorat (2)	615072	108	280	3389	46	70	147

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Wohyń (2)	615082	178	540	4250	50	106	212
Dęblin (1)	616011	38	3404	11804	51	6	6
Kłoczew (2)	616022	143	389	4096	52	5	195
Nowodwór (2)	616032	72	161	2567	52	5	195
Ryki - miasto (4)	616044	27	3483	6423	52	5	195
Ryki - obszar wiejski (5)	616045	135	436	6342	52	5	195
Stężycza (2)	616052	116	287	3200	51	5	5
Ułęż (2)	616062	84	213	2094	52	5	195
Świdnik (1)	617011	20	8654	26714	56	5	5
Mełgiew (2)	617022	95	841	5216	56	5	5
Piaski - miasto (4)	617034	8	547	1664	56	5	5
Piaski - obszar wiejski (5)	617035	162	432	4769	56	5	5
Rybczewice (2)	617042	99	296	2205	56	5	5
Trawniki (2)	617052	84	710	5624	56	5	5
Tomaszów Lubelski (1)	618011	13	5925	13154	53	51	64
Bełzec (2)	618022	29	576	1897	53	51	64
Jarczów (2)	618032	106	175	1994	53	51	64
Krynice (2)	618042	74	146	2064	53	51	64
Lubycza Królewska (2)	618052	212	937	4342	53	51	64
Łaszczów (2) *	618062	129	909	3835	54	51	64
Rachanie (2)	618072	89	386	3371	53	51	64
Susiec (2)	618082	191	382	4580	53	51	64
Tarnawatka (2)	618092	83	301	2361	53	51	64
Telatyn (2)	618102	112	165	2575	54	51	64
Tomaszów Lubelski (2)	618112	176	492	6488	53	51	64
Tyszowce - miasto (4) *	618124	18	328	1245	54	51	64
Tyszowce - obszar wiejski (5) *	618125	110	77	2299	54	51	64
Ułhówek (2)	618132	147	271	3189	53	51	64
Włodawa (1)	619011	18	2919	9206	56	106	122
Hanna (2)	619022	139	178	1916	56	106	122
Hańsk (2)	619032	176	634	2405	95	35	37
Stary Brus (2)	619042	132	97	1324	56	106	122
Urszulin (2)	619052	172	251	2348	45	5	5
Włodawa (2)	619062	245	506	3582	56	106	122
Wola Uhruska (2)	619072	154	272	2510	45	5	5
Wyryki (2)	619082	220	156	1636	56	106	122
Adamów (2)	620012	111	667	2885	56	89	126
Grabowiec (2)	620022	128	201	2606	56	89	126
Komarów-Osada (2)	620032	124	181	3125	56	89	126
Krasnobród - miasto (4)	620044	7	483	1868	40	89	126
Krasnobród - obszar wiejski (5)	620045	120	343	2358	40	89	126
Łabunie (2)	620052	87	318	3721	56	89	126
Miączyn (2)	620062	156	224	3525	56	89	126
Nielisz (2)	620072	112	227	3405	56	89	126
Radecznicza (2)	620082	110	459	3656	39	46	50
Sitno (2)	620092	112	347	4155	56	89	126
Skierbieszów (2)	620102	139	210	3053	56	89	126
Stary Zamość (2)	620112	98	201	3127	56	89	126
Sułów (2)	620122	93	271	2853	55	89	126
Szczebrzeszyn - miasto (4)	620134	29	716	3298	55	89	126
Szczebrzeszyn - obszar wiejski (5)	620135	95	339	3876	55	89	126
Zamość (2)	620142	196	1401	12647	56	89	126
Zwierzyniec - miasto (4)	620154	5	1755	2052	55	89	126
Zwierzyniec - obszar wiejski (5)	620155	149	71	2289	55	89	126
Biała Podlaska (1)	661011	49	13779	39174	38	106	174
Chełm (1)	662011	35	14687	45415	45	5	5
Lublin (1)	663011	147	105724	236827	56	5	5
Zamość (1)	664011	30	17631	44222	56	89	126
Kostrzyn nad Odrą (1) *	801011	46	4793	11618	57	4	4

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Bogdaniec (2)	801022	112	391	4390	61	114	189
Deszczno (2)	801032	168	730	4955	61	114	189
Kłodawa (2)	801042	235	366	4330	61	114	189
Lubiszyn (2)	801052	205	587	4265	61	114	189
Santok (2)	801062	169	879	4995	61	114	189
Witnica - miasto (4)	801074	8	1814	4383	57	114	189
Witnica - obszar wiejski (5)	801075	271	408	3799	57	114	189
Gubin (1)	802011	21	2329	11164	63	41	45
Bobrowice (2)	802022	185	431	2026	63	41	45
Bytnica (2)	802032	209	206	1708	63	41	45
Dąbie (2)	802042	171	476	3367	63	41	45
Gubin (2)	802052	379	713	4532	63	41	45
Krosno Odrzańskie - miasto (4)	802064	8	2765	8343	63	41	45
Krosno Odrzańskie - obszar wiejski (5)	802065	204	246	4172	63	41	45
Maszewo (2)	802072	214	115	1843	63	41	45
Bledzew (2)	803012	247	236	2995	208	123	213
Międzyrzecz - miasto (4)	803024	10	4899	12301	208	123	213
Międzyrzecz - obszar wiejski (5)	803025	305	297	3983	208	123	213
Przytoczna (2)	803032	185	602	3852	208	123	213
Pszczew (2)	803042	178	435	2711	208	123	213
Skwierzyna - miasto (4)	803054	36	2423	6737	208	123	213
Skwierzyna - obszar wiejski (5)	803055	249	158	1723	208	123	213
Trzciel - miasto (4)	803064	3	420	1591	209	123	213
Trzciel - obszar wiejski (5)	803065	175	776	2522	209	123	213
Nowa Sól (1)	804011	22	9465	26422	9	41	136
Bytom Odrzański - miasto (4)	804024	2	490	2909	9	41	136
Bytom Odrzański - obszar wiejski (5)	804025	50	4	599	9	41	136
Kolsko (2)	804032	81	155	2068	67	41	136
Koźuchów - miasto (4)	804044	6	2198	6194	9	41	136
Koźuchów - obszar wiejski (5)	804045	173	91	4009	9	41	136
Nowa Sól (2)	804052	176	724	4274	9	41	136
Nowe Miasteczko - miasto (4)	804064	3	385	1810	9	41	136
Nowe Miasteczko - obszar wiejski (5)	804065	74	125	1666	9	41	136
Otyń (2)	804072	92	625	3998	9	41	136
Siedlisko (2)	804082	92	306	2219	9	41	136
Cybinka - miasto (4)	805014	5	731	1727	60	62	227
Cybinka - obszar wiejski (5)	805015	274	268	2623	60	62	227
Górzycza (2)	805022	145	282	2588	60	62	227
Ośno Lubuskie - miasto (4)	805034	8	876	2448	58	62	227
Ośno Lubuskie - obszar wiejski (5)	805035	190	219	1630	58	62	227
Rzepin - miasto (4)	805044	11	2373	4235	59	62	227
Rzepin - obszar wiejski (5)	805045	180	269	2150	59	62	227
Słubice - miasto (4)	805054	19	4153	11671	60	62	227
Słubice - obszar wiejski (5)	805055	167	520	1645	60	62	227
Dobiegów - miasto (4)	806014	6	616	2085	61	114	189
Dobiegów - obszar wiejski (5)	806015	345	192	2349	61	114	189
Drezdenko - miasto (4)	806024	11	2403	6752	61	114	189
Drezdenko - obszar wiejski (5)	806025	389	829	4237	61	114	189
Stare Kurowo (2)	806032	78	520	2611	61	114	189
Strzelce Krajeńskie - miasto (4)	806044	5	3112	6907	61	114	189
Strzelce Krajeńskie - obszar wiejski (5)	806045	314	197	4639	61	114	189
Zwierzyn (2)	806052	100	412	2797	61	114	189
Krzyszczewo (2)	807012	194	439	2803	62	114	189
Lubniewice - miasto (4)	807024	12	234	1299	62	62	81
Lubniewice - obszar wiejski (5)	807025	118	27	741	62	62	81
Słońsk (2)	807032	159	922	2985	57	4	4
Sulęcín - miasto (4)	807044	9	4137	6579	62	62	81
Sulęcín - obszar wiejski (5)	807045	311	832	3776	62	62	81
Torzyn - miasto (4)	807054	9	813	1629	62	62	81

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Torzym - obszar wiejski (5)	807055	366	480	2702	62	62	81
Lubrza (2)	808012	122	372	2119	63	56	71
Łagów (2)	808022	200	466	3350	63	56	71
Skąpe (2)	808032	181	966	3370	63	56	71
Szczaniec (2)	808042	113	238	2519	63	56	71
Świebodzin - miasto (4)	808054	11	8637	14401	63	56	71
Świebodzin - obszar wiejski (5)	808055	216	883	5023	63	56	71
Zbąszynek - miasto (4)	808064	3	1212	3353	209	100	206
Zbąszynek - obszar wiejski (5)	808065	91	1943	1994	209	100	206
Babimost - miasto (4)	809014	4	1290	2713	205	100	206
Babimost - obszar wiejski (5)	809015	89	53	1476	205	100	206
Bojadła (2)	809022	102	490	2117	63	58	58
Czerwieńsk - miasto (4)	809034	9	1348	2859	63	41	45
Czerwieńsk - obszar wiejski (5)	809035	186	288	3469	63	41	45
Kargowa - miasto (4)	809044	5	936	2370	205	100	159
Kargowa - obszar wiejski (5)	809045	124	178	1311	205	100	159
Nowogród Bobrzański - miasto (4)	809054	15	443	3525	63	41	45
Nowogród Bobrzański - obszar wiejski (5)	809055	244	405	2587	63	41	45
Sulechów - miasto (4)	809064	7	3459	12080	63	41	45
Sulechów - obszar wiejski (5)	809065	229	1485	5515	63	41	45
Świdnica (2)	809072	161	646	3842	63	41	45
Trzebiechów (2)	809082	81	418	2081	63	41	45
Zabór (2)	809092	93	291	2265	63	41	45
Zielona Góra (2)	809102	220	2209	11076	63	41	45
Gozdnica (1)	810011	24	250	2180	64	24	25
Żagań (1)	810021	40	4935	17786	66	24	25
Brzeźnica (2)	810032	122	169	2374	66	24	25
Iłowa - miasto (4)	810044	9	1004	2529	64	24	25
Iłowa - obszar wiejski (5)	810045	144	104	2019	64	24	25
Małomice - miasto (4)	810054	5	439	2363	66	24	207
Małomice - obszar wiejski (5)	810055	75	81	1171	66	24	207
Niegosławice (2)	810062	135	202	2886	66	24	207
Szprotawa - miasto (4)	810074	11	2272	8304	66	24	207
Szprotawa - obszar wiejski (5)	810075	223	844	5683	66	24	207
Wymiarki (2)	810082	63	364	1582	64	24	25
Żagań (2)	810092	281	336	4434	66	24	25
Łęknica (1)	811011	16	472	1698	65	24	25
Żary (1)	811021	33	12167	25779	66	24	25
Brody (2)	811032	241	241	2229	63	41	45
Jasień - miasto (4)	811044	3	585	2935	66	24	45
Jasień - obszar wiejski (5)	811045	124	94	1759	66	24	45
Lipinki Łużyckie (2)	811052	89	289	2044	66	24	25
Lubsko - miasto (4)	811064	13	2686	9681	66	24	45
Lubsko - obszar wiejski (5)	811065	170	224	2945	66	24	45
Przewóz (2)	811072	178	262	2031	66	24	25
Trzebień (2)	811082	166	491	3651	65	24	25
Tuplice (2)	811092	66	323	2106	65	24	25
Żary (2)	811102	294	430	7481	66	24	25
Sława - miasto (4) *	812014	14	1654	2575	67	58	208
Sława - obszar wiejski (5) *	812015	313	1014	5241	67	58	208
Szlichtyngowa - miasto (4) *	812024	2	421	857	67	58	208
Szlichtyngowa - obszar wiejski (5) *	812025	98	121	2317	67	58	208
Wschowa - miasto (4) *	812034	8	2848	9515	67	58	208
Wschowa - obszar wiejski (5) *	812035	189	718	4444	67	58	208
Gorzów Wielkopolski (1)	861011	86	36521	84890	61	114	189
Zielona Góra (1)	862011	58	38087	79512	63	41	45
Bełchatów (1)	1001011	35	11392	44716	68	48	61
Bełchatów (2)	1001022	180	888	5804	68	48	61

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Drużbice (2)	1001032	113	311	3034	68	48	61
Kleszczów (2)	1001042	125	17477	2594	68	48	61
Kluki (2)	1001052	119	461	2355	68	48	61
Rusiec (2)	1001062	100	839	3211	68	48	61
Szczerców (2)	1001072	129	1451	4648	68	48	61
Zelów - miasto (4)	1001084	11	1257	5189	68	48	61
Zelów - obszar wiejski (5)	1001085	156	303	4292	68	48	61
Kutno (1)	1002011	34	14825	31653	78	38	42
Bedlno (2)	1002022	125	354	3638	78	38	42
Dąbrowice (2)	1002032	46	109	1184	78	38	42
Krośnice - miasto (4)	1002044	3	719	2969	78	38	42
Krośnice - obszar wiejski (5)	1002045	92	323	2656	78	38	42
Krzyżanów (2)	1002052	103	448	2670	78	38	42
Kutno (2)	1002062	122	447	5197	78	38	42
Łanięta (2)	1002072	55	229	1625	78	38	42
Nowe Ostrowy (2)	1002082	72	171	2392	78	38	42
Oporów (2)	1002092	68	134	1661	78	38	42
Strzelce (2)	1002102	90	400	2512	78	38	42
Żychlin - miasto (4)	1002114	9	1979	5745	78	38	42
Żychlin - obszar wiejski (5)	1002115	68	279	2524	78	38	42
Buczek (2)	1003012	92	303	3025	69	84	141
Łask - miasto (4)	1003024	15	5166	12507	69	84	141
Łask - obszar wiejski (5)	1003025	130	650	6081	69	84	141
Sędziejowice (2)	1003032	120	858	3921	69	84	141
Widawa (2)	1003042	179	829	4608	69	88	125
Wodzierady (2)	1003052	82	268	1862	69	66	87
Łęczycza (1)	1004011	9	3600	9993	78	107	177
Daszyna (2)	1004022	81	191	2532	78	107	177
Góra Świętej Małgorzaty (2)	1004032	90	208	2612	78	107	177
Grabów (2)	1004042	156	347	3872	206	107	177
Łęczycza (2)	1004052	151	844	5058	78	107	177
Piątek (2)	1004062	133	626	3946	78	66	87
Świnice Warckie (2)	1004072	93	252	2438	78	107	177
Witonia (2)	1004082	60	173	2121	78	107	177
Łowicz (1)	1005011	23	8390	20073	117	66	87
Bielawy (2)	1005022	164	515	3581	78	66	87
Chąśno (2)	1005032	72	142	1787	117	66	87
Domaniewice (2)	1005042	86	215	2844	78	66	87
Kiernozia (2)	1005052	76	146	2140	117	6	6
Kocierzew Południowy (2)	1005062	94	222	2697	117	66	87
Łowicz (2)	1005072	133	313	4320	117	66	87
Łyszkowice (2)	1005082	107	600	4094	78	66	87
Nieborów (2)	1005092	104	603	5897	117	6	6
Zduny (2)	1005102	129	359	3526	117	66	87
Andrespol (2)	1006022	23	1200	7613	78	66	87
Brójce (2)	1006032	69	546	3453	78	66	87
Koluszki - miasto (4)	1006074	10	2864	8952	78	66	87
Koluszki - obszar wiejski (5)	1006075	148	795	5988	78	66	87
Nowosolna (2)	1006082	54	631	2478	78	66	87
Rzgów - miasto (4) *	1006104	16	1345	2204	70	66	87
Rzgów - obszar wiejski (5) *	1006105	50	576	3706	70	66	87
Tuszyn - miasto (4)	1006114	23	1855	4630	70	66	87
Tuszyn - obszar wiejski (5)	1006115	107	554	2749	70	66	87
Białaczów (2)	1007012	115	792	3326	72	25	113
Drzewica - miasto (4)	1007024	5	927	2637	71	25	113
Drzewica - obszar wiejski (5)	1007025	113	344	4108	71	25	113
Mniszków (2)	1007032	124	337	2795	79	48	12
Opoczno - miasto (4)	1007044	25	6940	15088	72	25	113
Opoczno - obszar wiejski (5)	1007045	166	946	7569	72	25	113

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Paradyż (2)	1007052	81	418	2566	72	25	113
Poświętne (2)	1007062	141	170	1896	72	66	87
Sławno (2)	1007072	129	660	4278	72	25	113
Żarnów (2)	1007082	141	357	3528	185	25	113
Konstantynów Łódzki (1)	1008011	27	3635	11556	78	66	87
Pabianice (1)	1008021	33	15615	45467	78	66	87
Dłutów (2)	1008032	101	254	2577	78	66	87
Dobroń (2)	1008042	95	755	4446	69	66	87
Ksawerów (2) *	1008052	14	1575	4674	78	66	87
Lutomiersk (2)	1008062	134	854	4396	78	66	87
Pabianice (2)	1008072	88	478	3592	78	66	87
Działoszyn - miasto (4)	1009014	5	1200	4015	76	48	234
Działoszyn - obszar wiejski (5)	1009015	116	1443	3982	76	48	234
Kielczygłów (2)	1009022	89	265	2558	68	48	61
Nowa Brzeźnica (2)	1009032	135	279	3004	68	39	43
Pajęczno - miasto (4)	1009044	20	1538	4431	68	48	234
Pajęczno - obszar wiejski (5)	1009045	94	365	2989	68	48	234
Rząśnia (2)	1009052	86	312	2726	68	48	61
Siemkowice (2)	1009062	98	248	2983	76	48	234
Strzelce Wielkie (2)	1009072	78	314	2954	68	48	61
Sulmierzyce (2)	1009082	83	336	2866	68	48	61
Aleksandrów (2)	1010012	144	142	2581	79	48	12
Czarnocin (2)	1010022	72	420	2455	70	66	87
Gorzkowice (2)	1010032	102	762	5251	79	48	61
Grabica (2)	1010042	128	295	3669	79	48	12
Łęki Szlacheckie (2)	1010052	109	229	2050	79	48	12
Moszczenica (2)	1010062	111	1741	7852	79	48	12
Ręczno (2)	1010072	89	155	2145	79	48	12
Rozprza (2)	1010082	163	818	7343	79	48	12
Sulejów - miasto (4)	1010094	26	926	3989	79	48	12
Sulejów - obszar wiejski (5)	1010095	162	301	5615	79	48	12
Wola Krzysztoporska (2)	1010102	171	1041	6987	79	48	12
Wolbórz (2) *	1010112	152	2039	4683	79	48	12
Dalików (2)	1011012	115	398	2211	78	66	87
Pęczniew (2)	1011022	128	297	2137	78	122	101
Poddebice - miasto (4)	1011034	6	2011	5199	78	122	101
Poddebice - obszar wiejski (5)	1011035	218	241	4941	78	122	101
Uniejów - miasto (4)	1011044	12	279	1838	217	122	101
Uniejów - obszar wiejski (5)	1011045	117	54	2570	217	122	101
Wartkowice (2)	1011052	140	982	3764	78	122	101
Zadzim (2)	1011062	145	346	3121	78	66	87
Radomsko (1)	1012011	51	12647	31336	68	95	142
Dobryczyce (2)	1012022	51	474	2549	68	95	142
Gidle (2)	1012032	116	350	3963	68	95	142
Gomunice (2)	1012042	62	524	3689	68	48	61
Kamieńsk - miasto (4)	1012054	12	458	1812	68	48	61
Kamieńsk - obszar wiejski (5)	1012055	84	238	1896	68	48	61
Kobiele Wielkie (2)	1012062	102	235	2674	68	95	142
Kodrąb (2)	1012072	106	245	2797	68	95	142
Lgota Wielka (2)	1012082	63	257	2686	68	48	61
Ładzice (2)	1012092	83	879	2996	68	95	142
Masłowice (2)	1012102	116	198	2544	68	95	142
Przedbórz - miasto (4)	1012114	6	571	2452	68	95	142
Przedbórz - obszar wiejski (5)	1012115	184	127	2114	68	95	142
Radomsko (2)	1012122	86	580	3475	68	95	142
Wielgomłynny (2)	1012132	123	300	2929	68	95	142
Żytno (2)	1012142	197	471	3292	68	39	43
Rawa Mazowiecka (1)	1013011	14	5068	11921	73	40	44
Biała Rawska - miasto (4)	1013024	10	624	2092	73	40	44

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Biała Rawska - obszar wiejski (5)	1013025	198	174	4849	73	40	44
Cielądz (2)	1013032	93	112	2407	73	40	44
Rawa Mazowiecka (2)	1013042	164	798	5131	73	40	44
Regnów (2)	1013052	46	74	1027	73	40	44
Sadkowie (2)	1013062	121	167	3302	73	6	6
Sieradz (1)	1014011	51	12489	29082	74	88	125
Błaszki - miasto (4)	1014024	2	589	1360	219	65	86
Błaszki - obszar wiejski (5)	1014025	199	728	7715	219	65	86
Brąszewice (2)	1014032	107	228	2525	74	88	125
Brzeźnio (2)	1014042	129	641	3751	74	88	125
Burzenin (2)	1014052	120	260	3431	74	88	125
Goszczanów (2)	1014062	122	214	3315	75	65	86
Klonowa (2)	1014072	95	167	1729	74	88	125
Sieradz (2)	1014082	182	764	5976	74	88	125
Warta - miasto (4)	1014094	11	903	2153	75	88	125
Warta - obszar wiejski (5)	1014095	244	409	5561	75	88	125
Wróblew (2)	1014102	113	687	3699	74	88	125
Złoczew - miasto (4)	1014114	14	620	2193	74	88	125
Złoczew - obszar wiejski (5)	1014115	103	157	2282	74	88	125
Bolimów (2)	1015012	112	280	2459	78	6	6
Głuchów (2)	1015022	111	419	3423	73	66	87
Godzianów (2)	1015032	44	153	1557	78	66	87
Kowiesy (2)	1015042	86	88	1780	73	6	6
Lipce Reymontowskie (2)	1015052	43	241	2002	78	66	87
Maków (2)	1015062	83	311	3517	78	66	87
Nowy Kawęczyn (2)	1015072	104	266	1890	78	66	87
Skierniewice (2)	1015082	131	704	3995	78	66	87
Słupia (2)	1015092	41	110	1537	78	66	87
Tomaszów Mazowiecki (1)	1016011	41	15242	43060	72	66	165
Będków (2)	1016022	58	149	2073	70	66	87
Budziszewice (2)	1016032	30	103	1334	72	66	87
Czerniewice (2)	1016042	128	370	3150	73	40	44
Inowłódz (2)	1016052	98	678	2330	72	66	165
Lubochnia (2)	1016062	131	519	4711	72	66	165
Rokiciny (2)	1016072	91	657	3610	78	66	87
Rzeczycza (2)	1016082	109	221	2858	73	40	44
Tomaszów Mazowiecki (2)	1016092	150	935	6214	72	66	165
Ujazd (2)	1016102	97	1341	5026	72	66	165
Żelechlinek (2)	1016112	92	200	1980	73	66	87
Biała (2)	1017012	74	262	3310	76	18	19
Czarnożyły (2)	1017022	70	589	2691	76	18	19
Konopnica (2)	1017032	83	232	2375	76	18	19
Mokrsko (2)	1017042	78	445	3239	76	18	19
Osjaków (2)	1017052	101	435	2815	76	18	19
Ostrówek (2)	1017062	101	269	2813	76	18	19
Państwów (2)	1017072	115	411	3770	76	18	19
Skomlin (2)	1017082	55	172	2076	76	18	19
Wieluń - miasto (4)	1017094	17	10491	16088	76	18	19
Wieluń - obszar wiejski (5)	1017095	113	503	5226	76	18	19
Wierzchnas (2)	1017102	119	220	4004	76	18	19
Bolesławiec (2)	1018012	65	348	2555	77	112	185
Czastary (2)	1018022	63	348	2450	77	112	185
Galewice (2)	1018032	135	697	3777	77	112	185
Lututów (2)	1018042	76	532	2831	76	18	19
Łubnice (2)	1018052	61	408	2469	77	126	225
Sokolniki (2)	1018062	80	2084	3002	77	112	185
Wieruszów - miasto (4)	1018074	6	3560	5874	77	112	185
Wieruszów - obszar wiejski (5)	1018075	91	857	3277	77	112	185
Zduńska Wola (1)	1019011	25	12056	28970	74	84	115

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	pojęcie taksonomiczne - słabe kryterium grupowania	pojęcie taksonomiczne - mocne kryterium grupowania
Szadek - miasto (4)	1019024	18	658	1266	74	84	115
Szadek - obszar wiejski (5)	1019025	133	269	3215	74	84	115
Zapolice (2)	1019032	81	301	2939	74	84	115
Zduńska Wola (2)	1019042	112	1816	6899	74	84	115
Głowno (1)	1020011	20	2477	9667	78	66	87
Ozorków (1)	1020021	15	3597	13431	78	66	87
Zgierz (1)	1020031	42	11265	38191	78	66	87
Aleksandrów Łódzki - miasto (4)	1020044	14	2624	13554	78	66	87
Aleksandrów Łódzki - obszar wiejski (5)	1020045	102	904	3894	78	66	87
Głowno (2)	1020052	105	485	2908	78	66	87
Ozorków (2)	1020062	96	1006	3988	78	66	87
Parzęczew (2)	1020072	104	238	3209	78	66	87
Stryków - miasto (4)	1020084	8	1619	2375	78	66	87
Stryków - obszar wiejski (5)	1020085	150	2025	5408	78	66	87
Zgierz (2)	1020092	199	1345	7234	78	66	87
Brzeziny (1) *	1021011	22	2211	8282	78	66	87
Brzeziny (2) *	1021022	107	362	3226	78	66	87
Dmosin (2) *	1021032	100	336	2862	78	66	87
Jeżów (2) *	1021042	64	431	2177	78	66	87
Rogów (2) *	1021052	66	332	2913	78	66	87
Łódź (1)	1061011	293	208108	499477	78	66	87
Piotrków Trybunalski (1)	1062011	67	23814	52023	79	48	12
Skierzwice (1)	1063011	33	13042	32776	78	66	87
Bochnia (1)	1201011	30	8597	19180	80	13	150
Bochnia (2)	1201022	132	998	11406	80	13	150
Drwinia (2)	1201032	108	318	3922	80	13	14
Lipnica Murowana (2)	1201042	60	300	3248	80	13	150
Łapanów (2)	1201052	72	653	4426	80	13	14
Nowy Wiśnicz - miasto (4)	1201064	5	787	1708	80	13	150
Nowy Wiśnicz - obszar wiejski (5)	1201065	77	554	5954	80	13	150
Rzezawa (2)	1201072	86	599	6336	80	13	150
Trzciana (2)	1201082	44	446	2990	80	13	14
Żegocina (2)	1201092	35	456	3007	80	13	150
Borzęcin (2)	1202012	103	458	5022	81	13	94
Brzesko - miasto (4)	1202024	12	6591	11295	81	13	94
Brzesko - obszar wiejski (5)	1202025	91	567	11364	81	13	94
Czchów - miasto (4) *	1202034	14	355	1339	81	13	94
Czchów - obszar wiejski (5) *	1202035	52	351	4158	81	13	94
Dębno (2)	1202042	82	1183	8536	81	35	37
Gnojnik (2)	1202052	55	490	4332	81	13	94
Iwkowa (2)	1202062	47	292	3513	81	13	14
Szczurowa (2)	1202072	135	697	5923	81	13	94
Alwernia - miasto (4)	1203014	9	857	2311	95	13	14
Alwernia - obszar wiejski (5)	1203015	67	260	5741	95	13	14
Babice (2)	1203022	54	489	5544	95	110	179
Chrzanów - miasto (4)	1203034	38	9852	26359	95	110	179
Chrzanów - obszar wiejski (5)	1203035	41	400	6319	95	110	179
Libiąż - miasto (4)	1203044	36	2403	11839	94	110	179
Libiąż - obszar wiejski (5)	1203045	21	212	3278	94	110	179
Trzebinia - miasto (4)	1203054	31	6411	11936	95	110	179
Trzebinia - obszar wiejski (5)	1203055	75	662	9736	95	110	179
Bolesław (2)	1204012	36	131	1726	82	35	37
Dąbrowa Tarnowska - miasto (4)	1204024	23	2896	7527	82	35	37
Dąbrowa Tarnowska - obszar wiejski (5)	1204025	93	158	5336	82	35	37
Gręboszów (2)	1204032	49	123	2117	82	35	37
Mędrzechów (2)	1204042	44	173	2112	82	35	37
Olesno (2)	1204052	78	377	4709	82	35	37
Radgoszcz (2)	1204062	88	348	4252	82	35	37
Szczucin (2) *	1204072	119	1163	8248	82	35	37

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Gorlice (1)	1205011	24	10310	18982	83	20	21
Biecz - miasto (4)	1205024	18	906	2964	83	20	21
Biecz - obszar wiejski (5)	1205025	80	563	7244	83	20	21
Bobowa (2) *	1205032	50	491	5292	83	20	21
Gorlice (2)	1205042	103	886	9677	83	20	21
Lipinki (2)	1205052	66	446	4069	83	20	21
Łużna (2)	1205062	56	359	4867	83	20	21
Moszczenica (2)	1205072	38	156	2767	83	20	21
Ropa (2)	1205082	49	296	2941	83	20	21
Sękowa (2)	1205092	195	263	2864	83	20	21
Uście Gorlickie (2)	1205102	288	1095	3797	83	20	21
Czernichów (2)	1206012	85	678	8075	95	13	14
Igołomia-Wawrzeńczyce (2)	1206022	63	327	4795	95	13	14
Iwanowice (2)	1206032	71	368	5146	95	13	14
Jerzmanowice-Przegonia (2)	1206042	68	361	6453	89	13	14
Kocmyrzów-Luborzyca (2)	1206052	81	470	8333	95	13	14
Krzyszowice - miasto (4)	1206064	5	3047	6351	95	110	179
Krzyszowice - obszar wiejski (5)	1206065	134	1159	13527	95	110	179
Liszki (2)	1206072	72	979	10051	95	13	14
Michałowice (2)	1206082	51	439	4924	95	13	14
Mogilany (2)	1206092	44	1194	7194	95	13	14
Skała - miasto (4)	1206104	3	510	2344	95	13	14
Skała - obszar wiejski (5)	1206105	72	290	3615	95	13	14
Skawina - miasto (4)	1206114	21	8814	15201	84	13	14
Skawina - obszar wiejski (5)	1206115	79	1315	11217	84	13	14
Słomniki - miasto (4)	1206124	3	580	2786	95	13	14
Słomniki - obszar wiejski (5)	1206125	110	680	5729	95	13	14
Sułoszowa (2)	1206132	53	238	3516	89	13	14
Świątniki Górne - miasto (4) *	1206144	4	421	1293	95	13	14
Świątniki Górne - obszar wiejski (5) *	1206145	16	506	4158	95	13	14
Wielka Wieś (2)	1206152	48	1018	5943	95	13	14
Zabierzów (2)	1206162	99	4796	14300	95	13	14
Zielonki (2)	1206172	49	2047	10224	95	13	14
Limanowa (1)	1207011	19	6244	9411	85	30	31
Mszana Dolna (1)	1207021	27	1394	4512	88	61	112
Dobra (2)	1207032	110	571	5488	85	30	31
Jodłownik (2)	1207042	72	398	4661	85	30	31
Kamienica (2)	1207052	95	447	4297	85	30	31
Laskowa (2)	1207062	72	431	4240	85	30	31
Limanowa (2)	1207072	152	702	13460	85	30	31
Łukowica (2)	1207082	70	343	5317	85	30	31
Mszana Dolna (2)	1207092	170	636	9451	88	61	112
Niedzwiedź (2)	1207102	74	500	4056	88	61	112
Słopnice (2) *	1207112	57	233	3249	85	30	31
Tymbark (2)	1207122	33	1327	3608	85	30	31
Charsznica (2)	1208012	78	520	4669	89	110	181
Gołcza (2)	1208022	90	294	3701	89	13	14
Kozłów (2)	1208032	86	263	2985	95	13	14
Książ Wielki (2)	1208042	138	391	3216	95	13	14
Miechów - miasto (4)	1208054	15	3402	7595	95	13	14
Miechów - obszar wiejski (5)	1208055	133	289	4851	95	13	14
Raclawice (2)	1208062	59	112	1455	90	13	14
Słaboszów (2)	1208072	77	117	2227	95	13	14
Dobczyce - miasto (4)	1209014	13	2645	4017	95	13	14
Dobczyce - obszar wiejski (5)	1209015	53	407	4875	95	13	14
Lubień (2)	1209022	75	546	5586	88	13	14
Myślenice - miasto (4)	1209034	30	5524	11402	95	13	14
Myślenice - obszar wiejski (5)	1209035	124	2405	14043	95	13	14
Pcim (2)	1209042	89	1294	6226	95	13	14

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Raciechowice (2)	1209052	61	352	3572	95	13	14
Siepraw (2)	1209062	32	687	4875	95	13	14
Sułkowice - miasto (4)	1209074	16	1167	3959	95	13	14
Sułkowice - obszar wiejski (5)	1209075	44	537	4660	95	13	14
Tokarnia (2)	1209082	69	339	4874	95	13	14
Wiśniowa (2)	1209092	67	341	4089	95	13	14
Grybów (1)	1210011	17	817	3733	96	30	31
Chelmiec (2)	1210022	112	2619	14822	96	30	31
Gródek nad Dunajcem (2)	1210032	88	624	5118	96	30	31
Grybów (2)	1210042	153	1934	13018	96	30	31
Kamionka Wielka (2)	1210052	65	433	5387	96	30	31
Korzenna (2)	1210062	107	643	7714	96	30	31
Krynica-Zdrój - miasto (4) *	1210074	40	3167	6985	86	103	167
Krynica-Zdrój - obszar wiejski (5) *	1210075	105	505	3307	86	103	167
Łabowa (2)	1210082	119	351	3084	86	30	31
Łącko (2)	1210092	133	849	8755	96	30	31
Łososina Dolna (2)	1210102	85	815	5748	96	30	31
Muszyna - miasto (4)	1210114	24	891	3085	86	103	167
Muszyna - obszar wiejski (5)	1210115	117	564	3905	86	103	167
Nawojowa (2)	1210122	50	653	4527	96	30	31
Piwniczna-Zdrój - miasto (4) *	1210134	38	630	3556	96	30	31
Piwniczna-Zdrój - obszar wiejski (5) *	1210135	88	187	2730	96	30	31
Podegrodzie (2)	1210142	65	586	6843	96	30	31
Rytro (2)	1210152	42	239	2200	96	30	31
Stary Sącz - miasto (4)	1210164	15	1317	5699	96	30	31
Stary Sącz - obszar wiejski (5)	1210165	86	475	7860	96	30	31
Nowy Targ (1)	1211011	51	8650	21416	87	61	112
Szczawnica (1) *	1211021	33	1007	4467	87	61	112
Szczawnica (1) *	1211022	55	1007	4467	0	0	0
Czarny Dunajec (2)	1211032	217	1045	12908	93	61	79
Czorsztyn (2)	1211042	62	667	4416	87	61	112
Jabłonka (2)	1211052	213	1390	9928	87	61	112
Krościenko nad Dunajcem (2)	1211062	57	557	3942	87	61	112
Lipnica Wielka (2)	1211072	67	193	3327	87	61	112
Łapsze Niżne (2)	1211082	126	416	5479	87	61	112
Nowy Targ (2)	1211092	208	946	13353	87	61	112
Ochotnica Dolna (2)	1211102	141	704	4550	87	61	112
Raba Wyżna (2)	1211112	89	602	8332	88	61	112
Rabka-Zdrój - miasto (4) *	1211124	36	3224	7905	88	61	112
Rabka-Zdrój - obszar wiejski (5) *	1211125	33	359	2531	88	61	112
Spytkowice (2) *	1211132	32	449	2348	88	61	112
Szaflary (2)	1211142	54	680	6182	87	61	112
Bukowno (1)	1212011	65	2929	6799	89	110	181
Bolesław (2)	1212032	41	2274	4918	89	110	181
Klucze (2)	1212042	120	2133	9507	89	110	181
Olkusz - miasto (4)	1212054	26	8276	26036	89	110	181
Olkusz - obszar wiejski (5)	1212055	124	847	7887	89	110	181
Trzyciąż (2)	1212062	96	468	4171	89	13	14
Wolbrom - miasto (4)	1212074	10	3995	5855	89	110	181
Wolbrom - obszar wiejski (5)	1212075	136	675	8613	89	110	181
Oświęcim (1)	1213011	30	11650	25561	175	27	193
Brzeszcze - miasto (4)	1213024	19	6068	7306	175	27	176
Brzeszcze - obszar wiejski (5)	1213025	27	481	6664	175	27	176
Chelmek - miasto (4)	1213034	8	2141	5961	175	27	176
Chelmek - obszar wiejski (5)	1213035	19	420	2442	175	27	193
Kęty - miasto (4)	1213044	23	5371	12780	175	27	176
Kęty - obszar wiejski (5)	1213045	53	966	8893	175	27	176
Osiek (2)	1213052	41	443	4883	175	27	176
Oświęcim (2)	1213062	75	1361	10621	175	27	193

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Polanka Wielka (2)	1213072	24	259	2483	175	27	193
Przeciszów (2)	1213082	35	256	4198	175	27	193
Zator - miasto (4)	1213094	12	357	2370	94	27	193
Zator - obszar wiejski (5)	1213095	40	313	3277	94	130	72
Koniusza (2)	1214012	89	508	5336	90	13	14
Koszyce (2)	1214022	66	234	3378	182	13	14
Nowe Brzesko (2) *	1214032	54	319	3590	90	13	14
Pałacznica (2)	1214042	48	231	2211	90	13	14
Proszowice - miasto (4)	1214054	7	1833	4108	90	13	14
Proszowice - obszar wiejski (5)	1214055	93	263	6228	90	13	14
Radziemice (2)	1214062	58	263	2031	90	13	14
Jordanów (1)	1215011	21	1971	3295	91	125	217
Sucha Beskidzka (1)	1215021	28	4743	6185	92	125	217
Budzów (2)	1215032	74	269	4936	92	125	217
Bystra-Sidzina (2)	1215042	81	299	3686	91	125	217
Jordanów (2)	1215052	92	435	6298	91	125	217
Maków Podhalański - miasto (4)	1215064	20	895	3445	92	125	217
Maków Podhalański - obszar wiejski (5)	1215065	89	1270	6054	92	125	217
Stryżawa (2)	1215072	114	703	7189	92	125	217
Zawoja (2)	1215082	128	584	5273	92	125	217
Zembrzyce (2)	1215092	39	477	3314	92	125	217
Ciężkowice - miasto (4) *	1216014	10	391	1483	83	35	37
Ciężkowice - obszar wiejski (5) *	1216015	93	381	5092	83	35	37
Gromnik (2)	1216022	70	411	4990	95	35	37
Lisia Góra (2)	1216032	104	747	8364	95	35	37
Pleśna (2)	1216042	83	551	6858	95	35	37
Radłów (2) *	1216052	86	509	5858	95	35	37
Ryglice - miasto (4) *	1216064	25	243	1665	95	35	37
Ryglice - obszar wiejski (5) *	1216065	92	252	5002	95	35	37
Rzepiennik Strzyżewski (2)	1216072	71	334	3898	83	20	21
Skrzyszów (2)	1216082	86	2478	7922	95	35	37
Tarnów (2)	1216092	83	2467	14235	95	35	37
Tuchów - miasto (4)	1216104	18	1091	4068	95	35	37
Tuchów - obszar wiejski (5)	1216105	82	436	6503	95	35	37
Wierzchosławice (2)	1216112	75	878	6548	95	35	37
Wietrzychowice (2)	1216122	48	371	2422	95	35	37
Wojnicz (2) *	1216134	9	1628	7874	81	35	37
Wojnicz (2) *	1216135	70	1628	7874	81	35	37
Zakliczyn - miasto (4) *	1216144	4	474	921	81	35	37
Zakliczyn - obszar wiejski (5) *	1216145	118	364	6456	81	35	37
Żabno - miasto (4)	1216154	11	884	2778	95	35	37
Żabno - obszar wiejski (5)	1216155	91	1394	8864	95	35	37
Szerzyny (2) *	1216162	82	352	4826	132	17	18
Zakopane (1)	1217011	84	7357	16874	93	61	79
Biały Dunajec (2)	1217022	35	275	4112	93	61	79
Bukowina Tatrzańska (2)	1217032	132	708	7524	93	61	79
Kościelisko (2)	1217042	137	453	4910	93	61	79
Poronin (2)	1217052	84	522	6617	93	61	79
Andrychów - miasto (4)	1218014	10	6202	14195	94	130	72
Andrychów - obszar wiejski (5)	1218015	91	924	13013	94	130	72
Brzeźnica (2)	1218022	68	501	6177	84	13	14
Kalwaria Zebrzydowska - miasto (4)	1218034	5	1828	2745	94	130	72
Kalwaria Zebrzydowska - obszar wiejski (5)	1218035	70	1380	9147	94	130	72
Lanckorona (2)	1218042	40	549	3586	94	130	72
Mucharz (2)	1218052	37	212	2358	94	130	72
Spytkowice (2)	1218062	47	571	5784	94	13	14
Stryżów (2)	1218072	46	300	4055	94	130	72
Tomice (2)	1218082	42	560	4524	94	130	72

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Wadowice - miasto (4)	1218094	11	10230	12524	94	130	72
Wadowice - obszar wiejski (5)	1218095	103	1484	11214	94	130	72
Wieprz (2)	1218102	74	1019	6924	94	130	72
Biskupice (2)	1219012	41	357	5418	95	13	14
Gdów (2)	1219022	109	992	10140	95	13	14
Kłaj (2)	1219032	65	973	6243	80	13	14
Niepołomice - miasto (4)	1219044	27	3238	5702	95	13	14
Niepołomice - obszar wiejski (5)	1219045	69	1311	8590	95	13	14
Wieliczka - miasto (4)	1219054	13	4322	12325	95	13	14
Wieliczka - obszar wiejski (5)	1219055	87	1335	18793	95	13	14
Kraków (1)	1261011	327	258474	501892	95	13	14
Nowy Sącz (1)	1262011	58	28896	54558	96	30	31
Tarnów (1)	1263011	72	41064	75791	95	35	37
Białobrzegi - miasto (4)	1401014	8	1470	4916	98	15	16
Białobrzegi - obszar wiejski (5)	1401015	70	241	1800	98	15	16
Promna (2)	1401022	120	272	3347	98	6	6
Radzanów (2)	1401032	83	163	2184	98	6	6
Stara Błotnica (2)	1401042	96	256	3017	98	15	16
Stromiec (2)	1401052	156	268	3267	98	6	6
Wyśmierzyce - miasto (4)	1401064	17	123	536	98	15	16
Wyśmierzyce - obszar wiejski (5)	1401065	89	4	1128	98	15	16
Ciechanów (1)	1402011	33	12855	30578	108	6	55
Ciechanów (2)	1402022	141	238	3649	108	6	55
Głinojeck - miasto (4)	1402034	7	379	2016	108	6	6
Głinojeck - obszar wiejski (5)	1402035	147	570	2776	108	6	6
Gołymín-Ośrodek (2)	1402042	111	256	2390	108	6	55
Grudusk (2)	1402052	97	178	2323	108	6	6
Ojrzeń (2)	1402062	120	207	2627	108	6	55
Opinogóra Górna (2)	1402072	139	180	3569	108	6	55
Regimin (2)	1402082	111	222	2978	108	6	55
Sońsk (2)	1402092	154	727	4760	108	6	6
Garwolin (1)	1403011	22	7622	10350	97	6	6
Łaskarzew (1)	1403021	15	685	3172	97	6	6
Borowie (2)	1403032	80	188	2947	97	6	6
Garwolin (2)	1403042	136	1629	7233	97	6	6
Górzno (2)	1403052	91	219	3551	97	6	6
Łaskarzew (2)	1403062	88	298	3221	97	6	6
Maciejowice (2)	1403072	175	818	4136	97	6	6
Miastków Kościelny (2)	1403082	85	298	2897	97	6	6
Parysów (2)	1403092	64	482	2308	97	6	6
Piława - miasto (4)	1403104	7	681	2660	97	6	6
Piława - obszar wiejski (5)	1403105	70	531	3910	97	6	6
Sobolew (2)	1403112	95	678	4911	97	6	6
Trojanów (2)	1403122	151	378	4495	97	6	6
Wilga (2)	1403132	119	464	3167	97	6	6
Żelechów - miasto (4)	1403144	12	674	2522	97	6	6
Żelechów - obszar wiejski (5)	1403145	75	54	2450	97	6	6
Gostynin (1)	1404011	32	5150	12628	116	23	24
Gostynin (2)	1404022	271	940	7460	116	23	24
Pacyna (2)	1404032	90	337	2344	116	6	6
Sanniki (2)	1404042	95	447	4002	116	6	6
Szczawin Kościelny (2)	1404052	127	235	3111	116	23	24
Milanówek (1)	1405011	13	2270	9937	117	6	6
Podkowa Leśna (1)	1405021	10	380	2393	117	6	6
Baranów (2)	1405032	75	161	3046	117	6	6
Grodzisk Mazowiecki - miasto (4)	1405044	13	7692	17457	117	6	6
Grodzisk Mazowiecki - obszar wiejski (5)	1405045	98	1533	6851	117	6	6
Jaktorów (2)	1405052	55	778	6429	117	6	6
Żabia Wola (2)	1405062	103	1105	4063	117	6	6

nazwa gminy	gmina				numer lokalnego rynku pracy		
	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Belsk Duży (2)	1406012	108	1756	4173	98	6	6
Błędów (2)	1406022	134	318	4726	98	6	6
Chynów (2)	1406032	135	623	5908	107	6	6
Goszczyn (2)	1406042	57	179	1790	98	6	6
Grójec - miasto (4)	1406054	9	5642	9970	98	6	6
Grójec - obszar wiejski (5)	1406055	112	589	5201	98	6	6
Jasieniec (2)	1406062	108	210	3292	98	6	6
Mogielnica - miasto (4)	1406074	13	585	1528	98	131	232
Mogielnica - obszar wiejski (5)	1406075	128	262	3889	98	131	232
Nowe Miasto nad Pilicą - miasto (4)	1406084	11	686	2430	71	131	232
Nowe Miasto nad Pilicą - obszar wiejski (5)	1406085	148	245	2586	71	131	232
Pniewy (2)	1406092	102	413	2839	98	6	6
Warka - miasto (4)	1406114	26	1869	7256	107	6	6
Warka - obszar wiejski (5)	1406115	177	240	4780	107	6	6
Garbatka-Letnisko (2)	1407012	74	938	3210	112	15	198
Głowaczów (2)	1407022	186	377	4186	112	15	198
Gniewoszków (2)	1407032	84	161	2453	48	57	73
Grabów nad Pilicą (2)	1407042	125	118	2184	107	6	6
Kozienice - miasto (4)	1407054	10	5109	12727	112	15	198
Kozienice - obszar wiejski (5)	1407055	234	3574	7364	112	15	198
Magnuszew (2)	1407062	141	384	3884	112	6	6
Sieciechów (2)	1407072	62	219	2459	51	15	198
Legionowo (1)	1408011	14	6107	34600	117	6	6
Jabłonna (2)	1408022	65	1321	9110	117	6	6
Nieporęt (2)	1408032	100	2150	8662	117	6	6
Serock - miasto (4)	1408044	12	621	2407	117	6	6
Serock - obszar wiejski (5)	1408045	96	640	4972	117	6	6
Wieliszew (2) *	1408052	102	1079	5705	117	6	6
Chotcza (2)	1409012	90	93	1455	99	15	60
Ciepielów (2)	1409022	135	473	3362	99	15	60
Lipsko - miasto (4)	1409034	16	2160	3929	99	15	60
Lipsko - obszar wiejski (5)	1409035	119	178	3539	99	15	60
Rzeczniów (2)	1409042	103	200	2727	99	37	41
Sienno (2)	1409052	147	269	3742	99	15	60
Solec nad Wisłą (2)	1409062	130	289	3243	99	15	60
Huszlew (2)	1410012	118	150	1600	100	6	6
Łosice - miasto (4)	1410024	24	1933	4680	100	68	89
Łosice - obszar wiejski (5)	1410025	97	431	2288	100	68	89
Olszanka (2)	1410032	88	97	1814	100	68	89
Platerów (2)	1410042	129	220	2928	100	68	89
Sarnaki (2)	1410052	197	362	2955	100	6	6
Stara Kornica (2)	1410062	119	228	2715	100	68	89
Maków Mazowiecki (1)	1411011	10	2420	6393	101	6	67
Czerwonka (2)	1411022	110	73	1501	101	6	67
Karniewo (2)	1411032	129	223	3160	101	6	67
Krasnosielc (2)	1411042	168	315	3710	101	6	6
Młynarze (2)	1411052	75	119	995	101	82	110
Płoniawy-Bramura (2)	1411062	135	316	3502	101	6	67
Różan - miasto (4)	1411074	7	502	1737	101	6	67
Różan - obszar wiejski (5)	1411075	78	17	991	101	6	67
Rzewnie (2)	1411082	112	183	1511	101	6	6
Sypniewo (2)	1411092	128	153	2070	101	6	6
Szelków (2)	1411102	113	143	2182	101	6	67
Mińsk Mazowiecki (1)	1412011	13	8993	24598	117	6	6
Ceglów (2)	1412042	96	522	3822	117	6	6
Dębe Wielkie (2)	1412052	78	1109	5260	117	6	6
Dobre (2)	1412062	125	434	3556	117	6	6
Halinów - miasto (4) *	1412074	3	539	2142	117	6	6

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Halinów - obszar wiejski (5) *	1412075	60	722	6052	117	6	6
Jakubów (2)	1412082	87	222	2945	117	6	6
Kałużyn - miasto (4)	1412094	12	522	1816	117	6	6
Kałużyn - obszar wiejski (5)	1412095	82	181	1922	117	6	6
Latowicz (2)	1412102	114	286	3244	117	6	6
Mińsk Mazowiecki (2)	1412112	112	1651	8172	117	6	6
Mrozy (2)	1412122	145	585	5322	117	6	6
Siennica (2)	1412132	111	483	4247	117	6	6
Stanisławów (2)	1412142	107	994	3807	117	6	6
Sulejówek (1) *	1412151	19	2330	11892	117	6	6
Mława (1)	1413011	35	11668	19260	102	120	68
Dzierzgowo (2)	1413022	151	153	1911	102	120	68
Lipowiec Kościelny (2)	1413032	114	342	2950	102	120	68
Radzanów (2)	1413042	99	154	2012	102	120	68
Strzegowo (2) *	1413052	215	386	4614	102	6	6
Stupsk (2)	1413062	118	358	2917	102	120	68
Szreńsk (2)	1413072	109	317	2642	102	120	68
Szydłowo (2)	1413082	122	807	2737	102	6	6
Wieczfnia Kościelna (2)	1413092	120	832	2483	103	120	68
Wiśniewo (2)	1413102	99	277	3068	102	120	68
Nowy Dwór Mazowiecki (1)	1414011	28	5661	18397	117	6	6
Czosnów (2)	1414022	128	2223	5651	117	6	6
Leoncin (2)	1414032	158	622	3207	117	6	6
Nasielsk - miasto (4)	1414044	13	1532	4837	57	6	6
Nasielsk - obszar wiejski (5)	1414045	193	320	7261	57	6	6
Pomiechówek (2)	1414052	103	428	5640	117	6	6
Zakroczym - miasto (4)	1414064	20	532	2182	117	6	6
Zakroczym - obszar wiejski (5)	1414065	52	56	1765	117	6	6
Baranowo (2)	1415012	198	599	3705	213	63	83
Czarnia (2)	1415022	93	143	1429	104	6	6
Czerwin (2)	1415032	170	201	3049	104	82	110
Goworowo (2)	1415042	219	295	4987	104	6	6
Kadzidło (2)	1415052	258	616	6531	104	82	110
Lelis (2)	1415062	196	287	4931	104	82	110
Łyse (2)	1415072	247	1331	4619	104	104	184
Myszyniec - miasto (4)	1415084	11	502	1899	104	104	184
Myszyniec - obszar wiejski (5)	1415085	217	137	4029	104	104	184
Olszewo-Borki (2)	1415092	197	593	5899	104	82	110
Rzekuń (2)	1415102	136	1087	5675	104	82	110
Troszyn (2)	1415112	156	243	2844	104	82	110
Ostrów Mazowiecka (1)	1416011	22	6606	14543	105	82	53
Andrzejewo (2)	1416022	119	193	2484	105	82	53
Boguty-Pianki (2)	1416032	89	88	1628	152	6	6
Brok - miasto (4)	1416044	28	129	1163	105	82	53
Brok - obszar wiejski (5)	1416045	82	13	551	105	82	53
Małkinia Górna (2)	1416052	138	1521	7680	105	6	6
Nur (2)	1416062	96	152	1661	117	6	6
Ostrów Mazowiecka (2)	1416072	278	580	7485	105	82	53
Stary Lubotyń (2)	1416082	110	108	2220	105	82	53
Szulborze Wielkie (2)	1416092	47	68	1007	117	6	6
Wąsewo (2)	1416102	120	178	2607	105	82	53
Zaręby Kościelne (2)	1416112	89	225	2112	105	6	6
Józefów (1)	1417011	24	3577	11728	117	6	6
Otwock (1)	1417021	47	8548	26966	117	6	6
Celestynów (2)	1417032	89	1602	7205	117	6	6
Karczew - miasto (4)	1417044	28	2636	7049	117	6	6
Karczew - obszar wiejski (5)	1417045	53	242	3580	117	6	6
Kołbiel (2)	1417052	108	1198	4936	117	6	6
Osieck (2)	1417062	68	310	2073	117	6	6

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Sobienie-Jeziory (2)	1417072	98	287	3721	117	6	6
Wiązowna (2)	1417082	102	2862	6393	117	6	6
Góra Kalwaria - miasto (4)	1418014	14	2360	7109	107	6	6
Góra Kalwaria - obszar wiejski (5)	1418015	131	1239	8568	107	6	6
Konstancin-Jeziorna - miasto (4)	1418024	18	3788	10586	107	6	6
Konstancin-Jeziorna - obszar wiejski (5)	1418025	61	294	4345	107	6	6
Lesznowola (2)	1418032	69	5203	11065	106	6	6
Piaseczno - miasto (4)	1418044	16	17895	24828	107	6	6
Piaseczno - obszar wiejski (5)	1418045	112	1712	17015	107	6	6
Prażmów (2)	1418052	86	343	5573	107	6	6
Tarczyn - miasto (4) *	1418064	5	1013	2495	106	6	6
Tarczyn - obszar wiejski (5) *	1418065	109	143	4278	106	6	6
Bielsk (2)	1419012	125	764	5493	116	23	24
Bodzanów (2)	1419022	136	508	5071	116	23	24
Brudzeń Duży (2)	1419032	159	569	4791	116	23	24
Bulkowo (2)	1419042	117	420	3492	116	23	24
Drobin - miasto (4)	1419054	10	352	1886	116	23	24
Drobin - obszar wiejski (5)	1419055	133	283	3179	116	23	24
Gąbin - miasto (4)	1419064	28	608	2646	116	23	24
Gąbin - obszar wiejski (5)	1419065	118	346	4055	116	23	24
Łąck (2)	1419072	94	669	3137	116	23	24
Mała Wieś (2)	1419082	109	466	3742	116	23	24
Nowy Duninów (2)	1419092	146	132	2459	116	23	24
Radzanowo (2)	1419102	104	375	4589	116	23	24
Słubice (2)	1419112	96	306	2720	116	6	6
Słupno (2)	1419122	75	859	3532	116	23	24
Stara Biała (2)	1419132	111	947	6348	116	23	24
Staroźreby (2)	1419142	138	261	4395	116	23	24
Wyszogród - miasto (4)	1419154	13	413	1756	116	6	6
Wyszogród - obszar wiejski (5)	1419155	83	77	1868	116	6	6
Płońsk (1)	1420011	12	6412	14887	117	6	6
Raciąż (1)	1420021	8	885	3014	117	6	6
Baboszewo (2)	1420032	163	754	4748	117	6	6
Czerwińsk nad Wisłą (2)	1420042	144	316	4693	117	6	6
Dzierżążnia (2)	1420052	102	121	2235	117	6	6
Joniec (2)	1420062	73	162	1550	117	6	6
Naruszewo (2)	1420072	159	339	3938	117	6	6
Nowe Miasto (2)	1420082	118	311	2865	117	6	6
Płońsk (2)	1420092	127	445	4349	117	6	6
Raciąż (2)	1420102	244	322	4918	117	6	6
Sochocin (2)	1420112	122	406	3507	117	6	6
Załuski (2)	1420122	108	234	3249	117	6	6
Piastów (1)	1421011	6	2167	15571	117	6	6
Pruszków (1)	1421021	19	13704	35995	117	6	6
Brwinów - miasto (4)	1421034	10	1088	7559	117	6	6
Brwinów - obszar wiejski (5)	1421035	59	1111	6171	117	6	6
Michałowice (2)	1421042	35	2866	10139	117	6	6
Nadarzyn (2)	1421052	73	2809	6782	117	6	6
Raszyn (2)	1421062	44	6325	12705	117	6	6
Przasnysz (1)	1422011	25	4470	11047	108	6	56
Chorzele - miasto (4)	1422024	18	890	1765	109	6	56
Chorzele - obszar wiejski (5)	1422025	354	205	4137	109	6	56
Czernice Borowe (2)	1422032	120	164	2301	108	6	56
Jednorożec (2)	1422042	232	375	4065	109	6	56
Krasne (2)	1422052	101	248	2282	101	6	67
Krzynowłoga Mała (2)	1422062	185	131	2061	109	6	56
Przasnysz (2)	1422072	184	205	4155	108	6	56
Borkowice (2)	1423012	86	394	2606	71	25	40
Gielniów (2)	1423022	79	299	2776	71	25	40

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podjęcie taksonomiczne - słabe kryterium grupowania	podjęcie taksonomiczne - mocne kryterium grupowania
Klów (2)	1423032	87	197	1916	71	6	6
Odrzywół (2)	1423042	99	292	2311	71	6	6
Potworów (2)	1423052	82	192	2457	117	6	6
Przysucha - miasto (4)	1423064	7	1873	4119	71	25	40
Przysucha - obszar wiejski (5)	1423065	174	379	3521	71	25	40
Rusinów (2)	1423072	83	145	2529	71	25	40
Wieniawa (2)	1423082	104	252	3186	71	15	16
Gzy (2)	1424012	104	312	2383	108	6	6
Obryte (2)	1424022	141	319	2727	117	6	6
Pokrzywnica (2)	1424032	121	353	2874	117	6	6
Pułusk - miasto (4)	1424044	23	4021	12767	117	6	6
Pułusk - obszar wiejski (5)	1424045	111	388	2704	117	6	6
Świercze (2)	1424052	93	144	2889	117	6	6
Winnica (2)	1424062	115	380	2430	57	6	6
Zatory (2)	1424072	121	181	2723	117	6	6
Pionki (1)	1425011	18	2757	13152	112	15	198
Gózd (2)	1425022	78	423	4904	112	15	16
Ilża - miasto (4)	1425034	16	1263	3237	190	37	41
Ilża - obszar wiejski (5)	1425035	240	325	6323	190	37	41
Jastrzębia (2)	1425042	90	333	3855	112	15	16
Jedlińsk (2)	1425052	139	1245	8131	112	15	16
Jedlnia-Letnisko (2)	1425062	66	731	7244	112	15	16
Kowała (2)	1425072	75	605	6577	112	15	16
Pionki (2)	1425082	230	526	5910	112	15	198
Przytyk (2)	1425092	134	383	4113	112	15	16
Skaryszew - miasto (4)	1425104	27	413	2560	112	15	16
Skaryszew - obszar wiejski (5)	1425105	144	768	5608	112	15	16
Wierzbica (2)	1425112	94	858	6337	112	15	16
Wolanów (2)	1425122	83	677	5045	112	15	16
Zakrzew (2)	1425132	96	611	6903	112	15	16
Domanice (2)	1426012	47	114	1530	110	70	91
Korczew (2)	1426022	105	151	1657	100	70	91
Kotuń (2)	1426032	150	599	4970	110	6	6
Mokobody (2)	1426042	119	330	2989	110	70	91
Mordy - miasto (4)	1426054	5	233	1139	110	70	91
Mordy - obszar wiejski (5)	1426055	166	58	2499	110	70	91
Paprotnia (2)	1426062	81	101	1501	110	70	91
Przesmyki (2)	1426072	117	119	1993	100	70	91
Siedlce (2)	1426082	141	3036	10118	117	70	91
Skórzec (2)	1426092	119	391	4197	110	70	91
Suchożebry (2)	1426102	101	356	2774	110	70	91
Wiśniew (2)	1426112	126	438	3315	110	70	91
Wodynie (2)	1426122	115	139	2669	110	70	91
Zbuczyn (2) *	1426132	212	549	5661	110	70	91
Sierpc (1)	1427011	19	5460	12231	116	23	120
Gozdowo (2)	1427022	126	310	3584	116	23	24
Mochowo (2)	1427032	144	286	3689	116	23	24
Rościszewo (2)	1427042	115	113	2352	116	23	120
Sierpc (2)	1427052	150	376	4057	116	23	120
Szczutowo (2)	1427062	113	230	2523	116	23	120
Zawidz (2)	1427072	185	328	3985	116	23	120
Sochaczew (1)	1428011	26	8977	24921	117	6	6
Brochów (2) *	1428022	120	322	2675	117	6	6
Ilów (2)	1428032	129	367	3758	117	6	6
Młodziszyn (2)	1428042	117	201	3277	117	6	6
Nowa Sucha (2)	1428052	90	325	3689	117	6	6
Rybno (2)	1428062	73	139	2072	117	6	6
Sochaczew (2)	1428072	91	670	5488	117	6	6
Teresin (2)	1428082	88	3966	7054	117	6	6

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Sokołów Podlaski (1)	1429011	18	5989	12114	110	47	57
Bielany (2)	1429022	110	218	2178	110	47	57
Ceranów (2)	1429032	111	127	1333	111	47	57
Jabłonna Lacka (2)	1429042	149	717	2822	110	47	57
Kosów Lacki - miasto (4) *	1429054	12	562	1277	111	47	57
Kosów Lacki - obszar wiejski (5) *	1429055	188	62	2395	111	47	57
Repki (2)	1429062	168	231	3215	110	47	57
Sabnie (2)	1429072	108	215	2227	110	47	57
Sokołów Podlaski (2)	1429082	137	324	3621	110	47	57
Sterdyń (2)	1429092	130	214	2432	110	47	57
Chlewiska (2)	1430012	124	275	3523	189	37	33
Jastrząb (2)	1430022	55	274	3069	189	37	33
Mirów (2)	1430032	53	153	2202	189	37	33
Orońsko (2)	1430042	82	462	3463	112	15	16
Szydłowiec - miasto (4)	1430054	22	2080	8217	112	37	33
Szydłowiec - obszar wiejski (5)	1430055	116	467	4169	189	37	33
Błonie - miasto (4)	1432014	9	4575	7733	117	6	6
Błonie - obszar wiejski (5)	1432015	77	994	4865	117	6	6
Izabelin (2)	1432022	69	1354	6570	117	6	6
Kampinos (2)	1432032	85	450	2494	117	6	6
Leszno (2)	1432042	124	884	5767	117	6	6
Łomianki - miasto (4)	1432054	8	4536	10468	117	6	6
Łomianki - obszar wiejski (5)	1432055	31	1796	4348	117	6	6
Ożarów Mazowiecki - miasto (4)	1432064	6	2499	5232	117	6	6
Ożarów Mazowiecki - obszar wiejski (5)	1432065	65	2307	8580	117	6	6
Stare Babice (2)	1432072	60	2372	10217	117	6	6
Węgrów (1)	1433011	36	4015	8206	113	47	155
Grębków (2)	1433022	131	188	2657	113	6	6
Korytnica (2)	1433032	180	257	3802	113	6	6
Liw (2)	1433042	169	400	4442	113	47	155
Łochów - miasto (4)	1433054	13	919	4257	117	6	6
Łochów - obszar wiejski (5)	1433055	184	339	6562	117	6	6
Miedzna (2)	1433062	116	299	2420	113	47	155
Sadowne (2)	1433072	145	315	3583	117	6	6
Stoczek (2)	1433082	145	222	2971	113	6	6
Wierzbno (2)	1433092	103	152	1694	113	6	6
Kobyłka (1)	1434011	20	2498	11713	117	6	6
Marki (1)	1434021	26	4829	15484	117	6	6
Ząbki (1)	1434031	11	3692	16072	117	6	6
Zielonka (1)	1434041	79	3264	10617	117	6	6
Dąbrówka (2)	1434052	109	313	4300	117	6	6
Jadów (2)	1434062	117	308	4622	117	6	6
Klembów (2)	1434072	86	654	5598	117	6	6
Poświętne (2)	1434082	104	272	3635	117	6	6
Radzymin - miasto (4)	1434094	23	1822	4935	117	6	6
Radzymin - obszar wiejski (5)	1434095	106	1331	7460	117	6	6
Strachówka (2)	1434102	107	137	1739	117	6	6
Łuszczyca - miasto (4)	1434114	8	1051	4920	117	6	6
Łuszczyca - obszar wiejski (5)	1434115	95	261	6998	117	6	6
Wołomin - miasto (4)	1434124	17	8388	24117	117	6	6
Wołomin - obszar wiejski (5)	1434125	45	1102	8220	117	6	6
Brańszczyk (2)	1435012	167	644	4975	117	6	6
Długosiodło (2)	1435022	167	305	4389	117	6	6
Rząśnik (2)	1435032	167	263	3923	117	6	6
Somianka (2)	1435042	117	450	3142	117	6	6
Wyszaków - miasto (4)	1435054	21	6804	18051	117	6	6
Wyszaków - obszar wiejski (5)	1435055	143	371	6513	117	6	6
Zabrodzie (2)	1435062	92	357	3416	117	6	6
Kazanów (2)	1436012	95	201	2668	112	15	16

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	pojęcie taksonomiczne - słabe kryterium grupowania	pojęcie taksonomiczne - mocne kryterium grupowania
Policzna (2)	1436022	113	230	3358	114	15	198
Przyłęk (2)	1436032	132	308	3744	114	57	73
Tczów (2)	1436042	72	259	2881	114	15	16
Zwoleń - miasto (4)	1436054	16	2253	5306	114	15	16
Zwoleń - obszar wiejski (5)	1436055	145	155	4185	114	15	16
Biezuń - miasto (4)	1437014	12	284	1146	115	92	133
Biezuń - obszar wiejski (5)	1437015	109	71	1901	115	92	133
Kuczbork-Osada (2)	1437022	122	297	2931	115	92	133
Lubowidz (2)	1437032	192	350	4235	115	92	133
Lutocin (2)	1437042	126	141	2569	115	92	133
Siemiatkowo (2) *	1437052	113	157	2035	102	6	6
Żuromin - miasto (4)	1437064	11	2403	5672	115	92	133
Żuromin - obszar wiejski (5)	1437065	122	198	3171	115	92	133
Żyrardów (1)	1438011	14	6931	26809	117	6	6
Mszczonów - miasto (4)	1438024	9	2444	4159	73	6	6
Mszczonów - obszar wiejski (5)	1438025	144	718	2871	73	6	6
Puszcza Mariańska (2)	1438032	143	357	5275	117	6	6
Radziejowice (2)	1438042	73	732	2897	73	6	6
Wisłok (2)	1438052	150	578	5724	117	6	6
Ostrołęka (1)	1461011	28	14491	36248	104	82	110
Płock (1)	1462011	88	43901	85448	116	23	24
Radom (1)	1463011	112	48161	147061	112	15	16
Siedlce (1)	1464011	32	23233	51636	110	70	91
M.st.Warszawa (1) *	1465011	517	759447	1113637	117	6	6
Brzeg (1)	1601011	15	8608	24981	128	49	10
Skarbimierz (2) *	1601022	110	765	4564	128	49	10
Grodziszewo - miasto (4)	1601034	10	2234	5860	118	49	10
Grodziszewo - obszar wiejski (5)	1601035	275	648	6850	118	49	10
Lewin Brzeski - miasto (4)	1601044	12	772	3781	128	10	10
Lewin Brzeski - obszar wiejski (5)	1601045	148	445	4844	128	10	10
Lubsza (2)	1601052	213	552	5553	128	49	10
Olszanka (2)	1601062	93	476	3063	128	49	10
Baborów - miasto (4)	1602014	12	393	2055	119	78	103
Baborów - obszar wiejski (5)	1602015	105	174	2080	119	78	103
Branice (2)	1602022	122	1239	4898	119	78	103
Głubczyce - miasto (4)	1602034	13	4465	8703	119	78	103
Głubczyce - obszar wiejski (5)	1602035	281	1204	6679	119	78	103
Kietrz - miasto (4)	1602044	19	912	4088	171	98	54
Kietrz - obszar wiejski (5)	1602045	121	321	3380	119	78	103
Kędzierzyn-Koźle (1)	1603011	124	17189	42668	127	132	146
Bierawa (2)	1603022	119	1125	5124	127	132	146
Cisiek (2)	1603032	71	417	4258	127	132	146
Pawłowiczki (2)	1603042	153	757	5363	127	132	146
Polska Cerekiew (2)	1603052	60	382	2918	127	132	146
Reńska Wieś (2)	1603062	98	446	5384	127	132	146
Byczyna - miasto (4)	1604014	6	585	2398	77	10	132
Byczyna - obszar wiejski (5)	1604015	177	332	3732	77	10	132
Kluczbork - miasto (4)	1604024	12	6924	17149	128	10	132
Kluczbork - obszar wiejski (5)	1604025	205	739	7955	128	10	132
Lasowice Wielkie (2)	1604032	211	324	4523	128	10	132
Wołczyn - miasto (4)	1604044	7	1043	3914	128	10	132
Wołczyn - obszar wiejski (5)	1604045	234	357	5239	128	10	132
Gogolin - miasto (4)	1605014	20	729	3862	120	132	236
Gogolin - obszar wiejski (5)	1605015	81	1450	3619	120	132	236
Krapkowice - miasto (4)	1605024	21	5985	12289	127	132	236
Krapkowice - obszar wiejski (5)	1605025	76	191	3974	127	132	236
Strzeleczy (2)	1605032	118	555	4841	126	132	236
Walce (2)	1605042	69	486	3724	126	0	0
Zdzieszowice - miasto (4)	1605054	12	4844	9358	127	10	10

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Zdzieszowice - obszar wiejski (5)	1605055	45	559	2583	127	10	10
Domaszowice (2)	1606012	114	254	2423	128	126	121
Namysłów - miasto (4)	1606024	23	4647	10987	128	126	121
Namysłów - obszar wiejski (5)	1606025	267	431	6069	128	126	121
Pokój (2)	1606032	133	479	3698	128	10	10
Świerczów (2)	1606042	110	178	2294	128	126	121
Wilków (2)	1606052	101	261	2945	128	126	121
Głucholazy - miasto (4)	1607014	7	3156	9633	122	10	117
Głucholazy - obszar wiejski (5)	1607015	161	843	6811	122	10	117
Kamiennik (2)	1607022	89	224	2349	122	10	117
Korfantów - miasto (4)	1607034	10	764	1220	121	10	117
Korfantów - obszar wiejski (5)	1607035	170	271	4901	121	10	117
Łambinowice (2)	1607042	124	728	5312	121	10	10
Nysa - miasto (4)	1607054	28	10339	31383	122	10	117
Nysa - obszar wiejski (5)	1607055	190	1181	8159	122	10	117
Otmuchów - miasto (4)	1607064	28	1409	3384	123	10	117
Otmuchów - obszar wiejski (5)	1607065	159	515	5682	123	10	117
Paczków - miasto (4)	1607074	7	1607	5263	124	10	117
Paczków - obszar wiejski (5)	1607075	73	433	3629	124	10	117
Pakosławice (2)	1607082	74	331	2541	122	10	117
Skoroszyce (2)	1607092	104	700	4151	118	49	10
Dobrodzień - miasto (4)	1608014	20	1406	2660	125	113	187
Dobrodzień - obszar wiejski (5)	1608015	143	573	4028	125	113	187
Gorzów Śląski - miasto (4)	1608024	19	437	1667	76	113	77
Gorzów Śląski - obszar wiejski (5)	1608025	135	330	3212	76	113	77
Olesno - miasto (4)	1608034	15	3349	6636	125	113	77
Olesno - obszar wiejski (5)	1608035	225	244	5287	125	113	77
Praszka - miasto (4)	1608044	9	2147	5991	76	113	77
Praszka - obszar wiejski (5)	1608045	94	240	3453	76	113	77
Radłów (2)	1608052	117	181	2818	125	113	77
Rudniki (2)	1608062	100	529	5171	76	113	77
Zębowice (2)	1608072	96	210	2627	125	113	187
Chrzastowice (2)	1609012	82	448	4410	128	10	10
Dąbrowa (2)	1609022	131	610	6254	128	10	10
Dobrzeń Wielki (2)	1609032	91	3498	9204	128	10	10
Komprachcice (2)	1609042	56	485	7550	128	10	10
Łubniany (2)	1609052	126	790	5944	128	10	10
Murów (2)	1609062	160	777	3769	128	10	10
Niemodlin - miasto (4)	1609074	13	1287	4637	128	10	10
Niemodlin - obszar wiejski (5)	1609075	170	513	4459	128	10	10
Ozimek - miasto (4)	1609084	3	2912	6984	128	10	10
Ozimek - obszar wiejski (5)	1609085	123	2626	7224	128	10	10
Popielów (2)	1609092	175	433	5478	128	10	10
Prószków - miasto (4) *	1609104	16	655	1698	128	10	10
Prószków - obszar wiejski (5) *	1609105	105	471	4659	128	10	10
Tarnów Opolski (2)	1609112	82	1275	6458	128	10	10
Tułowice (2)	1609122	81	485	3773	121	10	10
Turawa (2)	1609132	172	804	6086	128	10	10
Biała - miasto (4)	1610014	15	582	1676	126	10	10
Biała - obszar wiejski (5)	1610015	182	294	5557	126	10	10
Głogówek - miasto (4)	1610024	22	1255	3566	126	132	146
Głogówek - obszar wiejski (5)	1610025	148	394	5212	126	132	146
Lubrza (2)	1610032	83	367	2857	126	10	10
Prudnik - miasto (4)	1610044	20	5368	14726	126	10	10
Prudnik - obszar wiejski (5)	1610045	102	249	3928	126	10	10
Izbicko (2)	1611012	85	296	3562	128	10	10
Jemielnica (2)	1611022	113	316	4975	0	0	0
Kolonowskie - miasto (4)	1611034	56	587	2217	169	0	0
Kolonowskie - obszar wiejski (5)	1611035	28	152	1782	0	0	0

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Leśnica - miasto (4)	1611044	15	362	1851	127	0	0
Leśnica - obszar wiejski (5)	1611045	80	164	3596	127	132	146
Strzelce Opolskie - miasto (4)	1611054	30	5267	13441	128	10	10
Strzelce Opolskie - obszar wiejski (5)	1611055	172	755	8487	128	10	10
Ujazd - miasto (4)	1611064	15	315	1040	127	132	146
Ujazd - obszar wiejski (5)	1611065	68	343	2863	127	132	146
Zawadzkie - miasto (4)	1611074	16	1889	5489	169	134	239
Zawadzkie - obszar wiejski (5)	1611075	66	360	2867	169	134	239
Opole (1)	1661011	97	47829	86845	128	10	10
Czarna (2)	1801032	185	125	1467	141	16	118
Lutowiska (2)	1801052	476	276	1451	141	16	118
Ustrzyki Dolne - miasto (4)	1801084	17	2058	6474	141	16	118
Ustrzyki Dolne - obszar wiejski (5)	1801085	461	744	4851	141	16	118
Brzozów - miasto (4)	1802014	11	3940	4930	129	16	17
Brzozów - obszar wiejski (5)	1802015	92	703	11185	129	16	17
Domaradz (2)	1802022	57	325	3524	129	2	2
Dydnia (2)	1802032	129	616	4822	141	16	17
Haczów (2)	1802042	72	371	5533	146	7	7
Jasienica Rosielna (2)	1802052	57	560	4343	129	16	17
Nozdrzec (2)	1802062	121	470	4889	129	16	17
Dębica (1)	1803011	34	17718	31072	130	77	102
Brzostek (2) *	1803022	122	685	7698	130	77	102
Czarna (2)	1803032	148	667	7667	130	77	102
Dębica (2)	1803042	138	3553	14772	130	77	102
Jodłowa (2)	1803052	60	247	3212	130	77	102
Pilzno - miasto (4)	1803064	16	1491	2835	130	77	102
Pilzno - obszar wiejski (5)	1803065	149	647	7473	130	77	102
Żyraków (2)	1803072	110	1017	8090	130	77	102
Jarosław (1)	1804011	35	13306	26753	131	29	30
Radymno (1)	1804021	14	1054	3597	131	29	30
Chłopice (2)	1804032	49	273	3365	131	29	30
Jarosław (2)	1804042	113	679	7832	131	29	30
Laszki (2)	1804052	138	399	4135	131	29	30
Pawłosiów (2)	1804062	47	315	5072	131	29	30
Pruchnik (2) *	1804072	78	489	5723	131	29	30
Radymno (2)	1804082	182	827	6724	131	29	30
Rokietnica (2)	1804092	57	249	2515	131	29	30
Rożwienica (2)	1804102	71	399	3745	131	29	30
Wiązownica (2)	1804112	245	663	6735	131	29	30
Jaśło (1)	1805011	37	15543	24393	132	17	18
Brzyska (2)	1805022	45	196	3722	132	17	18
Dębowiec (2)	1805032	86	536	5073	132	17	18
Jaśło (2)	1805042	93	839	9590	132	17	18
Końce (2) *	1805052	61	624	5300	132	17	18
Krempna (2)	1805062	204	176	1266	132	17	18
Nowy Żmigród (2)	1805072	104	457	5544	132	17	18
Osiek Jasielski (2)	1805082	60	181	3188	132	17	18
Skotyszyn (2)	1805092	78	1310	7560	132	17	18
Tarnowiec (2)	1805112	63	802	5452	132	17	18
Cmolas (2)	1806012	134	542	4628	133	33	78
Kolbuszowa - miasto (4)	1806024	8	3292	6177	133	33	78
Kolbuszowa - obszar wiejski (5)	1806025	163	881	9252	133	33	78
Majdan Królewski (2)	1806032	155	313	5920	143	36	78
Niwiska (2)	1806042	95	305	3472	133	33	78
Raniżów (2)	1806052	97	732	4218	133	2	2
Dzikowiec (2) *	1806062	122	226	3873	133	33	78
Chorkówka (2)	1807012	77	506	8018	146	7	7
Dukla - miasto (4)	1807024	6	691	1398	146	7	7
Dukla - obszar wiejski (5)	1807025	329	602	8532	146	7	7

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Iwonicz-Zdrój - miasto (4)	1807034	6	895	1164	146	7	7
Iwonicz-Zdrój - obszar wiejski (5)	1807035	39	753	5584	146	7	7
Jedlicze - miasto (4)	1807044	11	2073	3563	146	7	7
Jedlicze - obszar wiejski (5)	1807045	48	382	5850	146	7	7
Korczyna (2)	1807052	94	1146	6537	146	7	7
Krościenko Wyżne (2)	1807062	16	252	3202	146	7	7
Miejsce Piastowe (2)	1807072	51	1218	8152	146	7	7
Rymanów - miasto (4)	1807084	12	1053	2237	146	7	7
Rymanów - obszar wiejski (5)	1807085	155	1059	7308	146	7	7
Wojaszówka (2)	1807092	82	452	5586	146	7	7
Leżajsk (1)	1808011	21	6326	9529	134	91	128
Grodzisko Dolne (2)	1808022	79	311	4877	134	91	128
Kuryłówka (2)	1808032	142	283	3361	134	91	128
Leżajsk (2)	1808042	198	845	11816	134	91	128
Nowa Sarzyna - miasto (4)	1808054	9	2354	4332	135	91	128
Nowa Sarzyna - obszar wiejski (5)	1808055	135	492	8795	135	91	128
Lubaczów (1)	1809011	26	3486	8307	136	79	106
Cieszanów - miasto (4)	1809024	15	358	1243	136	79	106
Cieszanów - obszar wiejski (5)	1809025	204	624	3226	136	79	106
Horyniec-Zdrój (2) *	1809032	203	536	2982	136	79	106
Lubaczów (2)	1809042	203	382	5316	136	79	106
Narol - miasto (4) *	1809054	12	350	1340	137	79	106
Narol - obszar wiejski (5) *	1809055	191	341	3652	137	79	106
Oleszyce - miasto (4)	1809064	5	385	1973	136	79	106
Oleszyce - obszar wiejski (5)	1809065	147	158	2045	136	79	106
Stary Dzików (2)	1809072	156	216	2696	136	79	106
Wielkie Oczy (2)	1809082	146	184	2243	136	79	106
Łañcut (1)	1810011	19	6012	11337	147	2	2
Białobrzegi (2)	1810022	56	1670	4802	147	2	2
Czarna (2)	1810032	78	825	6518	147	2	2
Łañcut (2)	1810042	107	1568	12263	147	2	2
Markowa (2)	1810052	69	386	3887	147	2	2
Rakszawa (2)	1810062	66	910	4262	147	2	2
Żołynia (2)	1810072	57	616	4087	134	2	2
Mielec (1)	1811011	47	22905	40235	138	33	35
Borowa (2)	1811022	55	356	3267	138	33	35
Czermin (2)	1811032	80	334	3972	138	33	35
Gawłuszowice (2)	1811042	34	224	1674	138	33	35
Mielec (2)	1811052	122	1028	7299	138	33	35
Padew Narodowa (2)	1811062	71	265	3325	138	33	35
Przeclaw (2) *	1811072	134	669	6582	138	33	35
Radomyśl Wielki - miasto (4)	1811084	9	379	1811	138	33	35
Radomyśl Wielki - obszar wiejski (5)	1811085	152	548	6287	138	33	35
Tuszów Narodowy (2)	1811092	89	642	4505	138	33	35
Wadowice Górne (2)	1811102	87	931	4246	138	33	35
Harasiuki (2)	1812012	169	268	3725	139	28	29
Jarocin (2)	1812022	91	246	3281	142	28	29
Jeżowe (2)	1812032	124	457	5868	142	28	29
Krzyszów (2)	1812042	62	188	2510	139	28	29
Nisko - miasto (4)	1812054	61	3463	10253	142	28	29
Nisko - obszar wiejski (5)	1812055	81	277	4233	142	28	29
Rudnik nad Sanem - miasto (4) *	1812064	37	1023	4244	142	28	29
Rudnik nad Sanem - obszar wiejski (5) *	1812065	42	53	2083	142	28	29
Ulanów - miasto (4)	1812074	8	549	972	139	28	29
Ulanów - obszar wiejski (5)	1812075	111	137	4381	139	28	29
Bircza (2)	1813012	254	491	3846	147	32	34
Dubiecko (2)	1813022	154	575	5447	147	32	34
Fredropol (2)	1813032	160	312	3265	147	32	34
Krasiczyn (2)	1813042	127	310	3047	147	32	34

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Krzywcza (2)	1813052	95	210	2934	147	32	34
Medyka (2)	1813062	61	1387	3813	147	32	34
Orły (2)	1813072	70	431	5167	147	32	34
Przemyśl (2)	1813082	108	409	5561	147	32	34
Stubno (2)	1813092	89	254	2402	147	32	34
Żurawica (2)	1813102	96	2267	7511	147	32	34
Przeworsk (1)	1814011	22	5648	10315	140	2	2
Adamówka (2)	1814022	134	267	2476	0	0	0
Gać (2)	1814032	36	202	2719	140	2	2
Jawornik Polski (2)	1814042	63	225	2725	147	2	2
Kańczuga - miasto (4)	1814054	8	1076	2176	140	2	2
Kańczuga - obszar wiejski (5)	1814055	97	180	5377	140	2	2
Przeworsk (2)	1814062	91	475	8677	140	2	2
Sieniawa - miasto (4)	1814074	7	361	1339	131	0	0
Sieniawa - obszar wiejski (5)	1814075	120	186	2704	146	0	0
Tryńcza (2)	1814082	70	1159	4740	140	2	2
Zarzecze (2)	1814092	49	601	4325	140	2	2
Iwierzycze (2)	1815012	66	589	4426	147	2	2
Ostrów (2)	1815022	96	424	4115	130	77	102
Ropczyce - miasto (4)	1815034	47	4061	9943	130	77	102
Ropczyce - obszar wiejski (5)	1815035	92	453	6454	130	77	102
Sędziszów Małopolski - miasto (4)	1815044	10	3093	4637	130	2	2
Sędziszów Małopolski - obszar wiejski (5)	1815045	144	527	9204	147	2	2
Wielopole Skrzyńskie (2)	1815052	93	674	4814	130	77	102
Dynów (1)	1816011	25	1019	3814	129	2	2
Błazowa - miasto (4)	1816024	4	450	1401	147	2	2
Błazowa - obszar wiejski (5)	1816025	109	242	4790	147	2	2
Boguchwała (2) *	1816032	9	2140	13070	147	2	2
Boguchwała (2) *	1816032	87	2140	13070	0	0	0
Chmielnik (2) *	1816042	53	585	3743	147	2	2
Dynów (2)	1816052	119	254	4131	129	2	2
Głogów Małopolski - miasto (4)	1816064	14	653	3306	147	2	2
Głogów Małopolski - obszar wiejski (5)	1816065	132	2214	7994	147	2	2
Hyzne (2)	1816072	51	323	4084	147	2	2
Kamień (2)	1816082	74	375	4042	142	28	29
Krasne (2)	1816092	39	1054	5941	147	2	2
Lubenia (2)	1816102	55	254	3887	147	2	2
Sokołów Małopolski - miasto (4)	1816114	16	747	2653	147	2	2
Sokołów Małopolski - obszar wiejski (5)	1816115	118	871	7262	147	2	2
Świlcza (2)	1816122	128	1516	11655	147	2	2
Trzebownisko (2)	1816132	90	2227	11619	147	2	2
Tyczyn - miasto (4)	1816144	10	773	2156	147	2	2
Tyczyn - obszar wiejski (5)	1816145	73	976	7915	147	2	2
Sanok (1)	1817011	38	14363	26267	141	16	17
Besko (2)	1817022	27	384	2664	141	16	17
Bukowsko (2)	1817032	137	381	3135	141	16	17
Komańcza (2)	1817042	456	1331	3452	146	7	7
Sanok (2)	1817052	232	523	10269	141	16	17
Tyrawa Wołoska (2)	1817062	69	148	1164	141	16	17
Zagórz - miasto (4)	1817074	22	1720	3248	141	16	17
Zagórz - obszar wiejski (5)	1817075	138	269	4857	141	16	17
Zarszyn (2)	1817082	105	503	5511	141	16	17
Stalowa Wola (1)	1818011	83	24425	44514	142	28	29
Bojanów (2)	1818022	179	262	4295	142	28	29
Pysznica (2)	1818032	147	342	5910	142	28	29
Radomyśl nad Sanem (2) *	1818042	134	435	4382	142	28	29
Zaklików (2)	1818052	202	808	5189	142	28	29
Zaleszany (2)	1818062	87	446	6460	142	28	29

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Czudec (2)	1819012	85	929	6848	147	2	2
Frysztak (2)	1819022	91	866	6336	146	2	2
Niebylec (2)	1819032	104	444	6247	147	2	2
Strzyżów - miasto (4)	1819044	14	2711	5761	147	2	2
Strzyżów - obszar wiejski (5)	1819045	127	453	7202	147	2	2
Wiśniowa (2)	1819052	83	418	5024	147	2	2
Baranów Sandomierski - miasto (4)	1820014	9	367	920	148	36	166
Baranów Sandomierski - obszar wiejski (5)	1820015	113	786	6396	148	36	166
Gorzycy (2)	1820022	69	2626	9041	188	28	29
Grębów (2)	1820032	187	661	5780	148	28	29
Nowa Dęba - miasto (4)	1820044	17	2930	7492	143	36	78
Nowa Dęba - obszar wiejski (5)	1820045	126	265	4349	143	36	166
Baligród (2) *	1821012	158	187	1967	145	16	118
Cisna (2) *	1821022	287	168	1108	144	0	0
Lesko - miasto (4) *	1821034	15	2039	3811	145	16	118
Lesko - obszar wiejski (5) *	1821035	96	297	3597	145	16	118
Olszanica (2) *	1821042	94	570	3146	145	16	118
Solina (2) *	1821052	185	536	3258	145	16	118
Krosno (1)	1861011	43	25399	31546	146	7	7
Przemyśl (1)	1862011	44	17354	43137	147	32	34
Rzeszów (1)	1863011	68	72741	109589	147	2	2
Tarnobrzeg (1)	1864011	85	11133	33261	148	36	166
Augustów (1)	2001011	81	6317	18957	152	127	226
Augustów (2)	2001022	267	508	3788	152	127	226
Bargłów Kościelny (2)	2001032	188	196	3331	152	127	226
Lipsk - miasto (4)	2001044	5	216	1595	152	127	226
Lipsk - obszar wiejski (5)	2001045	179	40	1647	152	127	226
Nowinka (2)	2001052	204	131	1687	152	127	226
Płaska (2)	2001062	372	172	1494	152	127	226
Sztabin (2)	2001072	363	186	3047	152	127	226
Choroszcz - miasto (4)	2002014	17	1689	3606	152	3	3
Choroszcz - obszar wiejski (5)	2002015	147	782	4745	152	3	3
Czarna Białostocka - miasto (4)	2002024	14	997	6200	152	3	3
Czarna Białostocka - obszar wiejski (5)	2002025	192	80	1220	152	3	3
Dobrzyniewo Duże (2) *	2002032	161	349	4949	152	3	3
Gródek (2)	2002042	429	837	3378	152	3	3
Juchnowiec Kościelny (2) *	2002052	172	2007	8496	152	3	3
Łapy - miasto (4)	2002064	12	3162	10951	152	3	3
Łapy - obszar wiejski (5)	2002065	116	346	3758	152	3	3
Michałow (2) *	2002072	409	736	4124	152	3	3
Poświętne (2)	2002082	115	113	2120	152	3	3
Supraśl - miasto (4)	2002094	6	541	2837	152	3	3
Supraśl - obszar wiejski (5)	2002095	183	1714	5321	152	3	3
Suraż - miasto (4)	2002104	34	65	595	152	3	3
Suraż - obszar wiejski (5)	2002105	43	15	593	152	3	3
Turośl Kościelna (2)	2002112	140	167	3205	152	3	3
Tykocin - miasto (4)	2002124	29	281	1179	152	3	3
Tykocin - obszar wiejski (5)	2002125	178	116	2642	152	3	3
Wasilków - miasto (4)	2002134	28	994	5919	152	3	3
Wasilków - obszar wiejski (5)	2002135	99	87	2408	152	3	3
Zabłudów - miasto (4)	2002144	14	278	1500	152	3	3
Zabłudów - obszar wiejski (5)	2002145	325	468	3538	152	3	3
Zawady (2)	2002152	112	122	1748	0	0	0
Bielsk Podlaski (1)	2003011	27	6613	17476	152	115	186
Brańsk (1)	2003021	32	612	2286	149	115	186
Bielsk Podlaski (2)	2003032	431	197	3403	152	115	186
Boćki (2)	2003042	232	143	2627	152	115	186
Brańsk (2)	2003052	227	220	3411	149	115	186

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Orla (2)	2003062	160	188	1531	152	115	186
Rudka (2)	2003072	70	190	1239	149	0	0
Wyszki (2)	2003082	206	172	2648	149	3	3
Grajewo (1)	2004011	19	4581	14869	150	105	171
Grajewo (2)	2004022	308	252	3518	150	105	171
Radziłów (2)	2004032	200	206	2971	152	104	170
Rajgród - miasto (4)	2004044	35	286	1101	150	105	171
Rajgród - obszar wiejski (5)	2004045	172	103	2252	150	105	171
Szczuczyn - miasto (4)	2004054	13	460	2250	150	105	171
Szczuczyn - obszar wiejski (5)	2004055	103	80	1814	150	105	171
Wąsosz (2)	2004062	118	165	2286	150	105	171
Hajnówka (1)	2005011	21	4821	14121	151	115	190
Białowieża (2)	2005022	203	646	1513	151	115	190
Czeremcha (2)	2005032	97	570	2095	152	3	3
Czyże (2)	2005042	135	81	1237	151	115	190
Dubicze Cerkiewne (2)	2005052	151	190	876	151	115	190
Hajnówka (2)	2005062	293	354	2317	151	115	190
Kleszczel - miasto (4)	2005074	47	222	843	152	115	190
Kleszczel - obszar wiejski (5)	2005075	96	14	649	151	115	190
Narew (2)	2005082	242	1425	2197	152	115	190
Narewka (2)	2005092	339	564	2265	151	115	190
Kolno (1)	2006011	25	2127	7042	153	104	184
Grabowo (2)	2006022	129	154	2116	153	104	170
Kolno (2)	2006032	281	227	4884	153	104	184
Mały Płock (2)	2006042	140	197	2932	153	104	170
Stawiski - miasto (4)	2006054	13	362	1538	152	104	170
Stawiski - obszar wiejski (5)	2006055	153	30	2387	153	104	170
Turośl (2)	2006062	199	216	2699	104	104	184
Jedwabne - miasto (4)	2007014	11	258	1237	152	104	170
Jedwabne - obszar wiejski (5)	2007015	148	36	2069	152	104	170
Łomża (2)	2007022	207	677	5848	152	104	170
Miastkowo (2)	2007032	115	322	2483	104	82	110
Nowogród - miasto (4)	2007044	21	202	1253	152	104	170
Nowogród - obszar wiejski (5)	2007045	80	32	1110	152	104	170
Piątnica (2)	2007052	220	1019	6181	152	104	170
Przytuły (2)	2007062	71	85	1318	152	104	170
Śniadowo (2)	2007072	163	353	3246	152	104	170
Wizna (2)	2007082	133	149	2484	152	104	170
Zbójna (2)	2007092	186	174	2452	104	104	184
Goniądz - miasto (4)	2008014	4	231	1110	152	3	3
Goniądz - obszar wiejski (5)	2008015	373	157	1848	150	3	3
Jasionówka (2)	2008022	97	101	1722	152	3	3
Jaświły (2)	2008032	175	276	3032	152	3	3
Knyszyn - miasto (4)	2008044	24	278	1834	152	3	3
Knyszyn - obszar wiejski (5)	2008045	103	115	1209	152	3	3
Krypno (2)	2008052	112	172	2287	152	3	3
Mońki - miasto (4)	2008064	8	1877	6794	152	3	3
Mońki - obszar wiejski (5)	2008065	154	162	2917	152	3	3
Trzcianne (2)	2008072	332	144	2653	152	3	3
Sejny (1)	2009011	4	1149	3742	152	116	108
Giby (2)	2009022	323	220	1705	152	116	108
Krasnopol (2)	2009032	172	158	2250	152	116	108
Puńsk (2)	2009042	139	302	2667	152	116	108
Sejny (2)	2009052	217	98	2406	152	116	108
Siemiatycze (1)	2010011	36	3445	9893	116	3	175
Drohiczyn - miasto (4)	2010024	16	242	1281	116	3	175
Drohiczyn - obszar wiejski (5)	2010025	192	91	2618	116	3	175
Dziadkowice (2)	2010032	116	154	1755	116	3	175
Grodzisk (2)	2010042	203	206	2525	116	3	175

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Mielnik (2)	2010052	196	325	1420	116	3	175
Milejczyce (2)	2010062	151	89	1079	116	3	175
Nurzec-Stacja (2)	2010072	215	235	2420	116	3	175
Perlejewo (2)	2010082	107	91	1728	0	0	0
Siemiatycze (2)	2010092	227	368	3548	116	3	175
Dąbrowa Białostocka - miasto (4)	2011014	23	924	3884	152	3	3
Dąbrowa Białostocka - obszar wiejski (5)	2011015	241	169	3477	152	3	3
Janów (2)	2011022	208	240	2488	152	3	3
Korycin (2)	2011032	117	251	2080	152	3	3
Krynki (2) *	2011042	166	252	1946	152	3	3
Kuźnica (2)	2011052	133	526	2452	152	3	3
Nowy Dwór (2)	2011062	121	92	1624	152	3	3
Sidra (2)	2011072	174	210	2203	152	3	3
Sokółka - miasto (4)	2011084	19	4905	12586	152	3	3
Sokółka - obszar wiejski (5)	2011085	295	147	4289	152	3	3
Suchowola - miasto (4) *	2011094	26	332	1321	152	3	3
Suchowola - obszar wiejski (5) *	2011095	230	85	2783	152	3	3
Szudziałowo (2)	2011102	302	226	1878	152	3	3
Bakałarzewo (2)	2012012	123	147	1772	198	116	108
Filipów (2)	2012022	151	218	2585	152	116	108
Jeleniewo (2)	2012032	131	110	1759	152	116	108
Przeośl (2)	2012042	124	128	1707	152	116	108
Raczki (2)	2012052	142	343	3562	152	116	108
Rutka-Tartak (2)	2012062	92	74	1291	152	116	108
Suwałki (2)	2012072	265	1211	3793	152	116	108
Szypłiszki (2)	2012082	156	263	2328	152	116	108
Wiżajny (2)	2012092	123	86	1585	152	116	108
Wysokie Mazowieckie (1)	2013011	15	2924	5995	152	3	224
Ciechanowiec - miasto (4)	2013024	26	951	3175	117	6	6
Ciechanowiec - obszar wiejski (5)	2013025	176	118	2518	117	6	6
Czyżew-Osada (2) *	2013032	130	1258	3905	152	3	3
Klukowo (2)	2013042	124	176	2604	152	6	6
Kobylin-Borzymy (2)	2013052	119	147	1949	152	3	3
Kulesze Kościelne (2)	2013062	115	99	1943	152	3	224
Nowe Piekuty (2)	2013072	110	173	2290	152	3	3
Sokoły (2)	2013082	156	321	3519	152	3	3
Szepietowo (2) *	2013092	152	720	4357	152	3	224
Wysokie Mazowieckie (2)	2013102	166	159	2882	152	3	224
Zambrów (1)	2014011	19	3730	14643	152	3	191
Kołaki Kościelne (2)	2014022	74	116	1421	152	3	191
Rutki (2)	2014032	200	230	3588	152	6	6
Szumowo (2)	2014042	141	281	2914	152	3	191
Zambrów (2)	2014052	299	534	5039	152	3	191
Białystok (1)	2061011	102	76469	196673	152	3	3
Łomża (1)	2062011	33	12747	42379	152	104	170
Suwałki (1)	2063011	66	16162	45371	152	116	108
Borzytuchom (2)	2201012	109	203	1715	154	34	36
Bytów - miasto (4)	2201024	9	5942	10998	154	34	36
Bytów - obszar wiejski (5)	2201025	188	427	4301	154	34	36
Czarna Dąbrówka (2)	2201032	298	389	3476	154	34	36
Kończygłowy (2)	2201042	174	550	2669	154	34	36
Lipnica (2)	2201052	309	549	2895	154	34	36
Miastko - miasto (4)	2201064	6	2163	7237	163	34	36
Miastko - obszar wiejski (5)	2201065	459	1156	5509	154	34	36
Parchowo (2)	2201072	131	208	2059	154	34	36
Studzienice (2)	2201082	176	371	2153	154	34	36
Trzebielino (2)	2201092	226	370	2368	154	86	123
Tuchomie (2)	2201102	107	331	2491	154	34	36
Chojnice (1)	2202011	21	13203	25761	156	52	93

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Brusy - miasto (4)	2202024	5	899	2805	156	52	93
Brusy - obszar wiejski (5)	2202025	395	304	5023	156	52	93
Chojnice (2)	2202032	458	1301	9931	156	52	93
Czersk - miasto (4)	2202044	10	2092	5950	155	52	93
Czersk - obszar wiejski (5)	2202045	370	904	6733	155	52	93
Konarzyny (2)	2202052	105	217	1302	156	52	93
Człuchów (1)	2203011	13	4352	9862	156	52	156
Czarne - miasto (4)	2203024	46	1145	3924	233	64	148
Czarne - obszar wiejski (5)	2203025	189	208	2124	233	64	148
Człuchów (2)	2203032	361	1258	6387	156	52	156
Debrzno - miasto (4)	2203044	8	504	3486	156	52	156
Debrzno - obszar wiejski (5)	2203045	216	188	2455	156	52	156
Koczała (2)	2203052	223	432	2268	158	52	156
Przechlewo (2)	2203062	244	1350	3958	156	52	156
Rzeczynica (2)	2203072	275	480	2344	159	97	157
Pruszcz Gdański (1)	2204011	16	6230	16180	163	69	90
Cedry Wielkie (2)	2204022	124	455	3989	163	69	90
Kolbudy (2) *	2204032	83	1312	8255	163	69	90
Pruszcz Gdański (2)	2204042	144	4115	11307	163	69	90
Przywidz (2)	2204052	129	247	3299	163	69	90
Pszczółki (2)	2204062	50	649	5173	163	69	90
Suchy Dąb (2)	2204072	85	232	2471	163	69	90
Trąbki Wielkie (2)	2204082	163	650	6002	163	69	90
Chmielno (2)	2205012	79	440	3877	163	69	90
Kartuzy - miasto (4)	2205024	7	3863	9757	163	69	90
Kartuzy - obszar wiejski (5)	2205025	200	1637	9656	163	69	90
Przodkowo (2)	2205032	85	604	4189	163	69	90
Sierakowice (2)	2205042	183	1624	9859	163	69	90
Somonino (2)	2205052	112	864	5464	163	69	90
Stężycza (2)	2205062	160	832	5263	163	69	90
Sulęcyno (2)	2205072	131	230	2859	163	69	90
Żukowo - miasto (4)	2205084	5	1796	4033	163	69	90
Żukowo - obszar wiejski (5)	2205085	159	2787	12120	163	69	90
Kościerzyna (1)	2206011	16	5787	14728	164	128	228
Dziemiany (2)	2206022	125	605	2504	164	128	228
Karsin (2)	2206032	169	728	3489	161	52	93
Kościerzyna (2)	2206042	311	2879	8327	164	128	228
Liniewo (2)	2206052	110	232	2829	164	69	90
Lipusz (2)	2206062	109	285	2048	164	128	228
Nowa Karczma (2)	2206072	113	335	3805	164	69	90
Stara Kiszewa (2)	2206082	213	335	3896	164	128	228
Kwidzyn (1)	2207011	22	16854	25180	157	9	9
Gardeja (2)	2207022	193	380	5055	157	9	9
Kwidzyn (2)	2207032	207	892	6615	157	9	9
Prabuty - miasto (4)	2207044	5	1116	5555	157	9	9
Prabuty - obszar wiejski (5)	2207045	192	185	2765	157	9	9
Ryjewo (2)	2207052	104	510	3746	157	9	9
Sadlinki (2)	2207062	112	191	3502	157	9	9
Lębork (1)	2208011	18	7714	22522	158	97	157
Łeba (1)	2208021	15	1029	2492	159	97	157
Cewice (2)	2208032	188	500	4348	158	97	157
Nowa Wieś Lęborska (2)	2208042	270	987	7880	158	97	157
Wicko (2)	2208052	215	877	3499	159	97	157
Malbork (1)	2209011	17	8113	25185	163	69	204
Lichnowy (2)	2209032	89	328	2813	163	69	90
Malbork (2)	2209042	101	294	2610	163	69	204
Miłoradz (2)	2209062	94	230	2152	163	69	204
Nowy Staw - miasto (4)	2209074	5	514	2808	163	69	204
Nowy Staw - obszar wiejski (5)	2209075	109	397	2152	163	69	204

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Stare Pole (2)	2209082	79	437	2976	193	69	204
Krynica Morska (1)	2210011	116	166	932	160	0	0
Nowy Dwór Gdański - miasto (4)	2210024	5	1874	6437	193	69	90
Nowy Dwór Gdański - obszar wiejski (5)	2210025	208	600	4998	193	69	90
Ostaszewo (2)	2210032	61	233	1992	193	69	90
Stegna (2)	2210042	170	663	6210	160	69	90
Sztutowo (2)	2210052	112	311	2332	160	69	90
Hel (1)	2211011	22	476	2684	161	97	151
Jastarnia (1)	2211021	8	594	2618	161	97	151
Puck (1)	2211031	5	2821	7285	164	97	151
Władysławowo (1)	2211041	39	2348	9766	164	97	151
Kosakowo (2)	2211052	50	1169	5513	164	97	151
Krokowa (2)	2211062	211	1573	6301	34	97	151
Puck (2)	2211072	242	2211	13708	164	97	151
Ustka (1)	2212011	10	2776	11000	163	86	123
Damnica (2)	2212022	168	697	3985	163	86	123
Dębica Kaszubska (2)	2212032	300	748	6027	163	86	123
Główczyce (2)	2212042	322	488	5991	159	86	123
Kępice - miasto (4)	2212054	6	466	2518	163	86	123
Kępice - obszar wiejski (5)	2212055	287	669	3642	163	86	123
Kobylnica (2)	2212062	244	2069	6445	163	86	123
Potęgowo (2)	2212072	228	744	4498	163	86	123
Słupsk (2)	2212082	262	3518	8989	163	86	123
Smoldzino (2)	2212092	260	306	2234	163	86	123
Ustka (2)	2212102	217	1705	4780	163	86	123
Czarna Woda (1)	2213011	28	533	1962	155	52	93
Skórcz (1)	2213021	4	1008	2306	162	44	48
Starogard Gdański (1)	2213031	25	11944	30865	162	44	48
Bobowo (2)	2213042	52	91	1812	162	44	48
Kaliska (2)	2213052	111	386	3171	162	44	48
Lubichowo (2)	2213062	161	250	3427	162	44	48
Osieczna (2)	2213072	123	227	1660	32	52	93
Osiek (2)	2213082	156	140	1507	162	44	48
Skarszewy - miasto (4)	2213094	11	1196	4375	162	44	48
Skarszewy - obszar wiejski (5)	2213095	158	676	4418	162	44	48
Skórcz (2)	2213102	97	297	2819	162	44	48
Smętowo Graniczne (2)	2213112	86	356	3337	29	69	90
Starogard Gdański (2)	2213122	195	1956	8773	162	44	48
Zblewo (2)	2213132	138	895	6800	162	44	48
Tczew (1)	2214011	22	14959	39322	163	69	90
Gniew - miasto (4)	2214024	6	1499	4368	163	69	90
Gniew - obszar wiejski (5)	2214025	188	373	5484	29	69	90
Morzeszczyn (2)	2214032	91	318	2384	162	69	90
Pelplin - miasto (4)	2214044	4	1301	5531	162	69	90
Pelplin - obszar wiejski (5)	2214045	137	727	5125	162	69	90
Subkowy (2)	2214052	78	314	3234	162	69	90
Tczew (2)	2214062	171	1971	7141	162	69	90
Reda (1)	2215011	30	2394	12778	164	97	151
Rumia (1)	2215021	30	4616	29688	164	97	151
Wejherowo (1)	2215031	26	8374	28892	164	97	151
Choczewo (2)	2215042	183	378	3473	34	97	151
Gniewino (2)	2215052	176	1124	4249	34	97	151
Linia (2)	2215062	120	261	3369	164	97	151
Luzino (2)	2215072	111	899	7865	164	97	151
Łęczyce (2)	2215082	233	850	7045	158	97	151
Szemud (2)	2215092	177	1406	7806	164	97	151
Wejherowo (2)	2215102	194	3285	11807	164	97	151
Dzierzgoń - miasto (4) *	2216014	4	1066	3611	193	111	196
Dzierzgoń - obszar wiejski (5) *	2216015	127	95	2437	193	111	196

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Mikołajki Pomorskie (2) *	2216022	92	200	2349	193	69	196
Stary Dzierżoń (2) *	2216032	186	187	2508	196	111	169
Stary Targ (2) *	2216042	141	484	4221	193	69	196
Sztum - miasto (4) *	2216054	5	2482	6720	193	69	196
Sztum - obszar wiejski (5) *	2216055	176	785	5194	193	69	196
Gdańsk (1)	2261011	266	135199	300695	163	69	90
Gdynia (1)	2262011	135	68810	164715	164	97	151
Słupsk (1)	2263011	43	27236	65500	163	86	123
Sopot (1)	2264011	17	10586	25146	164	69	90
Będzin (1)	2401011	37	13067	39246	179	110	181
Czeladź (1)	2401021	16	5485	22882	179	110	181
Wojkowice (1)	2401031	13	987	6036	179	110	181
Bobrowniki (2)	2401042	51	1011	7255	179	110	181
Mierzęcice (2)	2401052	49	564	4761	179	110	181
Psary (2)	2401062	46	482	7214	179	110	181
Siewierz - miasto (4)	2401074	39	2461	3552	179	110	181
Siewierz - obszar wiejski (5)	2401075	76	453	4267	179	110	181
Sławków (1) *	2401081	37	1022	4556	179	110	181
Szczyrk (1)	2402011	39	1021	3705	175	27	28
Bestwina (2)	2402022	38	898	6687	175	27	28
Buczkowice (2)	2402032	19	1175	6718	175	27	28
Czechowice-Dziedzice - miasto (4)	2402044	33	10792	23071	175	27	28
Czechowice-Dziedzice - obszar wiejski (5)	2402045	34	688	5359	175	27	28
Jasienica (2)	2402052	93	3151	13117	175	27	28
Jaworze (2)	2402062	21	1132	4230	175	27	28
Kozy (2)	2402072	27	1252	7388	175	27	28
Porąbka (2)	2402082	64	1759	9241	175	27	28
Wilamowice - miasto (4)	2402094	10	405	1729	175	27	28
Wilamowice - obszar wiejski (5)	2402095	47	1091	7646	175	27	28
Wilkowice (2)	2402102	34	1239	7807	175	27	28
Cieszyn (1)	2403011	29	10897	23218	165	133	160
Ustroń (1)	2403021	59	4868	9777	165	133	238
Wisła (1)	2403031	109	2081	7063	165	133	238
Brenna (2)	2403042	95	894	6374	165	133	238
Chybie (2)	2403052	32	1020	5813	165	133	130
Dębowiec (2)	2403062	43	555	3413	165	133	160
Goeszów (2)	2403072	66	1100	7640	165	133	160
Hażlach (2) *	2403082	49	602	6356	165	133	160
Istebna (2)	2403092	84	936	6863	165	133	238
Skoczów - miasto (4)	2403104	10	5113	9740	165	133	238
Skoczów - obszar wiejski (5)	2403105	54	998	6805	165	133	238
Strumień - miasto (4)	2403114	6	1076	2235	165	133	130
Strumień - obszar wiejski (5)	2403115	53	355	5495	165	133	130
Zebrzydowice (2)	2403122	41	885	7981	165	133	130
Błachownia - miasto (4)	2404014	37	1308	6233	176	39	43
Błachownia - obszar wiejski (5)	2404015	30	403	2158	176	39	43
Dąbrowa Zielona (2)	2404022	100	286	2475	176	39	43
Janów (2)	2404032	147	721	3712	176	39	43
Kamienica Polska (2)	2404042	46	619	3630	176	39	43
Kłomnice (2)	2404052	147	940	8550	176	39	43
Koniecpol - miasto (4)	2404064	43	1419	4032	166	39	43
Koniecpol - obszar wiejski (5)	2404065	103	69	2321	166	39	43
Konopiska (2)	2404072	79	1039	6664	176	39	43
Kruszyna (2)	2404082	94	255	3003	176	39	43
Lelów (2)	2404092	124	400	3061	166	39	43
Mstów (2)	2404102	119	627	6527	176	39	43
Mykanów (2)	2404112	142	1383	8814	176	39	43
Olsztyn (2)	2404122	109	1001	4248	176	39	43

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Poczesna (2)	2404132	60	2088	7937	176	39	43
Przyrów (2)	2404142	81	289	2491	176	39	43
Rędziny (2)	2404152	41	1547	6122	176	39	43
Starcza (2)	2404162	20	122	1751	176	39	43
Knurów (1)	2405011	34	11358	26860	177	98	152
Pyskowice (1)	2405021	31	2713	12265	177	98	152
Gierałtowiec (2)	2405032	38	595	6784	177	110	181
Pilchowice (2)	2405042	70	889	6629	177	110	181
Rudziniec (2)	2405052	159	688	6742	177	98	152
Sośnicowice - miasto (4) *	2405064	12	687	1159	124	98	152
Sośnicowice - obszar wiejski (5) *	2405065	104	461	4267	124	98	152
Toszek - miasto (4)	2405074	10	907	2492	177	98	152
Toszek - obszar wiejski (5)	2405075	90	252	3957	177	98	152
Wielowieś (2)	2405082	116	376	3682	177	98	152
Kłobuck - miasto (4)	2406014	47	4346	8513	167	39	43
Kłobuck - obszar wiejski (5)	2406015	83	1026	4552	167	39	43
Krzepice - miasto (4)	2406024	28	1413	2900	168	39	43
Krzepice - obszar wiejski (5)	2406025	51	129	2908	168	39	43
Lipie (2)	2406032	99	1418	4005	167	39	43
Miedźno (2)	2406042	113	264	4710	167	39	43
Opatów (2)	2406052	73	332	4097	168	39	43
Panki (2)	2406062	55	1679	3090	168	39	43
Popów (2)	2406072	102	983	3752	167	39	43
Przystajń (2)	2406082	89	293	3697	168	113	77
Wręczyca Wielka (2)	2406092	149	1607	10648	167	39	43
Lubliniec (1)	2407011	89	7243	15702	169	113	154
Boronów (2)	2407022	57	286	2126	169	98	219
Ciasna (2)	2407032	134	1060	4903	169	113	154
Herby (2)	2407042	86	1551	4585	169	113	154
Kochanowice (2)	2407052	80	459	4180	169	113	154
Koszęcin (2)	2407062	129	909	7310	169	113	154
Pawonków (2)	2407072	119	326	4087	169	113	154
Woźniki - miasto (4)	2407084	71	1001	2838	169	110	181
Woźniki - obszar wiejski (5)	2407085	57	418	3289	169	110	181
Łaziska Górne (1)	2408011	20	5944	14515	179	110	181
Mikołów (1)	2408021	79	10952	25289	179	110	181
Orzesze (1)	2408031	84	2197	11909	170	110	181
Ormontowice (2)	2408042	15	4310	3588	170	110	181
Wiry (2)	2408052	35	667	4060	179	110	181
Myszków (1)	2409011	74	6123	21866	179	39	181
Koziegłowy - miasto (4)	2409024	27	523	1472	179	39	181
Koziegłowy - obszar wiejski (5)	2409025	132	890	7330	179	39	181
Niegowa (2)	2409032	88	269	3507	179	39	181
Poraj (2)	2409042	57	1602	6596	176	39	43
Żarki - miasto (4)	2409054	25	925	2809	179	39	181
Żarki - obszar wiejski (5)	2409055	76	92	2323	179	39	181
Goczałkowice-Zdrój (2)	2410012	47	995	3885	175	110	130
Kobiór (2)	2410022	48	371	2992	179	110	181
Miedzna (2)	2410032	50	2143	10578	175	110	181
Pawłowice (2)	2410042	76	8144	11938	178	133	130
Pszczyna - miasto (4)	2410054	22	8081	16695	175	110	130
Pszczyna - obszar wiejski (5)	2410055	153	1878	15541	175	110	130
Suszec (2)	2410062	75	3823	6812	178	110	130
Racibórz (1)	2411011	75	15644	37777	171	98	54
Kornowac (2)	2411022	26	206	3039	171	110	181
Krzanowice - miasto (4) *	2411034	3	272	1397	171	98	54
Krzanowice - obszar wiejski (5) *	2411035	44	118	2475	171	98	54
Krzyżanowice (2)	2411042	70	958	7309	171	98	54
Kuźnia Raciborska - miasto (4)	2411054	31	1476	3675	172	98	54

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	pojęcie taksonomiczne - słabe kryterium grupowania	pojęcie taksonomiczne - mocne kryterium grupowania
Kuźnia Raciborska - obszar wiejski (5)	2411055	96	318	4220	172	98	54
Nędza (2)	2411062	57	299	4623	172	98	54
Pietrowice Wielkie (2)	2411072	68	1032	4545	171	98	54
Rudnik (2)	2411082	74	214	3197	171	98	54
Czerwionka-Leszczyny - miasto (4)	2412014	38	3440	18372	170	110	181
Czerwionka-Leszczyny - obszar wiejski (5)	2412015	77	866	8106	170	110	181
Gaszowice (2)	2412022	20	477	5591	178	110	181
Jejkowice (2)	2412032	8	430	2385	178	110	181
Lyski (2)	2412042	57	974	5665	171	110	181
Świerklany (2)	2412052	24	1284	7179	178	110	130
Kalety (1)	2413011	76	499	5604	169	98	219
Miasteczko Śląskie (1)	2413021	68	1765	4715	179	98	219
Radzionków (1) *	2413031	13	3248	11264	179	98	152
Tarnowskie Góry (1)	2413041	84	17193	40230	179	98	219
Krupski Młyn (2)	2413052	39	1300	2197	179	98	219
Ożarówice (2) *	2413062	46	1315	3382	179	98	219
Świerklaniec (2)	2413072	45	1645	7117	179	98	219
Tworóg (2)	2413082	125	920	5145	179	98	219
Zbrosławice (2)	2413092	148	1651	10009	177	98	152
Bieruń (1)	2414011	41	10421	13253	179	110	181
Imielin (1)	2414021	28	751	5066	179	110	181
Lędziny (1)	2414031	31	6417	10454	179	110	181
Bojszowy (2)	2414042	35	665	4233	179	110	181
Chełm Śląski (2)	2414052	23	573	3625	179	110	181
Pszów (1)	2415011	20	3949	8907	178	110	181
Radlin (1) *	2415021	13	5535	11374	178	110	181
Rydułtowy (1)	2415031	15	5455	14210	178	110	181
Wodzisław Śląski (1)	2415041	50	8712	31639	178	110	181
Godów (2)	2415052	38	804	8031	178	110	130
Gorzycy (2)	2415062	64	2205	12441	178	110	181
Lubomia (2)	2415072	42	833	5151	171	110	181
Markłowice (2)	2415082	14	350	3142	178	110	181
Mszana (2)	2415092	31	420	4433	178	110	130
Poręba (1)	2416011	40	1349	5665	179	110	181
Zawiercie (1)	2416021	85	13702	35063	179	110	181
Irządze (2)	2416032	71	162	1637	179	110	181
Kroczyce (2)	2416042	111	392	3879	179	110	181
Łazy - miasto (4)	2416054	9	1445	4697	179	110	181
Łazy - obszar wiejski (5)	2416055	124	316	5617	179	110	181
Ogrodzieniec - miasto (4)	2416064	29	430	2818	179	110	181
Ogrodzieniec - obszar wiejski (5)	2416065	56	307	3259	179	110	181
Pilica - miasto (4)	2416074	8	610	1185	179	110	181
Pilica - obszar wiejski (5)	2416075	135	769	4301	179	110	181
Szczekociny - miasto (4)	2416084	18	1091	2572	173	110	218
Szczekociny - obszar wiejski (5)	2416085	116	137	2616	173	110	218
Włodowice (2)	2416092	76	246	3422	179	110	181
Żarnowiec (2)	2416102	125	202	2985	89	110	181
Żywiec (1)	2417011	51	14817	20520	174	73	96
Czernichów (2)	2417022	56	636	4086	174	27	28
Gilowice (2)	2417032	28	395	3470	174	73	96
Jeleśnia (2)	2417042	170	3202	8213	174	73	96
Koszarawa (2)	2417052	32	100	1507	174	73	96
Lipowa (2)	2417062	59	373	6046	174	73	96
Łękawica (2)	2417072	43	244	2654	174	73	96
Łodygowice (2)	2417082	36	887	8325	174	27	28
Milówka (2)	2417092	99	563	6038	174	73	96
Radziechowy-Wieprz (2)	2417102	65	559	7626	174	73	96
Rajcza (2)	2417112	131	835	5460	174	73	96

nazwa gminy	gmina				numer lokalnego rynku pracy		
	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Ślemień (2)	2417122	45	262	2118	174	73	96
Świnna (2)	2417132	39	448	4809	174	73	96
Ujsoły (2)	2417142	110	308	2773	174	73	96
Węgierska Górka (2)	2417152	76	1778	9162	174	73	96
Bielsko-Biała (1)	2461011	125	61359	115599	175	27	28
Bytom (1)	2462011	69	33115	121356	179	98	152
Chorzów (1)	2463011	33	25281	72303	179	110	181
Częstochowa (1)	2464011	160	74642	161671	176	39	43
Dąbrowa Górnicza (1)	2465011	189	39612	89928	179	110	181
Gliwice (1)	2466011	134	66830	132612	177	98	152
Jastrzębie-Zdrój (1)	2467011	85	30768	62959	178	110	130
Jaworzno (1)	2468011	153	22124	62970	179	110	179
Katowice (1)	2469011	165	148725	204369	179	110	181
Mysłowice (1)	2470011	66	19977	50199	179	110	181
Piekary Śląskie (1)	2471011	40	12062	38881	179	110	181
Ruda Śląska (1)	2472011	78	36503	95606	179	110	181
Rybnik (1)	2473011	148	39040	93904	178	110	181
Siemianowice Śląskie (1)	2474011	25	12148	47879	179	110	181
Sosnowiec (1)	2475011	91	48133	153692	179	110	181
Świętochłowice (1)	2476011	13	8270	36309	179	110	181
Tychy (1)	2477011	82	39697	90460	179	110	181
Zabrze (1)	2478011	80	39167	125513	177	98	152
Żory (1)	2479011	65	10487	44973	178	110	130
Busko-Zdrój - miasto (4)	2601014	12	5715	11425	180	54	69
Busko-Zdrój - obszar wiejski (5)	2601015	223	982	9275	180	54	69
Gnojno (2)	2601022	96	279	2758	180	8	8
Nowy Korczyn (2)	2601032	117	393	3717	180	54	69
Pacanów (2)	2601042	125	491	4530	180	54	69
Solec-Zdrój (2)	2601052	84	387	2942	180	54	69
Stopnica (2)	2601062	127	334	4654	180	54	69
Tuczepy (2)	2601072	84	624	2261	79	11	11
Wiślica (2)	2601082	100	305	3333	180	54	69
Imielno (2)	2602012	101	208	2659	181	8	134
Jędrzejów - miasto (4)	2602024	11	4543	11123	181	8	134
Jędrzejów - obszar wiejski (5)	2602025	216	629	7687	181	8	134
Małogoszcz - miasto (4) *	2602034	10	1182	2747	191	8	8
Małogoszcz - obszar wiejski (5) *	2602035	136	253	4663	191	8	8
Nagłowice (2)	2602042	117	242	3084	181	8	134
Oksa (2)	2602052	91	235	2918	181	8	59
Sędziszów - miasto (4)	2602064	8	2520	4724	181	8	222
Sędziszów - obszar wiejski (5)	2602065	138	300	3670	181	8	222
Słupia (Jędrzejowska) (2)	2602072	108	203	2678	181	8	222
Sobków (2)	2602082	144	424	5116	184	8	8
Wodzisław (2)	2602092	177	482	4385	181	8	134
Bejsce (2)	2603012	57	251	2484	182	13	14
Czarnocin (2)	2603022	70	163	2389	182	13	14
Kazimierza Wielka - miasto (4)	2603034	5	1518	3751	182	13	14
Kazimierza Wielka - obszar wiejski (5)	2603035	135	374	6629	182	13	14
Opatowiec (2)	2603042	69	188	2146	182	13	14
Skalbmierz - miasto (4)	2603054	7	308	834	183	13	14
Skalbmierz - obszar wiejski (5)	2603055	79	63	3263	183	36	39
Bieliny (2)	2604012	88	389	6062	192	8	8
Bodzentyn - miasto (4)	2604024	9	392	1434	189	8	8
Bodzentyn - obszar wiejski (5)	2604025	151	215	5603	189	8	8
Chęciny - miasto (4)	2604034	14	1068	2904	184	8	8
Chęciny - obszar wiejski (5)	2604035	113	585	6539	184	8	8
Chmielnik - miasto (4)	2604044	8	872	2554	180	8	8
Chmielnik - obszar wiejski (5)	2604045	134	322	4528	180	8	8
Daleszyce (2) *	2604054	16	944	9271	192	8	8

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Daleszyce (2) *	2604055	205	944	9271	192	8	8
Górno (2)	2604062	83	659	8196	192	8	8
Łagów (2)	2604072	113	309	4030	192	8	8
Łopuszno (2)	2604082	177	688	5387	192	8	8
Masłów (2)	2604092	86	454	6183	192	8	8
Miedziana Góra (2)	2604102	71	750	6541	192	8	8
Mniów (2)	2604112	95	346	5669	192	8	8
Morawica (2)	2604122	140	2008	8589	192	8	8
Nowa Słupia (2)	2604132	86	881	5836	190	8	8
Piekoszków (2)	2604142	103	1587	9675	192	8	8
Pierzchnica (2)	2604152	105	291	2775	192	8	8
Raków (2)	2604162	191	256	3284	79	8	8
Sitkówka-Nowiny (2)	2604172	46	2529	4558	184	8	8
Strawczyn (2)	2604182	86	570	6101	192	8	8
Zagnańsk (2)	2604192	125	1171	8002	192	8	8
Fałków (2)	2605012	132	275	2847	185	25	26
Gowarczów (2)	2605022	101	212	3078	185	25	26
Końskie - miasto (4)	2605034	18	7527	13752	185	25	26
Końskie - obszar wiejski (5)	2605035	232	1531	9500	185	25	26
Radoszyce (2)	2605042	147	534	5486	185	25	26
Ruda Maleniecka (2)	2605052	110	273	1931	185	25	26
Słupia (Konecka) (2)	2605062	106	209	2040	185	8	8
Smyków (2)	2605072	62	288	2226	185	25	26
Stąporków - miasto (4)	2605084	11	1046	4113	185	25	26
Stąporków - obszar wiejski (5)	2605085	221	935	7304	185	25	26
Bačkowie (2)	2606012	96	271	3039	188	60	145
Iwaniska (2)	2606022	105	346	4151	188	60	145
Lipnik (2)	2606032	81	214	3377	188	36	39
Opatów - miasto (4)	2606044	9	2544	4429	188	60	145
Opatów - obszar wiejski (5)	2606045	105	317	3434	188	60	145
Ożarów - miasto (4)	2606054	8	853	3483	186	60	76
Ożarów - obszar wiejski (5)	2606055	176	978	3931	186	60	76
Sadowie (2)	2606062	82	282	2560	186	60	145
Tarłów (2)	2606072	163	304	3284	186	60	76
Wojciechowice (2)	2606082	86	439	2683	186	60	145
Ostrowiec Świętokrzyski (1)	2607011	46	17250	48607	186	60	76
Bałtów (2)	2607022	105	192	2312	186	60	76
Bodzechów (2)	2607032	122	1012	8195	186	60	76
Ćmielów - miasto (4)	2607044	13	811	1958	186	60	76
Ćmielów - obszar wiejski (5)	2607045	105	98	2698	186	60	76
Kunów - miasto (4)	2607054	7	505	1949	186	60	76
Kunów - obszar wiejski (5)	2607055	107	177	4155	186	60	76
Waśniów (2)	2607062	112	202	4253	186	60	76
Działoszyce - miasto (4)	2608014	2	165	653	187	54	199
Działoszyce - obszar wiejski (5)	2608015	104	40	2545	187	54	199
Kije (2)	2608022	100	367	2698	180	8	8
Michałów (2)	2608032	112	301	2764	180	54	199
Pińczów - miasto (4)	2608044	14	3424	7911	180	54	199
Pińczów - obszar wiejski (5)	2608045	199	752	5970	180	54	199
Złota (2)	2608052	82	184	2896	180	54	199
Sandomierz (1)	2609011	29	8201	15962	188	36	39
Dwikozy (2)	2609022	85	628	5566	188	36	39
Klimontów (2)	2609032	99	517	5045	188	36	39
Koprzywnica - miasto (4) *	2609044	18	268	1545	148	36	166
Koprzywnica - obszar wiejski (5) *	2609045	51	99	2636	148	36	166
Łonów (2)	2609052	87	516	4536	148	36	166
Obrazów (2)	2609062	72	225	4042	188	36	39
Samborzec (2)	2609072	85	309	5449	188	36	39
Wilczyce (2)	2609082	70	126	2334	188	36	39

nazwa gminy	gmina				numer lokalnego rynku pracy		
	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Zawichost - miasto (4)	2609094	20	243	1121	188	36	39
Zawichost - obszar wiejski (5)	2609095	60	154	1653	188	36	39
Skarżysko-Kamienna (1)	2610011	64	11751	31811	189	37	33
Bliżyn (2)	2610022	141	340	5287	189	37	33
Łączna (2)	2610032	62	419	3196	189	8	8
Skarżysko Kościelne (2)	2610042	53	285	3729	189	37	33
Suchedniów - miasto (4)	2610054	59	1086	5605	189	37	33
Suchedniów - obszar wiejski (5)	2610055	16	290	1137	189	37	33
Starachowice (1)	2611011	32	15148	33810	190	37	41
Brody (2)	2611022	161	1277	6598	190	37	41
Mirzec (2)	2611032	111	403	5072	190	37	41
Pawłów (2)	2611042	137	722	8819	190	37	41
Wąchock - miasto (4)	2611054	16	378	1709	190	37	41
Wąchock - obszar wiejski (5)	2611055	66	274	2541	189	37	33
Bogoria (2)	2612012	123	397	4615	79	11	11
Łubnice (2)	2612022	84	173	2515	79	11	11
Oleśnica (2)	2612032	53	136	2328	79	11	11
Osiek - miasto (4)	2612044	17	365	1210	79	11	11
Osiek - obszar wiejski (5)	2612045	112	562	3488	79	11	11
Połaniec - miasto (4)	2612054	17	1386	5930	79	11	11
Połaniec - obszar wiejski (5)	2612055	58	2031	2151	79	11	11
Rytwiany (2)	2612062	126	524	3780	79	11	11
Staszów - miasto (4)	2612074	20	5001	10263	79	11	11
Staszów - obszar wiejski (5)	2612075	207	1756	6547	79	11	11
Szydłów (2)	2612082	108	672	2940	79	11	11
Kluczewsko (2)	2613012	137	235	2984	191	8	59
Krasocin (2)	2613022	193	753	6481	191	8	59
Moskorzew (2)	2613032	73	141	1737	173	110	218
Radków (2)	2613042	88	111	1517	191	8	59
Secemin (2)	2613052	163	341	2972	166	8	59
Włoszczowa - miasto (4)	2613064	30	4813	7063	191	8	59
Włoszczowa - obszar wiejski (5)	2613065	224	526	5662	191	8	59
Kielce (1)	2661011	110	67543	139166	192	8	8
Bartoszyce (1)	2801011	11	5539	16879	199	43	143
Górowo Iławeckie (1)	2801021	3	778	2894	199	43	143
Bartoszyce (2)	2801032	428	827	6786	199	43	143
Biszynek - miasto (4)	2801044	2	418	1651	195	43	47
Biszynek - obszar wiejski (5)	2801045	201	123	2591	195	43	47
Górowo Iławeckie (2)	2801052	416	534	4487	199	43	143
Sępól - miasto (4)	2801064	5	114	1319	199	43	143
Sępól - obszar wiejski (5)	2801065	241	242	2824	199	43	143
Braniewo (1)	2802011	12	3179	11736	193	90	127
Braniewo (2)	2802022	307	425	3917	193	90	127
Frombork - miasto (4)	2802034	8	509	1622	193	90	127
Frombork - obszar wiejski (5)	2802035	116	7	809	193	90	127
Lelkowo (2)	2802042	198	185	1838	193	90	127
Pieniężno - miasto (4)	2802054	4	381	1993	195	43	47
Pieniężno - obszar wiejski (5)	2802055	239	98	2453	195	43	47
Płoskinia (2)	2802062	170	264	1636	193	90	127
Wilczęta (2)	2802072	148	172	1901	193	90	127
Działdowo (1)	2803011	11	5796	13908	102	120	205
Działdowo (2)	2803022	273	1152	5799	102	120	205
Howo-Osada (2)	2803032	104	789	4401	102	120	68
Lidzbark - miasto (4)	2803044	6	1764	5370	102	120	205
Lidzbark - obszar wiejski (5)	2803045	249	459	3612	102	120	205
Płońnica (2)	2803052	163	330	3402	102	120	205
Rybno (2)	2803062	148	866	4402	102	111	52
Elbląg (2)	2804012	192	2124	4166	163	90	127
Godkowo (2)	2804022	167	188	2003	193	90	127

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Gronowo Elbląskie (2)	2804032	89	415	3203	193	90	127
Markusy (2)	2804042	110	226	2499	193	90	127
Milejewo (2)	2804052	96	440	1942	193	90	127
Młynary - miasto (4)	2804064	3	393	1116	193	90	127
Młynary - obszar wiejski (5)	2804065	155	141	1663	193	90	127
Pasłęk - miasto (4)	2804074	11	2101	7939	193	90	127
Pasłęk - obszar wiejski (5)	2804075	253	264	4313	193	90	127
Rychliki (2)	2804082	132	233	2474	193	90	127
Tolkmicko - miasto (4)	2804094	2	374	1730	193	90	127
Tolkmicko - obszar wiejski (5)	2804095	206	252	2558	193	90	127
Elk (1)	2805011	21	13899	37231	198	105	197
Elk (2)	2805022	380	889	6302	198	105	197
Kalinowo (2)	2805032	284	305	4255	198	105	197
Prostki (2)	2805042	231	464	4594	150	105	197
Stare Juchy (2)	2805052	197	261	2362	198	105	197
Giżycko (1)	2806011	14	5454	19125	199	124	215
Giżycko (2)	2806042	289	1541	5018	199	124	215
Kruklanki (2)	2806052	202	268	1932	199	124	215
Miłki (2)	2806062	170	446	2413	199	124	215
Ryn - miasto (4)	2806084	4	552	1933	199	124	215
Ryn - obszar wiejski (5)	2806085	208	115	1846	199	124	215
Wydminy (2)	2806102	233	356	4032	199	124	215
Ława (1)	2807011	22	10066	21384	194	111	169
Lubawa (1)	2807021	17	4348	6024	195	111	52
Ława (2)	2807032	424	1225	7316	194	111	169
Kisielice - miasto (4)	2807044	3	303	1486	194	111	169
Kisielice - obszar wiejski (5)	2807045	170	59	2446	194	111	169
Lubawa (2)	2807052	236	803	6165	195	111	52
Susz - miasto (4)	2807064	7	1122	3672	194	111	169
Susz - obszar wiejski (5)	2807065	252	542	4408	194	111	169
Zalewo - miasto (4)	2807074	8	710	1419	196	111	169
Zalewo - obszar wiejski (5)	2807075	246	230	2916	196	111	169
Kętrzyn (1)	2808011	10	6809	18934	199	43	47
Barciany (2)	2808022	295	521	4169	199	43	47
Kętrzyn (2)	2808032	285	747	5121	199	43	47
Korsze - miasto (4)	2808044	4	557	2936	195	43	47
Korsze - obszar wiejski (5)	2808045	246	329	3708	195	43	47
Reszel - miasto (4)	2808054	4	1000	3216	199	43	47
Reszel - obszar wiejski (5)	2808055	175	103	1994	199	43	47
Srokowo (2)	2808062	194	251	2683	199	43	47
Lidzbark Warmiński (1)	2809011	14	3850	10462	199	43	143
Kiwity (2)	2809022	145	103	2115	199	43	143
Lidzbark Warmiński (2)	2809032	373	352	4172	199	43	143
Lubomino (2)	2809042	149	210	2313	195	43	47
Orneta - miasto (4)	2809054	10	1564	6012	195	43	47
Orneta - obszar wiejski (5)	2809055	234	103	2000	195	43	47
Mrażowo (1)	2810011	15	6076	14370	197	43	47
Mikołajki - miasto (4)	2810024	9	1077	2544	197	43	47
Mikołajki - obszar wiejski (5)	2810025	247	246	2907	197	43	47
Mrażowo (2)	2810032	295	714	4818	197	43	47
Piecki (2)	2810042	314	817	4951	197	43	47
Sorkwity (2)	2810052	185	290	2895	197	43	47
Janowiec Kościelny (2)	2811012	136	158	2047	103	43	47
Janowo (2)	2811022	192	157	1714	103	43	47
Kozłowo (2)	2811032	254	465	3781	103	43	47
Nidzica - miasto (4)	2811044	7	3787	9615	103	43	47
Nidzica - obszar wiejski (5)	2811045	372	495	4187	103	43	47
Nowe Miasto Lubawskie (1)	2812011	11	2742	7020	195	111	183
Biskupiec (2)	2812022	242	1130	5888	194	111	169

nazwa gminy	gmina				numer lokalnego rynku pracy		
	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Grodziczno (2)	2812032	154	469	3638	195	111	52
Kurzętnik (2)	2812042	149	1102	5421	195	111	183
Nowe Miasto Lubawskie (2)	2812052	138	353	4732	195	111	183
Kowale Oleckie (2)	2813032	252	482	3344	198	87	124
Olecko - miasto (4)	2813044	12	3864	10441	198	87	124
Olecko - obszar wiejski (5)	2813045	254	615	3107	198	87	124
Świątajno (2)	2813052	215	260	2424	198	87	124
Wieliczki (2)	2813062	141	289	2051	198	87	124
Barczewo - miasto (4)	2814014	5	1395	4912	195	43	47
Barczewo - obszar wiejski (5)	2814015	315	1541	6102	195	43	47
Biskupiec - miasto (4)	2814024	5	2914	6975	195	43	47
Biskupiec - obszar wiejski (5)	2814025	285	334	5471	195	43	47
Dobre Miasto - miasto (4)	2814034	5	2800	6830	199	43	47
Dobre Miasto - obszar wiejski (5)	2814035	254	326	3399	199	43	47
Dywity (2)	2814042	161	1559	6283	195	43	47
Gietrzwałd (2)	2814052	172	755	3570	195	43	47
Jeziorany - miasto (4)	2814064	3	523	2166	195	43	47
Jeziorany - obszar wiejski (5)	2814065	208	408	2956	195	43	47
Jonkowo (2)	2814072	169	1113	3749	195	43	47
Kolno (2)	2814082	179	296	2115	195	43	47
Olsztynek - miasto (4)	2814094	8	1875	4998	195	43	47
Olsztynek - obszar wiejski (5)	2814095	363	845	3740	195	30	31
Purda (2)	2814102	318	403	4904	195	43	47
Stawiguda (2)	2814112	223	760	3526	195	43	47
Świątki (2)	2814122	164	297	2580	195	43	47
Ostróda (1)	2815011	14	7869	21842	195	117	200
Dąbrówno (2)	2815022	166	470	2708	195	117	200
Grunwald (2)	2815032	180	292	3514	195	117	200
Łukta (2)	2815042	186	499	2774	195	43	47
Małdyty (2)	2815052	189	793	3905	200	117	114
Miłakowo - miasto (4) *	2815064	9	553	1717	200	117	114
Miłakowo - obszar wiejski (5) *	2815065	149	81	1960	200	117	114
Miłomłyn - miasto (4) *	2815074	12	314	1483	195	117	200
Miłomłyn - obszar wiejski (5) *	2815075	149	96	1647	195	117	200
Morąg - miasto (4)	2815084	6	3565	9451	200	117	114
Morąg - obszar wiejski (5)	2815085	305	520	6411	200	117	114
Ostróda (2)	2815092	401	3096	9794	195	117	200
Biała Piska - miasto (4)	2816014	3	597	2581	199	43	223
Biała Piska - obszar wiejski (5)	2816015	417	227	4837	195	43	223
Orzysz - miasto (4)	2816024	8	647	3676	195	43	223
Orzysz - obszar wiejski (5)	2816025	355	230	2251	195	43	223
Pisz - miasto (4)	2816034	10	4358	12400	199	43	223
Pisz - obszar wiejski (5)	2816035	624	507	4722	199	43	223
Ruciane-Nida - miasto (4)	2816044	17	580	3102	195	43	223
Ruciane-Nida - obszar wiejski (5)	2816045	341	351	2240	195	43	223
Szczytno (1)	2817011	11	6246	16808	117	43	47
Dźwierzuty (2)	2817022	263	344	4051	195	43	47
Jedwabno (2)	2817032	312	375	2206	103	43	47
Pasym - miasto (4) *	2817044	15	319	1739	195	43	47
Pasym - obszar wiejski (5) *	2817045	134	151	1599	195	43	47
Rozogi (2)	2817052	224	313	3366	104	43	47
Szczytno (2)	2817062	346	1226	6729	117	43	47
Świątajno (2)	2817072	280	454	3667	197	43	47
Wielbark (2)	2817082	348	495	3762	195	43	47
Banie Mazurskie (2) *	2818012	205	223	2355	198	87	129
Dubeninki (2) *	2818022	205	132	1921	198	87	129
Gołdap - miasto (4) *	2818034	17	3350	8579	198	87	129
Gołdap - obszar wiejski (5) *	2818035	345	295	3797	198	87	129
Budry (2) *	2819012	175	114	1843	199	124	105

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Pozezdrze (2) *	2819022	177	175	2266	199	124	215
Węgorzewo - miasto (4) *	2819034	11	1939	7583	199	124	105
Węgorzewo - obszar wiejski (5) *	2819035	330	254	3364	199	124	105
Elbląg (1)	2861011	80	29313	84584	193	90	127
Olsztyn (1)	2862011	88	60431	119290	195	43	47
Chodzież (1)	3001011	13	5767	12950	201	63	192
Budzyń (2)	3001022	209	1731	5317	201	63	192
Chodzież (2)	3001032	213	859	3483	201	63	192
Margonin - miasto (4)	3001044	5	414	1907	201	63	192
Margonin - obszar wiejski (5)	3001045	118	581	2180	201	63	192
Szamocin - miasto (4)	3001054	5	775	2689	201	63	192
Szamocin - obszar wiejski (5)	3001055	122	99	1886	201	63	192
Czarnków (1)	3002011	10	4862	7625	202	96	131
Czarnków (2)	3002022	346	1084	6745	202	96	131
Drawsko (2)	3002032	163	737	3756	202	63	83
Krzyż Wielkopolski - miasto (4)	3002044	6	1677	4094	61	96	189
Krzyż Wielkopolski - obszar wiejski (5)	3002045	168	192	1605	202	96	189
Lubasz (2)	3002052	168	476	4359	202	96	131
Połajewo (2)	3002062	141	421	3799	202	63	83
Trzcianka - miasto (4)	3002074	18	4604	10890	203	96	131
Trzcianka - obszar wiejski (5)	3002075	356	598	4223	203	96	131
Wieleń - miasto (4)	3002084	4	1899	3816	202	96	131
Wieleń - obszar wiejski (5)	3002085	426	434	4133	202	96	131
Gniezno (1)	3003011	41	16186	46038	221	63	116
Czarniejewo - miasto (4)	3003024	10	340	1702	221	63	116
Czarniejewo - obszar wiejski (5)	3003025	102	646	2878	221	63	116
Gniezno (2)	3003032	178	730	5413	221	63	116
Kiszkowo (2)	3003042	114	507	3397	221	63	83
Kłecko - miasto (4)	3003054	10	548	1782	221	63	116
Kłecko - obszar wiejski (5)	3003055	122	199	3072	221	63	116
Łubowo (2)	3003062	114	469	3441	221	63	83
Mieleszyn (2)	3003072	99	423	2543	221	63	116
Niechanowo (2)	3003082	106	448	3417	221	63	116
Trzemeszno - miasto (4)	3003094	5	1751	5233	221	109	116
Trzemeszno - obszar wiejski (5)	3003095	170	325	3960	221	109	116
Witkowo - miasto (4)	3003104	8	1069	5234	221	63	116
Witkowo - obszar wiejski (5)	3003105	176	221	3510	221	63	116
Borek Wielkopolski - miasto (4)	3004014	6	951	1606	204	75	98
Borek Wielkopolski - obszar wiejski (5)	3004015	122	248	3172	204	75	98
Gostyń - miasto (4)	3004024	10	7111	13540	204	75	98
Gostyń - obszar wiejski (5)	3004025	126	1370	4552	204	75	98
Krobia - miasto (4)	3004034	7	690	2587	204	75	98
Krobia - obszar wiejski (5)	3004035	123	1053	5434	204	75	98
Pępowo (2)	3004042	87	683	3619	204	75	211
Piaski (2)	3004052	101	1314	5161	204	75	98
Pogorzela - miasto (4)	3004064	4	433	1240	207	75	211
Pogorzela - obszar wiejski (5)	3004065	93	113	1892	207	75	211
Poniec - miasto (4)	3004074	3	858	1870	204	75	98
Poniec - obszar wiejski (5)	3004075	128	273	3102	204	75	98
Granowo (2)	3005012	67	568	3074	205	63	83
Grodzisk Wielkopolski - miasto (4)	3005024	18	6655	9001	205	100	158
Grodzisk Wielkopolski - obszar wiejski (5)	3005025	115	545	3092	205	100	158
Kamieniec (2)	3005032	134	609	4019	205	100	158
Rakoniewice - miasto (4)	3005044	3	901	2115	205	100	159
Rakoniewice - obszar wiejski (5)	3005045	197	1036	5686	205	100	159
Wielichowo - miasto (4)	3005054	1	294	1128	205	100	158
Wielichowo - obszar wiejski (5)	3005055	107	261	3219	205	100	158
Jaraczewo (2)	3006012	133	1077	5105	221	80	82

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Jarocin - miasto (4)	3006024	15	7574	17118	221	80	82
Jarocin - obszar wiejski (5)	3006025	185	1828	11652	221	80	82
Kotlin (2)	3006032	84	819	4438	212	80	82
Żerków - miasto (4)	3006044	2	410	1350	221	80	82
Żerków - obszar wiejski (5)	3006045	168	316	5161	221	80	82
Blizanów (2)	3007012	158	936	5775	219	65	86
Brzeziny (2)	3007022	127	336	3584	219	65	86
Ceków-Kolonia (2)	3007032	88	473	2874	219	65	86
Godziesze Wielkie (2)	3007042	105	765	5172	219	65	86
Koźminek (2)	3007052	89	502	4633	219	65	86
Lisków (2)	3007062	75	675	3217	219	65	86
Mycielin (2)	3007072	111	258	3000	219	65	86
Opatówek (2)	3007082	104	1212	6395	219	65	86
Stawiszyn - miasto (4)	3007094	1	191	1018	219	65	86
Stawiszyn - obszar wiejski (5)	3007095	78	486	3573	219	65	86
Szczytniki (2)	3007102	110	437	4934	219	65	86
Żelazków (2)	3007112	114	959	5650	219	65	86
Baranów (2)	3008012	74	3151	4794	77	126	225
Bralin (2)	3008022	85	966	3560	77	126	225
Kępno - miasto (4)	3008034	8	5312	9740	77	126	225
Kępno - obszar wiejski (5)	3008035	116	2033	5976	77	126	225
Łęka Opatowska (2)	3008042	78	1796	3248	77	126	225
Perzów (2)	3008052	75	582	2436	77	126	225
Rychtal (2)	3008062	97	246	2563	77	126	225
Trzcinica (2)	3008072	75	383	2886	77	126	225
Koło (1)	3009011	14	8307	15433	206	26	27
Babinek (2)	3009022	134	351	4733	206	26	27
Chodów (2)	3009032	78	183	2084	206	6	6
Dąbie - miasto (4)	3009044	9	189	1334	206	26	27
Dąbie - obszar wiejski (5)	3009045	121	108	2715	206	26	27
Grzegorzew (2)	3009052	73	454	3396	206	26	27
Kłodawa - miasto (4)	3009064	4	1619	4407	206	26	27
Kłodawa - obszar wiejski (5)	3009065	125	193	3883	206	26	27
Koło (2)	3009072	103	403	4345	206	26	27
Kościelec (2)	3009082	105	488	4105	206	26	27
Olszówka (2)	3009092	82	104	2834	206	26	27
Osiek Mały (2)	3009102	87	238	3592	206	26	27
Przedecz - miasto (4)	3009114	3	199	1104	206	26	27
Przedecz - obszar wiejski (5)	3009115	73	62	1500	206	26	27
Golina - miasto (4)	3010014	4	490	2901	220	53	66
Golina - obszar wiejski (5)	3010015	95	152	4364	220	53	66
Grodziec (2)	3010022	118	281	3137	220	53	66
Kazimierz Biskupi (2)	3010032	108	2162	6704	220	53	66
Kleczew - miasto (4)	3010044	7	2108	2613	220	53	66
Kleczew - obszar wiejski (5)	3010045	103	1033	3454	220	53	66
Kramsk (2)	3010052	132	730	6202	220	53	66
Krzymów (2)	3010062	93	885	4365	220	53	66
Rychwał - miasto (4)	3010074	10	319	1560	220	53	66
Rychwał - obszar wiejski (5)	3010075	108	125	3582	220	53	66
Rzgów (2)	3010082	105	257	4157	220	53	66
Skulsk (2)	3010092	85	356	3820	220	53	66
Sompolno - miasto (4)	3010104	6	435	2378	206	53	66
Sompolno - obszar wiejski (5)	3010105	131	1085	4153	206	53	66
Stare Miasto (2)	3010112	98	978	6706	220	53	66
Ślesin - miasto (4)	3010124	7	637	2021	220	53	66
Ślesin - obszar wiejski (5)	3010125	138	525	6356	220	53	66
Wierzbinek (2)	3010132	147	453	4523	220	53	66
Wilczyn (2)	3010142	83	295	3945	220	53	66
Kościan (1)	3011011	9	5846	15826	221	22	23

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Czempiń - miasto (4)	3011024	3	1120	3452	221	63	83
Czempiń - obszar wiejski (5)	3011025	139	904	3855	221	63	83
Kościan (2)	3011032	202	1692	9599	221	22	23
Krzywiń - miasto (4)	3011044	2	460	1006	221	22	23
Krzywiń - obszar wiejski (5)	3011045	177	810	5155	221	22	23
Śmigiel - miasto (4)	3011054	5	1513	3502	221	22	23
Śmigiel - obszar wiejski (5)	3011055	185	1056	7650	221	22	23
Sulmierzyce (1)	3012011	29	228	1768	207	55	70
Kobylin - miasto (4)	3012024	5	420	1938	207	75	211
Kobylin - obszar wiejski (5)	3012025	107	961	3096	207	75	211
Koźmin Wielkopolski - miasto (4) *	3012034	6	1470	4261	207	55	70
Koźmin Wielkopolski - obszar wiejski (5)	3012035	146	414	4255	207	55	70
Krotoszyn - miasto (4)	3012044	23	9336	18943	207	55	70
Krotoszyn - obszar wiejski (5)	3012045	234	818	6790	207	55	70
Rozdrażew (2)	3012052	79	383	3105	207	55	70
Zduny - miasto (4)	3012064	6	510	2957	207	55	70
Zduny - obszar wiejski (5)	3012065	79	214	1556	207	55	70
Krzemieniewo (2)	3013012	113	1145	5446	221	22	23
Lipno (2)	3013022	104	1521	3648	221	22	23
Osieczna - miasto (4)	3013034	5	488	1322	221	22	23
Osieczna - obszar wiejski (5)	3013035	124	913	4124	221	22	23
Rydzyzna - miasto (4)	3013044	2	1168	1658	221	22	23
Rydzyzna - obszar wiejski (5)	3013045	133	784	3570	221	22	23
Święciechowa (2)	3013052	135	874	4614	221	22	23
Wijewo (2)	3013062	62	296	2167	67	22	23
Włoszakowice (2)	3013072	128	1486	5395	205	22	23
Chrzypsko Wielkie (2)	3014012	85	524	1979	215	63	83
Kwilcz (2)	3014022	142	618	4016	208	100	83
Międzychód - miasto (4)	3014034	7	2854	7243	208	123	213
Międzychód - obszar wiejski (5)	3014035	299	1846	4808	208	123	213
Sieraków - miasto (4)	3014044	14	1152	3881	208	123	213
Sieraków - obszar wiejski (5)	3014045	189	289	1660	208	123	213
Kuślin (2)	3015012	106	556	3492	209	100	80
Lwówek - miasto (4)	3015024	3	734	1828	209	100	80
Lwówek - obszar wiejski (5)	3015025	181	299	3801	209	100	80
Miedzichowo (2)	3015032	209	306	2468	209	100	80
Nowy Tomyśl - miasto (4)	3015044	5	6381	10288	209	100	80
Nowy Tomyśl - obszar wiejski (5)	3015045	181	1162	5792	209	100	80
Opalenica - miasto (4)	3015054	6	2102	6058	209	100	231
Opalenica - obszar wiejski (5)	3015055	143	701	4074	209	100	231
Zbąszyń - miasto (4)	3015064	5	1848	4743	209	100	206
Zbąszyń - obszar wiejski (5)	3015065	175	565	3791	209	100	206
Oborniki - miasto (4)	3016014	14	4946	12051	218	63	83
Oborniki - obszar wiejski (5)	3016015	326	1888	8934	218	63	83
Rogoźno - miasto (4)	3016024	11	2561	7154	218	63	83
Rogoźno - obszar wiejski (5)	3016025	205	580	3940	218	63	83
Ryczywół (2)	3016032	155	527	4573	218	63	83
Ostrów Wielkopolski (1)	3017011	42	19989	47657	210	83	111
Nowe Skalmierzyce - miasto (4)	3017024	2	1097	3370	210	65	86
Nowe Skalmierzyce - obszar wiejski (5)	3017025	123	2906	6352	210	65	86
Odolanów - miasto (4)	3017034	5	1447	3228	210	83	111
Odolanów - obszar wiejski (5)	3017035	131	521	5387	210	83	111
Ostrów Wielkopolski (2)	3017042	208	1796	11185	210	83	111
Przygodzice (2)	3017052	163	723	7108	210	83	111
Raszków - miasto (4)	3017064	2	230	1387	210	83	111
Raszków - obszar wiejski (5)	3017065	133	776	5624	210	83	111
Sieroszewice (2)	3017072	163	548	5998	210	83	111
Sośnie (2)	3017082	188	392	4154	8	49	38
Czajków (2)	3018012	71	126	1517	211	31	32

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Doruchów (2)	3018022	99	328	3153	211	31	32
Grabów nad Prosną - miasto (4)	3018034	3	727	1306	211	31	32
Grabów nad Prosną - obszar wiejski (5)	3018035	121	237	3486	211	31	32
Kobyła Góra (2)	3018042	129	636	3653	211	31	32
Kraszewice (2)	3018052	75	292	2194	211	31	32
Mikstat - miasto (4)	3018064	3	541	1235	211	31	32
Mikstat - obszar wiejski (5)	3018065	84	140	2667	211	31	32
Ostrzeszów - miasto (4)	3018074	12	6628	9638	211	31	32
Ostrzeszów - obszar wiejski (5)	3018075	176	482	5450	211	31	32
Piła (1)	3019011	103	24115	49928	203	96	144
Białośliwie (2)	3019022	76	1244	3085	203	96	144
Kaczory (2)	3019032	151	1692	4863	203	96	144
Łobżenica - miasto (4)	3019044	3	627	2063	203	96	144
Łobżenica - obszar wiejski (5)	3019045	188	337	4102	203	96	144
Miasteczko Krajeńskie (2)	3019052	71	209	1976	203	96	144
Szydłowo (2)	3019062	266	537	4975	203	96	144
Ujście - miasto (4)	3019074	6	900	2493	203	96	144
Ujście - obszar wiejski (5)	3019075	122	419	2639	203	96	144
Wyrzysk - miasto (4)	3019084	4	1199	3403	203	96	144
Wyrzysk - obszar wiejski (5)	3019085	155	889	5488	203	96	144
Wysoka - miasto (4)	3019094	5	296	1741	203	96	144
Wysoka - obszar wiejski (5)	3019095	118	487	2553	203	96	144
Chocz (2)	3020012	74	233	2987	212	80	107
Czermin (2)	3020022	98	343	3055	212	80	107
Dobrzyca (2)	3020032	117	1498	5218	212	80	107
Gizalki (2)	3020042	108	385	2856	212	80	107
Gołuchów (2)	3020052	136	1039	6090	212	65	86
Pleszew - miasto (4)	3020064	13	5139	11496	212	80	107
Pleszew - obszar wiejski (5)	3020065	167	1059	7799	212	80	107
Luboń (1)	3021011	14	3552	17789	221	63	83
Puszczykowo (1)	3021021	16	2116	6093	221	63	83
Buk - miasto (4)	3021034	3	1672	3985	209	63	231
Buk - obszar wiejski (5)	3021035	88	484	3678	209	63	231
Czerwonak (2)	3021042	82	5330	16800	221	63	83
Dopiewo (2)	3021052	108	3505	9469	221	63	83
Kleszczewo (2)	3021062	74	689	3561	221	63	83
Komorniki (2)	3021072	66	5934	9832	221	63	83
Kostrzyn - miasto (4)	3021084	8	3462	5606	221	63	83
Kostrzyn - obszar wiejski (5)	3021085	147	522	4440	221	63	83
Kórnik - miasto (4)	3021094	6	1386	4568	221	63	83
Kórnik - obszar wiejski (5)	3021095	181	3878	7156	221	63	83
Mosina - miasto (4)	3021104	14	2438	8153	221	63	83
Mosina - obszar wiejski (5)	3021105	158	1967	8531	221	63	83
Murowana Goślina - miasto (4)	3021114	6	2308	7067	221	63	83
Murowana Goślina - obszar wiejski (5)	3021115	166	328	3676	221	63	83
Pobiedziska - miasto (4)	3021124	9	1108	5407	221	63	83
Pobiedziska - obszar wiejski (5)	3021125	181	843	5229	221	63	83
Rokietnica (2)	3021132	79	1537	6366	221	63	83
Stęszew - miasto (4)	3021144	6	2016	3579	221	63	83
Stęszew - obszar wiejski (5)	3021145	169	1235	5451	221	63	83
Suchy Las (2)	3021152	116	4426	9295	221	63	83
Swarzędz - miasto (4)	3021164	9	7930	20422	221	63	83
Swarzędz - obszar wiejski (5)	3021165	93	4287	7119	221	63	83
Tarnowo Podgórne (2)	3021172	102	14345	12602	213	63	83
Bojanowo - miasto (4)	3022014	2	588	1870	221	122	164
Bojanowo - obszar wiejski (5)	3022015	122	784	3780	221	122	164
Jutrosin - miasto (4)	3022024	2	505	1192	8	122	164
Jutrosin - obszar wiejski (5)	3022025	113	306	3141	8	122	164
Miejscza Górka - miasto (4)	3022034	3	656	2028	221	122	164

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Miejska Górka - obszar wiejski (5)	3022035	100	426	3824	221	122	164
Pakosław (2)	3022042	77	1229	2855	8	122	164
Rawicz - miasto (4)	3022054	8	8312	13963	221	122	164
Rawicz - obszar wiejski (5)	3022055	127	2350	5216	221	122	164
Słupca (1)	3023011	10	4906	9807	214	53	161
Lądek (2)	3023022	99	433	3448	214	53	161
Orchowo (2)	3023032	98	284	2440	25	109	201
Ostrowite (2)	3023042	104	376	3077	214	53	161
Powidz (2)	3023052	81	207	1304	221	63	116
Słupca (2)	3023062	145	505	5408	214	53	161
Strzałkowo (2)	3023072	142	1064	6078	214	53	161
Zagórów - miasto (4)	3023084	3	642	1916	214	53	161
Zagórów - obszar wiejski (5)	3023085	156	149	3692	214	53	161
Obrzycko (1)	3024011	4	257	1429	215	63	83
Duszniki (2)	3024022	156	1032	5268	213	63	83
Kaźmierz (2)	3024032	128	998	4733	213	63	83
Obrzycko (2)	3024042	111	724	2775	215	63	83
Ostroróg - miasto (4)	3024054	1	368	1319	215	63	83
Ostroróg - obszar wiejski (5)	3024055	84	376	1819	215	63	83
Pniewy - miasto (4)	3024064	9	2297	5035	215	100	83
Pniewy - obszar wiejski (5)	3024065	149	380	2787	215	100	83
Szamotuły - miasto (4)	3024074	10	4561	12163	215	63	83
Szamotuły - obszar wiejski (5)	3024075	166	1307	6337	215	63	83
Wronki - miasto (4)	3024084	6	5269	7761	215	96	83
Wronki - obszar wiejski (5)	3024085	295	842	4605	215	96	83
Dominowo (2)	3025012	79	169	1840	216	63	83
Krzykosy (2)	3025022	110	415	4098	216	63	83
Nowe Miasto nad Wartą (2)	3025032	120	1155	5632	216	63	83
Środa Wielkopolska - miasto (4)	3025044	18	6555	14350	216	63	83
Środa Wielkopolska - obszar wiejski (5)	3025045	190	1213	5571	216	63	83
Zaniemyśl (2)	3025052	107	883	3960	216	63	83
Brodnica (2)	3026012	96	631	2964	221	63	83
Dolsk - miasto (4)	3026024	6	246	911	204	63	83
Dolsk - obszar wiejski (5)	3026025	118	396	2661	204	63	83
Książ Wielkopolski - miasto (4)	3026034	2	781	1783	221	63	83
Książ Wielkopolski - obszar wiejski (5)	3026035	146	450	3500	221	63	83
Śrem - miasto (4)	3026044	12	8270	20285	221	63	83
Śrem - obszar wiejski (5)	3026045	194	2814	6108	221	63	83
Turek (1)	3027011	16	8834	19025	217	99	153
Brudzew (2)	3027022	113	751	3666	217	99	153
Dobra - miasto (4)	3027034	2	324	960	217	99	153
Dobra - obszar wiejski (5)	3027035	130	499	2914	217	99	153
Kawęczyn (2)	3027042	101	168	3145	217	99	153
Małanów (2)	3027052	107	1888	3911	217	99	153
Przykona (2)	3027062	111	1108	2479	217	99	153
Tuliszków - miasto (4)	3027074	7	313	2296	217	53	66
Tuliszków - obszar wiejski (5)	3027075	142	393	4344	217	99	153
Turek (2)	3027082	109	1310	4659	217	99	153
Władysławów (2)	3027092	91	831	4845	217	99	153
Wągrowiec (1)	3028011	18	5805	16280	221	63	210
Damaśławek (2)	3028022	105	386	3477	221	63	210
Gołańcz - miasto (4)	3028034	13	729	2169	201	63	210
Gołańcz - obszar wiejski (5)	3028035	179	129	3032	201	63	210
Mieścisko (2)	3028042	135	473	3689	221	63	83
Skoki - miasto (4)	3028054	11	244	2602	221	63	83
Skoki - obszar wiejski (5)	3028055	187	285	3040	221	63	83
Wapno (2)	3028062	44	220	1950	201	63	210
Wągrowiec (2)	3028072	348	805	7137	221	63	210
Przemęt (2)	3029012	225	1533	8540	205	22	23

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Siedlec (2)	3029022	205	1998	7578	205	100	159
Wolsztyn - miasto (4)	3029034	5	4900	8789	205	100	159
Wolsztyn - obszar wiejski (5)	3029035	245	3066	9936	205	100	159
Kołaczkowo (2)	3030012	116	508	3859	221	63	83
Miłosław - miasto (4)	3030024	4	775	2306	216	63	83
Miłosław - obszar wiejski (5)	3030025	128	1004	4338	216	63	83
Nekla - miasto (4) *	3030034	20	542	2177	221	63	83
Nekla - obszar wiejski (5) *	3030035	76	838	2194	221	63	83
Pyzdry - miasto (4)	3030044	12	524	2046	214	63	161
Pyzdry - obszar wiejski (5)	3030045	126	105	2404	214	63	161
Września - miasto (4)	3030054	13	7388	18927	221	63	83
Września - obszar wiejski (5)	3030055	209	1346	9549	221	63	83
Złotów (1)	3031011	12	4471	12192	218	121	209
Jastrowie - miasto (4)	3031024	72	1626	5487	203	64	148
Jastrowie - obszar wiejski (5)	3031025	282	287	1863	203	64	148
Krajenka - miasto (4)	3031034	4	739	2311	218	121	209
Krajenka - obszar wiejski (5)	3031035	187	225	2198	218	96	144
Lipka (2)	3031042	191	469	3508	218	121	209
Okonek - miasto (4)	3031054	6	605	2445	233	64	148
Okonek - obszar wiejski (5)	3031055	320	1081	3157	233	64	148
Tarnówka (2)	3031062	132	179	1957	218	121	209
Zakrzewo (2)	3031072	162	593	2909	218	121	209
Złotów (2)	3031082	292	1013	5760	218	121	209
Kalisz (1)	3061011	69	31957	70274	219	65	86
Konin (1)	3062011	82	25334	53396	220	53	66
Leszno (1)	3063011	32	20716	42235	221	22	23
Poznań (1)	3064011	262	224741	379394	221	63	83
Białogard (1)	3201011	26	5212	15757	242	42	46
Białogard (2)	3201022	328	303	4892	242	42	46
Karlıno - miasto (4)	3201034	9	1166	3848	242	42	46
Karlıno - obszar wiejski (5)	3201035	132	317	2065	242	42	46
Tychowo (2) *	3201042	351	986	4438	235	42	46
Bierzwnik (2)	3202012	239	551	2973	222	4	4
Choszczno - miasto (4)	3202024	10	3022	10618	222	4	4
Choszczno - obszar wiejski (5)	3202025	236	549	3998	222	4	4
Drawno - miasto (4)	3202034	5	325	1549	222	4	4
Drawno - obszar wiejski (5)	3202035	317	258	1850	222	4	4
Krzęcin (2)	3202042	140	167	2386	222	4	4
Pełczyce - miasto (4)	3202054	13	220	1773	231	1	1
Pełczyce - obszar wiejski (5)	3202055	188	344	3398	231	1	1
Recz - miasto (4)	3202064	12	632	1946	222	4	4
Recz - obszar wiejski (5)	3202065	168	95	1714	222	4	4
Czaplinek - miasto (4)	3203014	14	2179	4608	223	64	85
Czaplinek - obszar wiejski (5)	3203015	351	294	3061	223	64	85
Drawsko Pomorskie - miasto (4)	3203024	22	2367	7667	224	64	85
Drawsko Pomorskie - obszar wiejski (5)	3203025	323	290	3124	224	64	85
Kalisz Pomorski - miasto (4)	3203034	12	778	2648	238	96	100
Kalisz Pomorski - obszar wiejski (5)	3203035	468	205	1893	238	96	100
Ostrowice (2)	3203042	150	219	1617	224	64	85
Wierzchowo (2)	3203052	229	815	2926	225	64	85
Złocieniec - miasto (4)	3203064	32	1472	8901	225	64	85
Złocieniec - obszar wiejski (5)	3203065	163	220	1391	225	64	85
Goleniów - miasto (4)	3204024	12	5775	15000	243	119	203
Goleniów - obszar wiejski (5)	3204025	431	1739	7063	243	119	203
Maszewo - miasto (4)	3204034	6	574	1911	243	119	203
Maszewo - obszar wiejski (5)	3204035	204	174	3264	243	119	203
Nowogard - miasto (4)	3204044	12	3290	11335	243	119	203
Nowogard - obszar wiejski (5)	3204045	327	256	4854	243	119	203
Osina (2)	3204052	102	414	1824	243	119	203

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Przybiernów (2)	3204062	229	467	3307	243	4	4
Stepnica (2)	3204072	293	775	3111	243	119	203
Brojce (2)	3205012	118	225	2387	226	94	168
Gryfice - miasto (4)	3205024	12	3598	11135	226	94	168
Gryfice - obszar wiejski (5)	3205025	249	585	4290	226	94	168
Karnice (2)	3205032	133	199	2655	227	94	168
Ploty - miasto (4)	3205044	4	599	2684	241	94	168
Ploty - obszar wiejski (5)	3205045	235	270	3121	241	94	168
Rewal (2)	3205072	41	849	2277	227	94	168
Trzebiatów - miasto (4)	3205084	10	1361	6802	226	94	140
Trzebiatów - obszar wiejski (5)	3205085	215	668	4299	226	94	140
Banie (2)	3206012	206	414	4160	243	4	4
Cedynia - miasto (4)	3206024	2	258	1095	228	4	4
Cedynia - obszar wiejski (5)	3206025	179	275	1702	228	4	4
Chojna - miasto (4)	3206034	12	1364	4816	228	4	4
Chojna - obszar wiejski (5)	3206035	321	263	4234	228	4	4
Gryfino - miasto (4)	3206044	10	3089	15186	243	4	4
Gryfino - obszar wiejski (5)	3206045	244	3275	6339	243	4	4
Mieszkowice - miasto (4)	3206054	5	492	2277	228	4	4
Mieszkowice - obszar wiejski (5)	3206055	234	172	2350	228	1	214
Moryń - miasto (4)	3206064	6	221	1128	228	4	4
Moryń - obszar wiejski (5)	3206065	119	102	1819	228	4	4
Stare Czarnowo (2)	3206072	153	422	2599	243	4	4
Trzcianko-Zdrój - miasto (4)	3206084	2	455	1521	228	4	4
Trzcianko-Zdrój - obszar wiejski (5)	3206085	168	116	1969	228	4	4
Widuchowa (2)	3206092	209	332	3552	243	4	4
Dziwnów - miasto (4) *	3207014	5	312	1997	229	118	202
Dziwnów - obszar wiejski (5) *	3207015	33	219	834	229	118	202
Golczewo - miasto (4)	3207024	7	363	1764	229	94	168
Golczewo - obszar wiejski (5)	3207025	168	122	2093	229	4	4
Kamień Pomorski - miasto (4)	3207034	11	1848	6198	229	118	202
Kamień Pomorski - obszar wiejski (5)	3207035	198	274	3405	229	118	202
Międzyzdroje - miasto (4)	3207044	4	1114	3664	213	102	235
Międzyzdroje - obszar wiejski (5)	3207045	110	60	719	213	102	235
Świerzno (2)	3207052	140	329	2702	229	118	202
Wolin - miasto (4)	3207064	14	500	3405	213	102	235
Wolin - obszar wiejski (5)	3207065	313	411	4814	213	102	235
Kołobrzeg (1)	3208011	26	10850	29726	230	94	140
Dygowo (2)	3208022	129	407	3511	230	94	140
Gościno (2) *	3208032	116	645	3163	230	94	140
Kołobrzeg (2)	3208042	144	935	6301	230	94	140
Rymań (2)	3208052	146	351	2625	230	94	140
Siemyśl (2)	3208062	107	348	2329	230	94	140
Ustronie Morskie (2)	3208072	57	741	2399	230	94	140
Będzino (2)	3209012	181	977	6159	242	42	46
Biesiekierz (2)	3209022	117	891	3658	242	42	46
Bobolice - miasto (4)	3209034	5	792	2902	242	42	46
Bobolice - obszar wiejski (5)	3209035	362	395	3381	242	42	46
Manowo (2)	3209042	188	618	4210	242	42	46
Mielno (2)	3209052	62	854	3362	242	42	46
Polanów - miasto (4)	3209064	7	556	1954	242	42	46
Polanów - obszar wiejski (5)	3209065	387	458	3978	242	42	46
Sianów - miasto (4)	3209074	16	667	4319	242	42	46
Sianów - obszar wiejski (5)	3209075	211	744	4254	242	42	46
Świeszyno (2)	3209082	132	490	3688	242	42	46
Barlinek - miasto (4)	3210014	18	4655	9561	231	1	1
Barlinek - obszar wiejski (5)	3210015	241	187	3457	231	1	1
Boleszkowice (2)	3210022	130	187	1830	57	1	214
Dębno - miasto (4)	3210034	20	2897	9122	57	1	214

nazwa gminy	gmina				numer lokalnego rynku pracy		
	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Dębno - obszar wiejski (5)	3210035	298	470	4241	57	1	214
Myślibórz - miasto (4)	3210044	15	2737	7892	231	1	1
Myślibórz - obszar wiejski (5)	3210045	314	416	5749	231	1	1
Nowogródek Pomorski (2)	3210052	146	226	2044	231	1	1
Dobra (Szczecińska) (2)	3211012	110	2665	8740	243	4	4
Kołbaskowo (2)	3211022	105	678	6052	243	4	4
Nowe Warpno - miasto (4)	3211034	25	104	778	243	4	4
Nowe Warpno - obszar wiejski (5)	3211035	173	12	263	243	4	4
Police - miasto (4)	3211044	37	7911	24446	243	4	4
Police - obszar wiejski (5)	3211045	215	798	4726	243	4	4
Bielice (2)	3212012	84	129	1856	232	4	4
Kozielice (2)	3212022	95	232	1650	232	4	4
Lipiany - miasto (4)	3212034	6	830	2730	232	4	4
Lipiany - obszar wiejski (5)	3212035	89	19	1229	232	4	4
Przelewice (2)	3212042	162	247	3289	232	4	4
Pyrzyce - miasto (4)	3212054	39	2823	8468	232	4	4
Pyrzyce - obszar wiejski (5)	3212055	165	285	4331	232	4	4
Warnice (2)	3212062	86	295	2235	232	4	4
Darłowo (1)	3213011	20	1842	9509	242	42	46
Sławno (1)	3213021	16	3067	8529	163	86	123
Darłowo (2)	3213032	269	1133	4763	242	42	46
Malechowo (2)	3213042	227	490	4086	242	42	46
Postomino (2)	3213052	227	493	4367	163	86	123
Sławno (2)	3213062	284	598	5532	163	86	123
Stargard Szczeciński (1)	3214011	48	12655	47738	243	4	4
Chociwel - miasto (4)	3214024	4	629	2216	243	4	4
Chociwel - obszar wiejski (5)	3214025	157	83	1736	243	4	4
Dobrzany - miasto (4)	3214034	5	512	1583	243	52	156
Dobrzany - obszar wiejski (5)	3214035	130	51	1623	243	52	156
Dolice (2)	3214042	237	718	5061	243	4	4
Ińsko - miasto (4)	3214054	7	281	1328	243	4	4
Ińsko - obszar wiejski (5)	3214055	144	42	860	243	4	4
Kobylanka (2)	3214062	122	527	2466	243	4	4
Marianowo (2)	3214082	102	141	1933	243	4	4
Stara Dąbrowa (2)	3214092	113	167	2240	243	4	4
Stargard Szczeciński (2)	3214102	318	1172	7154	243	4	4
Suchań - miasto (4)	3214114	4	169	953	243	4	4
Suchań - obszar wiejski (5)	3214115	129	17	1812	243	4	4
Szczecinek (1)	3215011	37	9774	25172	233	64	148
Barwice - miasto (4)	3215024	7	468	2459	223	64	148
Barwice - obszar wiejski (5)	3215025	252	284	3146	223	64	148
Biały Bór - miasto (4)	3215034	13	296	1387	154	34	36
Biały Bór - obszar wiejski (5)	3215035	257	123	1888	154	34	36
Borne Sulinowo - miasto (4)	3215044	18	621	2440	233	64	148
Borne Sulinowo - obszar wiejski (5)	3215045	466	273	3171	233	64	148
Grzmiąca (2)	3215052	205	346	3100	233	64	148
Szczecinek (2)	3215062	510	624	6330	233	64	148
Świdwin (1)	3216011	22	2844	10309	234	74	97
Brzeźno (2)	3216022	111	140	1783	234	74	97
Połczyn-Zdrój - miasto (4)	3216034	7	1693	5566	235	74	97
Połczyn-Zdrój - obszar wiejski (5)	3216035	337	954	4762	235	74	97
Rąbino (2)	3216042	180	248	2441	234	74	97
Sławoborze (2)	3216052	189	374	2764	234	74	97
Świdwin (2)	3216062	247	883	3920	234	74	97
Wałcz (1)	3217011	38	5772	17490	236	96	100
Człopa - miasto (4)	3217024	6	273	1533	237	96	131
Człopa - obszar wiejski (5)	3217025	343	136	1655	237	96	131
Mirosławiec - miasto (4)	3217034	2	612	1729	238	96	100
Mirosławiec - obszar wiejski (5)	3217035	201	162	2069	238	96	100

gmina					numer lokalnego rynku pracy		
nazwa gminy	TERYT 2006	powierzchnia w km2	liczba osób pracujących	liczba ludności w wieku produkcyjnym	europejski algorytm regionalizacji	podejście taksonomiczne - słabe kryterium grupowania	podejście taksonomiczne - mocne kryterium grupowania
Tuczno - miasto (4)	3217044	9	504	1289	239	96	100
Tuczno - obszar wiejski (5)	3217045	240	78	1866	239	96	100
Wałcz (2)	3217052	576	567	7877	236	96	100
Dobra - miasto (4) *	3218014	2	206	1262	243	119	203
Dobra - obszar wiejski (5) *	3218015	114	147	1538	243	4	4
Łobez - miasto (4) *	3218024	12	2345	6984	240	72	4
Łobez - obszar wiejski (5) *	3218025	215	201	2349	240	72	4
Radowo Małe (2) *	3218032	180	191	2357	240	72	4
Resko - miasto (4) *	3218044	4	1144	2954	241	72	4
Resko - obszar wiejski (5) *	3218045	282	140	2319	241	72	4
Węgorzyno - miasto (4) *	3218054	7	540	1985	240	4	4
Węgorzyno - obszar wiejski (5) *	3218055	249	178	2581	240	4	4
Koszalin (1)	3261011	83	27457	72473	242	42	46
Szczecin (1)	3262011	301	112590	272549	243	4	4
Świnoujście (1)	3263011	197	7201	27572	213	102	235

Źródło: Opracowanie własne

Doktorantka uzyskała w roku 2010 stypendium w ramach projektu
pt. **Wsparcie stypendialne dla doktorantów na kierunkach uznanych
za strategiczne z punktu widzenia rozwoju Wielkopolski,**
finansowanego ze środków Europejskiego Funduszu Społecznego.

Część obliczeń wykonano w Poznańskim Centrum Superkomputerowo – Sieciowym.