

GŁÓWNY URZĄD STATYSTYCZNY



**ESTYMACJA DANYCH
Z BADANIA AKTYWNOŚCI
EKONOMICZNEJ LUDNOŚCI
NA POZIOMIE POWIATÓW
DLA LAT 1995–2002**

**Opracowanie pod kierunkiem
prof. dra hab. CZESŁAWA BRACHY**

Warszawa, grudzień 2003

KOMITET REDAKCYJNY GŁÓWNEGO URZĘDU STATYSTYCZNEGO

PRZEWODNICZĄCY

Tadeusz Toczyński

REDAKTOR GŁÓWNY

Halina Dmochowska

CZŁONKOWIE

Wojciech Adamczewski, Maria Bieńkowska, Stanisława Borkowska, Małgorzata Fronk,
Iwona Gruczyńska, Bożena Jakóbiak, Janina Janecka, Małgorzata Kałaska, Jacek Kotowski,
Liliana Kursa, Lucyna Nowak, Lucyna Przybylska, Grażyna Szydłowska

SEKRETARZ

Hanna Poławska

Publikowanie wyników NSP 2002

pod kierunkiem

Janusza Witkowskiego, Zastępcy Generalnego Komisarza
Spisowego, Wiceprezesa GUS

Opracowanie analityczne o charakterze naukowo-badawczym wykonane pod kierunkiem
prof. dr hab. Czesława Bracha

Zespół autorski:

Czesław Bracha

Bronisław Lednicki

Robert Wieczorkowski

Praca zrealizowana w Zakładzie Badań Statystyczno-Ekonomicznych GUS i PAN; kierujący
tematem – Stanisław Godowski, Z-ca Dyrektora

Współpraca:

Pracownicy GUS

Recenzent:

Dr hab. inż. Paweł J. Szablowski

Konsultacje dotyczące danych statystycznych:

Małgorzata Kałaska

Przedmowa

Niniejsza publikacja zawiera wyniki prac analitycznych i obliczeniowych o charakterze metodologicznym, wykonanych w Zakładzie Badań Statystyczno-Ekonomicznych GUS i PAN pod kierunkiem prof. dr hab. Czesław Brachy. Prace dotyczyły problemu wykorzystania wyników Narodowego Spisu Powszechnego Ludności i Mieszkań do estymacji danych o aktywności ekonomicznej ludności na poziomie województw, podregionów i powiatów oraz wykonania stosownych szacunków.

W analizach dotyczących rynku pracy niezbędne jest dysponowanie szczegółową informacją o liczbie osób aktywnych zawodowo, pracujących i bezrobotnych oraz ich charakterystyce demograficzno-społecznej i zawodowej na różnych poziomach podziału terytorialnego kraju. Badanie Aktywności Ekonomicznej Ludności (BAEL) realizowane przez Główny Urząd Statystyczny od 1992 roku dostarcza informacji dla całego kraju oraz w pewnym, ograniczonym zakresie dla województw. W celu wykonania szacunków ludności aktywnej zawodowo na niższym poziomie agregacji dla lat 1995-2002 Autorzy odwołali się do metod estymacji właściwych dla statystyki małych obszarów.

W opracowaniu przedstawiono estymatory zastosowane do rozszacowania wyników BAEL na poziom województw, podregionów i powiatów (wykorzystujące informacje uzyskane na podstawie NSP'2002), własności tych estymatorów oraz metodę szacowania ich precyzji; zamieszczono też opis schematu losowania próby do BAEL oraz stosowanej w tym badaniu metody estymacji. Załączone w aneksie publikacji tablice prezentują wyniki obliczeń wykonanych dla 2002 roku, natomiast pełny zestaw danych oszacowanych dla lat 1995-2002 zawiera dołączona do publikacji płyta CD.

Wyniki opracowania obejmują dane o aktywności ekonomicznej ludności według płci, wieku, poziomu wykształcenia, miejsca zamieszkania, a dla pracujących również według statusu zatrudnienia i sektorów ekonomicznych prezentowane na różnych poziomach podziału terytorialnego. Wyrażam przekonanie, że stanowiąc będą istotne wzbogacenie zasobów informacyjnych wykorzystywanych do analiz regionalnych oraz przy podejmowaniu decyzji dotyczących sytuacji na lokalnych rynkach pracy.

Tadeusz Toczyński

*Generalny Komisarz Spisowy
Prezes Głównego Urzędu Statystycznego*

Warszawa, luty 2004 r.

Od autorów

W roku 2002 przeprowadzony został Narodowy Spis Powszechny, który m.in. dostarczył danych o aktywności ekonomicznej ludności. Dane te mogą stanowić ważne źródło informacji służących do poprawy stochastycznych własności stosowanych dotychczas estymatorów. W związku z tym powstaje problem wykorzystania w procesie estymacji danych NSP 2002 do otrzymywania wiarygodnych ocen parametrów dla województw, podregionów (subregionów) i powiatów.

Istotne to jest zwłaszcza dla tych dziedzin statystyki, które dostarczają informacji niezbędnych do podejmowania decyzji z zakresu polityki gospodarczej i społecznej na szczeblu regionalnym. W dużym stopniu dotyczy to statystyki zatrudnienia i bezrobocia. Dysponowanie szczegółowymi danymi o aktywności ekonomicznej ludności na różnych poziomach agregacji niezbędne jest do prowadzenia analiz dotyczących rynku pracy. Zastosowanie w takich przypadkach klasycznych metod estymacji prowadzi na ogół do tak dużych błędów szacunku, że otrzymane oceny parametrów nie nadają się do praktycznego wykorzystania. Realizowane przez GUS, od 1993 roku, badania aktywności ekonomicznej ludności pozwalają na uzyskiwanie szczegółowych informacji o bezrobociu i zatrudnieniu dla całego kraju oraz w pewnym zakresie dla województw. Potrzeba uzyskiwania podstawowych informacji o bezrobociu i zatrudnieniu na poziomie podregionów i powiatów spowodowała konieczność odwołania się do metod estymacji adekwatnych dla *statystyki małych obszarów*.

W publikacji przedstawiono efekty prac nad dezagregacją danych dotyczących aktywności ekonomicznej ludności na poziom województw, podregionów i powiatów za lata 1995—2002. Opisane zostały zastosowane metody estymacji właściwe dla statystyki małych obszarów tj. estymator syntetyczny oraz estymator złożony będący kombinacją estymatora klasycznego i syntetycznego, własności tych estymatorów oraz metodę szacowania ich precyzji. Przedstawiony został również opis schematu losowania próby w badaniu aktywności ekonomicznej ludności oraz stosowana w tym badaniu metoda estymacji. Zagadnienia te są o tyle ważne, że badania te, obok NSP 2002, są źródłem danych ocen parametrów dla województw, podregionów i powiatów.

Zawarte w aneksie publikacji tablice przedstawiają dla województw i podregionów oszacowanie liczby pracujących, bezrobotnych i biernych zawodowo w korelacji z takimi cechami jak płeć, wiek i wykształcenie. Dla powiatów natomiast zamieszczono dane o aktywności ekonomicznej wg płci i miejsca zamieszkania, a w przypadku pracujących również według statusu zatrudnienia i sektorów ekonomicznych. Podane w aneksie wyniki obliczeń

dotyczą, ze względu na ograniczoną objętość publikacji, tylko roku 2002. Pełny zestaw danych obejmujący lata 1995–2002 wraz ze wskaźnikami zgodności i współczynnikami zmienności dostępny jest w załączonej do publikacji płycie CD.

Autorzy opracowania wdzięczni będą za krytyczne uwagi dotyczące zarówno zastosowanych metod jak i zakresu wykonanej pracy. Uwagi te wykorzystane zostaną w dalszych pracach związanych z dezagregacją danych, na mniejsze podzbiorowości, pochodzących z badania aktywności ekonomicznej ludności.

Zakład Badań
Statystyczno-Ekonomicznych
GUS i PAN

Spis treści

1. Wprowadzenie	7
1.1. Uwagi wstępne	7
1.2. Cel zakres opracowania	8
2. Schemat losowania próby oraz uogólnianie wyników badania	10
2.1. Koncepcja doboru próby	10
2.2. Schemat losowania próby dla okresu od II kwartału 1993 do I kwartału 1999 ...	11
2.3. Metoda estymacji parametrów i oceny precyzji	12
2.4. Zmiany w schemacie losowania wprowadzone w 1999 roku	15
3. Estymacja parametrów przy wykorzystaniu danych z NSP 2002	19
3.1 Szacowanie parametrów dla podregionów i województw	19
3.2. Metoda szacowania precyzji estymatorów	31
4. Zgodność estymatorów klasycznych i wykorzystujących dane z NSP 2002 ...	42
5. Uwagi końcowe	53
6. Bibliografia	56
7. Aneks	59
7.1. Opis tablic.....	59
7.2. Tablica 0. Aktywność ekonomiczna ludności w wieku 15 lat i więcej w latach 1995—2002 wg województw i podregionów.....	61
7.3. Płyta CD z pełnymi danymi dla lat 1995—2002	

1. Wprowadzenie

1.1. Uwagi wstępne

Pełne i wszechstronne rozpoznanie aktywności ekonomicznej ludności na poziomie województw, podregionów i powiatów jest istotną przesłanką przy podejmowaniu wielu decyzji gospodarczych, a także z zakresu polityki społecznej, zarówno na poziomie ogólnokrajowym, jak i regionalnym. Dostępność danych statystycznych dotyczących pracujących i bezrobotnych według szczegółowych przekrojów terytorialnych jest bardzo istotna przede wszystkim z punktu widzenia analizy sytuacji na rynku pracy. Prowadzenie pogłębionych analiz jest uwarunkowane posiadaniem szczegółowych informacji o pracujących i bezrobotnych według różnorodnych cech, w tym demograficznych, społeczno-zawodowych i ekonomicznych w różnych przekrojach terytorialnych (NTS2, NTS3, NTS4).

Możliwość dezagregacji terytorialnej danych w zakresie rynku pracy zależna jest od źródeł pozyskiwania danych. Podstawowym źródłem informacji jest Badanie Aktywności Ekonomicznej Ludności (BAEL). Badanie to jest najważniejszym badaniem reprezentacyjnym prowadzonym w GUS. Metodologia tego badania musi być zgodna z Regulacją Rady Wspólnot Europejskich Nr 3711/91 z dnia 16 grudnia 1991r. (Official Journal of the European Communities No L 351/1 z 20.12.1991) zastąpioną następnie Regulacją Rady Unii Europejskiej (OJ L 77/3 , 14.03.1998 r.). Prowadzenie badań rynku pracy (siły roboczej) jest obligatoryjne dla rządów wszystkich krajów członkowskich UE, przy czym metodologia tych badań określona jest w głównych punktach odpowiednimi artykułami wspomnianej wcześniej Regulacji. Z tego wynika, że GUS nie może dowolnie dobrać sobie metodologii BAEL, lecz musi się poddać rygorom narzuconym przez UE.

Ostatnia regulacja UE wymusiła zmiany w schemacie losowania próby do tego badania. Dlatego też od IV kwartału 1999 r. omawiane badanie prowadzone jest metodą ciągłą tj. przez wszystkie 13 tygodni w kwartale, a nie jak poprzednio tylko w środkowym tygodniu kwartału. Ponadto, alokacja próby pomiędzy województwa nie jest, jak poprzednio, proporcjonalna, ponieważ uznano za niezbędne uzyskiwanie niektórych ocen parametrów także na poziomie województw (tj. NTS 2). Próba do tego badania stanowi około 0,2% ogółu gospodarstw domowych. Pozwala to na wyodrębnienie z ogólnej liczby ludności w wieku 15 lat i więcej trzech subpopulacji ze względu na rodzaj aktywności ekonomicznej, tj. pracujących, bezrobotnych i biernych zawodowo. Metoda badania warunkuje stopień szczegółowości wyników — pozyskiwane są dane dla całego kraju oraz w podstawowym zakresie dla województw. Drugim źródłem informacji jest Narodowy Spis Powszechny Ludności i Mieszkań (w dalszej części publikacji będziemy używać skrótu: NSP 2002), który był przeprowadzony w 2002 roku. Wartość tego badania, jako badania pełnego (powszechnego) jest ogromna. Fakt, iż do zbadania aktywności ekonomicznej ludności w NSP 2002 zastosowano takie same definicje, jak w BAEL, daje możliwości wykorzystania obydwu źródeł informacji do opraco-

wania metody rozszacowywania danych z BAEL na niższe poziomy podziału terytorialnego kraju, niż to jest możliwe tylko na podstawie samego BAELu. Opracowanie metody estymacji danych, a następnie dokonanie szacunków, pozwoli na utworzenie bazy informacyjnej o ludności aktywnej zawodowo (pracujący, bezrobotni) oraz biernej zawodowo według różnych cech, zawierającej szacunkowe dane za lata 1995—2002. Będzie ona corocznie uzupełniana informacjami przygotowywanymi zgodnie z opracowaną metodologią, z uwzględnieniem wyników bieżących badań statystycznych. Uzyskane szeregi czasowe będą wykorzystywane m.in. w bieżących analizach dotyczących zasobów pracy, a także w pracach prognostycznych.

Należy zdawać sobie sprawę z tego, że informacje pochodzące z NSP 2002 będą się dezaktualizowały. Powstaje zatem problem zbadania szybkości tej dezaktualizacji. Jest rzeczą oczywistą, że nie można tego zagadnienia zrobić *ex ante*. Możliwe jest jednak zrobienie ekstrapolacji wstecz z wykorzystaniem wyników NSP 2002 i porównanie otrzymanych w ten sposób wyników z tymi, które zostały pozyskane wcześniej. Względne różnice jednych i drugich danych dadzą orientacyjną wiedzę o tym, jak długo będzie można wykorzystywać dane spisowe przy estymacji parametrów dla powiatów i podregionów.

W kolejnym etapie realizowanym w czasie, gdy dane z NSP 2002 będą dawały zadowalające wyniki, należy podjąć prace studialne nad możliwością i sposobem wykorzystania informacji pochodzących, między innymi, z rejestrów bezrobocia prowadzonych przez powiatowe biura pracy oraz innych źródeł (np. wszelkiego typu rejestrów dostępnych w GUS). Prace te będą wymagały zbadania własności różnych estymatorów dla małych obszarów. W szczególności ich obciążenia i precyzji.

1.2. Cel i zakres opracowania

W latach od 1995 do 1999 (I kwartał) stosowano schemat losowania próby właściwy dla realizacji badania w środkowym tygodniu kwartału. W II i III kwartale 1999 roku badanie nie było realizowane, po czym od IV kwartału badanie realizowane jest metodą ciągłą z tygodniową rotacją próby. Schemat ten stosowany jest do chwili obecnej. Parametry populacji będą szacowane z wykorzystaniem danych z NSP 2002 roku według metod odpowiednich do zastosowanych schematów losowania w poszczególnych okresach.

Podstawową jednostką terytorialną, dla której będą szacowane parametry będzie powiat. Ze względu na to, że liczebność próby w wielu powiatach jest zerowa lub niewielka, nie ma możliwości zastosowania klasycznych metod estymacji. W związku z tym, powiaty potraktowane zostaną jako „małe obszary” i przy estymacji parametrów dla tych jednostek zostaną wykorzystane estymatory syntetyczne. Estymatory te są, w przybliżeniu, tak samo efektywne jak uogólnione estymatory regresyjne. Pokazały to liczne badania symulacyjne (por. np. Cz. Bracha (1993, 1996)). Ponadto trudności w kojarzeniu danych z badań aktywności ekonomicznej z danymi z NSP na poziomie rejonów statystycznych i obwodów spisowych spowodowały konieczność ograniczenia się do estymatorów syntetycznych. Jako cechy do-

datkowe zostaną użyte odpowiednie wartości pochodzące z NSP 2002 r. W przypadkach, w których liczby jednostek pierwszego stopnia (jps) będą dostatecznie duże, zostaną zastosowane estymatory klasyczne (zbliżone do estymatorów Horwitza-Thompsona, z tą jednak różnicą, że wagi będą zmiennymi losowymi). Rozważone zostaną również estymatory złożone (ang. composite estimators (porównaj R. Griffiths (1996))) będące liniową, wypukłą kombinacją estymatorów klasycznych i syntetycznych. Oczywiście te estymatory będą mogły być stosowane w tych warunkach, w których sens będzie miało zastosowanie estymatorów klasycznych.

Bazą szacunków dla powiatów będą oceny parametrów dla województw uzyskane metodami klasycznymi. Dla większych powiatów (przede wszystkim dla miast-powiatów) przewiduje się dodatkowo szacowanie parametrów metodami klasycznymi.

W odniesieniu do podregionów traktowanych jako „małe obszary” rozważone zostaną dwie metody estymacji. Pierwsza metoda to estymatory klasyczne. W drugiej metodzie bazą szacunków dla podregionów będą oceny uzyskane dla województw metodami klasycznymi, zaś cechami dodatkowymi będą odpowiednie dane z NSP 2002 r.

Dla województw zastosowane zostaną podobne metody szacunku jak dla podregionów. Dokładniej mówiąc: po pierwsze oceny dla województw będą sumami ocen dla wchodzących w ich skład podregionów, po drugie oceny dla województw uzyskiwane będą metodami tradycyjnymi.

W przypadku stosowania estymatorów syntetycznych oceny dla kolumn zbiorczych w tablicach wynikowych stanowią będą sumy ocen częściowych. Tym samym uniknie się niezgodności pomiędzy ocenami na różnych poziomach agregacji.

Dla wszystkich ocen parametrów szacowana będzie precyzja. Zastosujemy w odniesieniu do badanych lat tę samą metodę — tzw. bootstrap. Należy ona do grupy nowoczesnych i efektywnych metod symulacyjnych, co potwierdzają liczne publikacje naukowe. W praktycznej realizacji metod symulacyjnych przy szacowaniu wariancji estymatorów dla wielostopniowych schematów losowania dokonuje się zazwyczaj uprzedniej agregacji danych na poziom jednostek losowania pierwszego stopnia (jps). Takie podejście, zmniejszające istotnie czasochłonność obliczeń, zastosowano również w omawianych analizach danych z BAEL. Uwzględniono również fakt warstwowania jednostek pierwszego stopnia definiując warstwy w sposób jak najbardziej odpowiadający warstwom zastosowanym w badaniu w poszczególnych analizowanych latach. Dla roku 1999 precyzja oszacowań będzie mniejsza niż dla innych lat, ponieważ badane były tylko dwa kwartały.

Dokładny opis metody estymacji parametrów oraz szacowania precyzji zostanie przedstawiony w dalszej części opracowania. W publikacji przedstawiona zostanie analiza efektywności estymatorów syntetycznych oraz estymatorów złożonych (porównaj R. Griffiths. (1996)) i klasycznych dla większych powiatów oraz wszystkich podregionów i województw.

Szacunki dotyczące aktywności ekonomicznej bazują na danych pochodzących z badań aktywności ekonomicznej ludności za lata 1993—2002. W związku z tym, uważamy za niezbędne przedstawienie opisu schematu losowania próby oraz metody estymacji parametrów stosowany w tych badaniach.

2. Schemat losowania próby oraz uogólnianie wyników badania¹⁾

2.1. Koncepcja doboru próby

W okresie 1993—2002 rok próba do BAEL była losowana według dwóch istotnie różniących się schematów (o czym wspominaliśmy wcześniej). Rok 1999 był przełomowy. W tymże roku bowiem GUS dostosował schemat losowania do nowych zaleceń Unii Europejskiej (pisaliśmy o tym fakcie w pierwszym rozdziale). Należy zwrócić uwagę na jeszcze jeden fakt. Od pierwszego stycznia 1999 roku zaczął obowiązywać nowy podział administracyjny kraju (zamiast poprzednich 49 województw utworzono 16 i wprowadzono powiaty). Fakt ten również wpłynął na schemat losowania próby. Ponadto, dane dla lat 1995—1998 zostały tak przeliczone, aby odpowiadały podziałowi terytorialnemu kraju obowiązującemu w NSP 2002 r. Badanie aktywności ekonomicznej ludności przeprowadzane jest metodą reprezentacyjną. Umożliwia ona uogólnienie wyników badania na populację generalną. Autorem schematów losowania, metod estymacji parametrów oraz precyzji szacunków był A. Szarkowski (1981, 2002)

Badanie aktywności ekonomicznej ludności realizowane jest na próbie wylosowanej przy zastosowaniu schematu losowania dwustopniowego, warstwowego z różnymi prawdopodobieństwami wyboru jednostek I-szego stopnia. Jednostkami losowania I-szego stopnia są rejony statystyczne w miastach oraz obwody spisowe na terenach wiejskich. Na drugim stopniu losowane są mieszkania, zaś jednostkami badania są gospodarstwa domowe i osoby w wieku 15 lat i więcej.

Celem badania jest zebranie informacji o aktywności ekonomicznej ludności w wieku 15 lat i więcej dla określonego kwartału. Każda próba kwartalna pierwszego stopnia składa się z czterech tzw. prób elementarnych, przy czym co kwartał dokonuje się częściowej wymiany prób elementarnych. Do ankietowania w danym kwartale przeznaczają się dwie próby elementarne badane w kwartale poprzednim, jedną próbę elementarną nowowprowadzoną do badania oraz jedną próbę elementarną niebadaną w kwartale poprzednim, a która została wprowadzona do badania dokładnie przed rokiem. Poszczególne próby elementarne losowane są w sposób wzajemnie niezależny; losując daną próbę elementarną nie bierze się zupełnie

¹⁾ Autorem schematu losowania i metody estymacji był A. Szarkowski. Prawa autorskie do tych rozwiązań posiada Departament Statystyki Społecznej, który umożliwił nam zamieszczenie w opracowaniu rozwiązań A. Szarkowskiego.

pod uwagę wyników losowania innych prób. W efekcie tego schematu, każda próba elementarna (w wariancie aktualnym od lutego 1994 r) aktualnie używana jest wg zasady 2-(2)-2: dwa kwartały w badaniu, dwa kwartały przerwy, znów dwa kwartały w badaniu. Szczegółowe informacje dotyczące wykorzystania poszczególnych prób podaje załączona Karta Rotacji na lata 1995—1999 (I kwartał).

2.2. Schemat losowania próby dla okresu od II kwartału 1993 do I kwartału 1999

Pojedyncze próby elementarne losowano do BAEL przy wykorzystaniu schematu losowania dwustopniowego. Jednostkami losowania pierwszego stopnia (jps) zarówno w miastach jak i na wsi były rejony statystyczne. Począwszy od maja 1997 roku na wsiach z przyczyn organizacyjnych jako jednostki pierwszego stopnia wykorzystywane są obwody spisowe. Rejony i obwody muszą spełniać określone wymogi odnośnie minimalnej liczby mieszkań. Jeżeli wymogi te nie są spełnione to jednostki losowania tworzy się łącząc ze sobą dwa lub więcej sąsiednich rejonów lub obwodów. Jednostkami losowania drugiego stopnia (jds) są mieszkania. Badaniu podlegają wszystkie osoby w wieku 15 lat i więcej zamieszkałe w mieszkaniach wylosowanych.

Początkowo wszystkie mieszkania w próbach elementarnych losowane były z prawdopodobieństwem wyboru równym 1/2000. Począwszy od maja 1997r. do BAEL wprowadza się próby elementarne, do których mieszkania miejskie losowane są z prawdopodobieństwem wyboru równym 1/2000, a mieszkania wiejskie z prawdopodobieństwem wyboru 1/1818. Związane to jest ze zmianą jednostki losowania pierwszego stopnia na terenach wiejskich z rejonu statystycznego na obwód spisowy. Zmiana taka mogła spowodować mniejszą efektywność schematu losowania, co zrekompensowane zostało zwiększeniem liczebności próby.

Losowanie jednostek pierwszego stopnia oraz mieszkań przeprowadza się z operatu utworzonego na podstawie rejestru podziału terytorialnego kraju TERYT prowadzonego przez Główny Urząd Statystyczny i Wojewódzkie Urzędy Statystyczne. Rejestr ten składa się m.in. z systemu BRE10 tj. zbioru rejonów statystycznych i obwodów spisowych oraz z systemu NOBC10 tj. z systemu identyfikacji adresowej ulic, nieruchomości i mieszkań. System TERYT aktualizowany jest corocznie na stan 1 stycznia. Proces aktualizacji trwa jednak ok. 10 miesięcy, tak że nowy rejestr mieszkań wykorzystywany był do losowania prób nowoprowadzanych dopiero w czwartym kwartale danego roku. Przy losowaniu prób elementarnych pomija się więc z konieczności mieszkania z najnowszego budownictwa.

Jednostki pierwszego stopnia przed losowaniem były warstwowane wg województw. Wewnątrz województw tworzono warstwę wiejską oraz od 2 do 5 warstw miejskich, w zależności od województwa. Warstwy miejskie wyróżniano ze względu na wielkość miasta. Przy losowaniu prób elementarnych wprowadzanych w okresie od listopada 1993 r. do sierpnia 1994 r. wyróżniono ogółem 150 warstw miejskich.

Na drugim stopniu, z pojedynczego obwodu wiejskiego losowano do próby ok. 8 mieszkań. Dla rejonów miejskich, szczególnie w miastach większych, liczbę mieszkań losowanych z pojedynczego rejonu starano się wyznaczyć na niższym poziomie tak, aby zwiększyć liczbę rejonów losowanych do próby. W każdej z warstw losowanie mieszkań przeprowadzano tak, aby ostateczne prawdopodobieństwo wyboru do próby elementarnej pojedynczego mieszkania było stale równe $1/2000$ dla miast, a $1/1818$ dla wsi.

Pojedyncza próba elementarna wprowadzana w okresie od listopada 1993 r. do sierpnia 1994 r. liczyła 908 jednostek pierwszego stopnia, a wartość oczekiwana liczby mieszkań w próbie elementarnej wynosi 5714,3. W okresie późniejszym wartości te uległy zwiększeniu tak, że w lutym 1999 roku pojedyncza próba elementarna liczyła około 1000 jps i średnio 6000 mieszkań. W miastach powyżej 100.000 mieszkańców wypadało średnio 4,98 mieszkania na rejon, w pozostałych miastach — 6,67, a na wsi 8,02.

Losowanie jednostek pierwszego stopnia w poszczególnych warstwach przeprowadzano z prawdopodobieństwami wyboru proporcjonalnymi do liczby mieszkań w jednostce metodą Hartleya-Rao. Polega ona na doborze systematycznym jednostek, po ich uprzednim losowym uporządkowaniu. Faktyczna liczba mieszkań losowanych z pojedynczej jednostki pierwszego stopnia ustalana była randomizacyjnie tak, aby jej wartość oczekiwana była równa wyliczonym liczbom teoretycznym. Po ustaleniu tej liczby, wyboru mieszkań dokonywano metodą losowania prostego.

2.3. Metoda estymacji parametrów i oceny precyzji

Metoda uogólniania wyników wzorowana jest w pewnej mierze na amerykańskim badaniu Current Population Survey i uwzględnia następujące okoliczności:

- 1) prawdopodobieństwa wyboru mieszkań (w miastach — $1/2000$, na wsi — $1/1818$),
- 2) poziom realizacji (kompletności) badania ze względu na klasę miejscowości,
- 3) stratyfikację ex post ze względu na płeć, grupy wieku oraz w podziale na miasto i wieś.

Krokiem wstępnym jest wyliczenie tzw. współczynników realizacji R_k przy czym $k = 1, 2, \dots, 6$. Oblicza się je według wzoru

$$(2.3.1) \quad R_k = \frac{Z_k}{Z_k + B_k},$$

gdzie:

Z_k — oszacowana liczba mieszkań zbadanych w k -tej klasie miejscowości,

B_k — oszacowana liczba mieszkań niezbadanych w k -tej klasie miejscowości, które powinny być zbadane.

Łączna suma $Z_k + B_k$ jest liczbą mniejszą od liczby mieszkań wybranych do próby, bowiem wyliczając B_k bierze się pod uwagę jedynie mieszkania niezbadane z powodu odmowy lub krótkotrwałej (do 2 miesięcy) nieobecności mieszkańców, natomiast pomija mieszkania zlikwidowane, zamienione w obiekty niemieszkalne, niezamieszkałe lub zamknięte z powodu długotrwałej (ponad 2 miesiące) nieobecności mieszkańców. Współczynniki realizacji wyliczamy w sześciu grupach wyróżnionych ze względu na klasę miejscowości. W badaniu lutowym 1999 roku wyniosły one (z dokładnością do trzech cyfr po przecinku):

— 0,614 dla Warszawy,

— 0,774 dla pozostałych miast pow. 500 tys. mieszkańców tj.: Kraków, Łódź, Poznań, Wrocław oraz trójmiasto Gdańsk-Gdynia-Sopot,

— 0,785 dla pozostałych miast pow. 100 tys. mieszkańców,

— 0,863 dla miast 20—100 tys. (bez Sopotu),

— 0,901 dla pozostałych miast,

— 0,951 dla wsi.

Dla Polski współczynnik ten był równy $R = 0,856$.

Następnie na podstawie danych BAEL szacowana jest liczba osób wg płci, 12 grup wieku oraz w podziale na miasta i wieś. Grupy wieku stosowane przy wyliczaniu wag są takie same, jak w publikowanych tablicach wynikowych tj.: 15—17 lat, 18—19, 20—24, 25—29, 30—34, 35—39, 40—44, 45—49, 50—54, 55—59, 60—64, 65 i więcej lat. Łącznie szacuje się liczbę osób dla 48 grup.

Liczbę osób w j -tej grupie \hat{g}_j ($j=1,2,\dots,48$) na podstawie danych z BAEL szacuje się ze wzoru:

$$(2.3.2) \quad \hat{g}_j = \sum_i \frac{F}{R_k} x_{ij},$$

gdzie:

F — odwrotność frakcji losowania mieszkań (dla miast $F = 2000$, dla wsi $F = 818$),

R_k — współczynnik realizacji dla k -tej klasy miejscowości,

x_{ij} — liczba osób przynależnych do j -tej grupy zbadanych w BAEL w i -tym mieszkaniu.

Następnie na podstawie bieżących szacunków demograficznych wylicza się liczebności populacyjne wyżej zdefiniowanych 48 grup ludności tj G_j ($j=1,2,\dots,48$). Wyliczone dla tych grup liczebności nie obejmują ludności poza gospodarstwami domowymi, ponieważ nie jest ona z założenia badana w BAEL. Dotyczy to osób z gospodarstw zbiorowych (łącznie z wojskiem skoszarowanym) i osób przebywających czasowo za granicą.

Dzieląc wartości populacyjne dla poszczególnych grup przez ich oszacowania na podstawie próby otrzymujemy mnożniki niezbędne w stratyfikacji ex post tj.

$$(2.3.3) \quad M_j = \frac{G_j}{\hat{g}_j} \quad (j=1,2,\dots,48).$$

Następnie obliczamy wagę finalną dla osoby z i -tego mieszkania przynależną do j -tej grupy ze względu na płeć i wiek oraz k -tą klasę miejscowości. Waga ta ma postać:

$$(2.3.4) \quad W_{ijk} = \frac{F}{R_k} M_j.$$

W przypadku mieszkań i gospodarstw domowych we wzorze na wagę finalną pominięty jest mnożnik M_j .

KARTA ROTACJI NA LATA 1995—1999 (I kwartał)

Numer próby	Lata i kwartały																			
	1995				1996				1997				1998				1999			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
8	x																			
9	x	x																		
10	-	x	x																	
11	-	-	x	x																
12	x	-	-	x	x															
13	x	x	-	-	x	x														
14		x	x	-	-	x	x													
15			x	x	-	-	x	x												
16				x	x	-	-	x	x											
17					x	x	-	-	x	x										
18						x	x	-	-	x	x									
19							x	x	-	-	x	x								
20								x	x	-	-	x	x							
21									x	x	-	-	x	x						
22										x	x	-	-	x	x					
23											x	x	-	-	x	x				
24												x	-	-	x	x				x
25													x	x	-	-	x			x
26														x	x	-	-			
27															x	x	-			
28																x				x
29																				x

2.4. Zmiany w schemacie losowania wprowadzone w 1999 roku

Od IV kwartału 1999 r. w BAEL wprowadzone zostały zmiany w metodzie badania i alokacji próby. Od tego czasu badanie prowadzone jest w sposób ciągły. Oznacza to, że w każdym z 13 tygodni danego kwartału¹⁾ ankieterzy odwiedzają określoną liczbę (aktualnie 1900) losowo wybranych mieszkań i zbierają dane o aktywności ekonomicznej w tygodniu poprzednim. Badaniem objęte są wszystkie osoby w wieku 15 lat i więcej, zamieszkałe w wylosowanych mieszkaniach. Próbka mieszkań przeznaczonych do odwiedzin jest zmieniana z tygodnia na tydzień.

¹⁾ Stosowane obecnie w BAEL pojęcie kwartału różni się nieco od kwartału kalendarzowego: każdy kwartał BAEL liczy 13 tygodni i jego początek wypada zawsze w poniedziałek. Tak na przykład pierwszy kwartał 2000 r. trwał od 3 stycznia do 3 kwietnia.

Próbki tygodniowe otrzymuje się z losowego podziału na 13 części próby kwartalnej, liczącej obecnie 24.700 mieszkań. Próba kwartalna została skonstruowana tak, aby każda z 13 próbek tygodniowych miała nie tylko jednakową wielkość, ale i jednakową strukturę.

Wyniki badania są opracowywane i publikowane w ujęciu kwartalnym. Z niewielkim uproszczeniem można powiedzieć, że wyniki kwartalne wyliczane są jako średnie wyników z 13 tygodni danego kwartału.

Dobór prób kwartalnych odbywa się wg zasad tzw. schematu rotacyjnego, który pozostaje niezmienny od drugiego kwartału 1993 roku, przy czym każda z prób elementarnych na dany kwartał, ze względu na metodę ciągłą, dzielona jest na 13 tygodniowych próbek elementarnych. Załączona „Karta rotacji na lata 1999—2004” podaje szczegóły rotacji kwartalnej.

Podobnie jak w latach poprzednich, losowanie prób do BAEL przebiega wg zasad losowania dwustopniowego. Jednostkami losowania pierwszego stopnia są w miastach rejony statystyczne, a na wsi obwody spisowe (w bardzo rzadkich przypadkach jednostki losowania tworzy się łącząc ze sobą dwa lub więcej sąsiednich rejonów lub obwodów). Jednostkami losowania drugiego stopnia są mieszkania. Jednostki pierwszego stopnia (jps) losowane są z zastosowaniem tzw. warstwowania. Podstawą podziału na warstwy jest podział na województwa. Ustalając liczbę mieszkań losowanych z poszczególnych województw, inaczej niż poprzednio, dla poprawienia precyzji wyników dla województw mniejszych (o mniejszej liczbie ludności i mieszkań) zastosowano tzw. nadreprezentowanie województw małych połączone z tzw. niedoreprezentowaniem województw dużych. I tak np. województwo mazowieckie czy śląskie liczą pięciokrotnie więcej mieszkań niż lubuskie czy opolskie, ale z tych pierwszych losuje się do próby BAEL tylko 2,25 razy tyle mieszkań, co z tych drugich¹⁾.

Liczba mieszkań losowanych z poszczególnych województw do tygodniowej próbki elementarnej waha się od 20 do 45, zaś liczba losowanych jps od 3 do 7. Zmiana alokacji próby pomiędzy województwa spowodowała konieczność wprowadzenia zmian w podziale województw na warstwy. Wewnątrz województw wyróżniono 3 względnie 4 warstwy, jedynie w woj. wielkopolskim utworzono 5 warstw, zaś w woj. śląskim udało się utworzyć 7 warstw. Warstwy wewnątrz województw utworzono ze względu na wielkość miejscowości, zaliczając wsie do najmniejszych. Przy tworzeniu warstw uwzględniana była specyfika danego województwa. Nie było sztywnego kryterium podziału. Łącznie utworzono 61 warstw.

¹⁾ Liczbę mieszkań w próbie losowaną z poszczególnych województw ustalono w oparciu o często stosowaną w tego rodzaju sytuacjach praktyczną zasadę, że liczebności prób powinny być w przybliżeniu proporcjonalne do **pierwiastka kwadratowego** z liczebności populacyjnych.

Losowanie jps wewnątrz warstw przeprowadzono ze zróżnicowanymi prawdopodobieństwami wyboru, proporcjonalnymi do liczby mieszkań w jps. Z każdej jps wybranego do próby losowano następnie ustaloną liczbę mieszkań (od 4 do 9). Przy ustalaniu tych liczb starano kierować się zasadą, aby (dla pogodzenia względów organizacyjnych i wymogu uzyskania możliwie najlepszej precyzji) losować po 8 mieszkań z jps pochodzących ze wsi i małych miast, po 6—7 mieszkań z jps z miast średnich i po 5 mieszkań z jps wielkomiejskich. Często jednak trzeba było robić wyjątki od tej zasady (np. po 6 mieszkań z jps w Warszawie, po 4 z jps w utworzonym *ad hoc* zespole Poznań-Swarzędz-Luboń, a po 7 mieszkań z jps we wsiach i małych miastach woj. łódzkiego).

Metoda uogólniania wyników pozostała bez zmian. Jedyne wagi F wynikające z prawdopodobieństw wyboru są zróżnicowane i ich wartość zależy od województwa. Wprowadzenie do BAEL metody ciągłej spowodowało zwiększenie trudności w realizacji badania i między innymi znaczny w skali kraju wzrost liczby odmów. Dla czwartego kwartału 2002 r współczynniki realizacji R wyniosły:

— 0,797 dla całej Polski,

— 0,520 dla Warszawy,

— 0,715 dla miast o liczbie mieszkańców od 500 tys. do 1 mln (Kraków, Łódź, Poznań, Wrocław wraz z trójmiastem Gdańsk-Gdynia-Sopot),

— 0,763 dla miast o liczbie mieszkańców od 100 do 500 tys.,

— 0,784 dla miast o liczbie mieszkańców od 20 do 100 tys.,

— 0,816 dla pozostałych miast,

— 0,892 dla wsi.

Procedura estymacji stosowana w BAEL doprowadza do złożonych „wielopiętrowych” estymatorów, zaliczających się do typu estymatorów ilorazowych. Wariancje takich estymatorów nie dają się szacować klasycznymi metodami podręcznikowymi i trzeba stosować pewne specjalne metody przybliżone. W nowej edycji BAEL zdecydowano się na użycie *metody linearyzacji*, zwanej także *metodą szeregu Taylora*. Metoda ta polega na zastąpieniu estymatora złożonego, rozpatrywanego jako funkcja prostych estymatorów podręcznikowych, *liniową* funkcją tych estymatorów prostych, otrzymaną ze znanego wzoru Taylora. Nieco informacji o tej metodzie można znaleźć np. w ustępie 11.18 podręcznika W.G. Cochrańa *Sampling Techniques*, 1977. Więcej informacji na ten temat można znaleźć w monografii K.M. Woltera (1985).

KARTA ROTACJI NA LATA 1999—2004

Numer próby	Lata i kwartały																												
	1999				2000				2001				2002				2003				2004								
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4					
1			x																										
2			x	x																									
3			—	x	x																								
4			—	—	x	x																							
5			x	—	—	x	x																						
6			x	x	—	—	x	x																					
7				x	x	—	—	x	x																				
8					x	x	—	—	x	x																			
9						x	x	—	—	x	x																		
10							x	x	—	—	x	x																	
11								x	x	—	—	x	x																
12									x	x	—	—	x	x															
13										x	x	—	—	x	x														
14											x	x	—	—	x	x													
15													x	x	—	—	x	x											
16														x	x	—	—	x	x										
17															x	x	—	—	x	x									
18																x	x	—	—	x	x								
19																	x	x	—	—	x	x							
20																		x	x	—	—	x	x						
21																			x	x	—	—	x	x					
22																				x	x	—	—	x	x				
23																					x	x	—	—	x	x			
24																						x	x	—	—	x	x		
25																							x	x	—	—	x	x	
26																								x	x	—	—	x	x

Próba 1, 2, 3, 4, 5
Próba 6 i dalsze

— próby o skróconym okresie badania.
— próby o normalnym okresie badania.

3. Estymacja parametrów przy wykorzystaniu danych z NSP 2002

3.1 Szacowanie parametrów dla podregionów i województw

Parametry dla województw i podregionów szacowane były metodami klasycznymi tj. takimi samymi jakie, stosuje się do tej pory w BAEL. Estymator parametru dla województwa i podregionu oznaczmy przez t_w , odpowiednio. Niezależnie od tego dla podregionów i województw zastosowano estymatory syntetyczne określone następującymi wzorami:

— dla podregionów (subregionów)

$$(3.1.1) \quad x_{ws} = t_w f_{ws},$$

gdzie f_{ws} jest udziałem wartości danej zmiennej (z NSP 2002) w s -tym podregionie w stosunku do w -tego województwa,

— dla województw

$$(3.1.2) \quad x_w = t f_w,$$

gdzie t jest estymatorem klasycznym dla całego kraju, natomiast f_w jest udziałem wartości danej zmiennej (z NSP 2002) w w -tym województwie w stosunku do całego kraju.

Trzecim rozpatrywanym typem estymatorów będzie estymator złożony zaproponowany przez R. Griffiths'a (1996). Jest on liniową wypukłą kombinacją estymatorów klasycznego i syntetycznego:

$$(3.1.3) \quad y_{ws} = v_{ws} t_{ws} + (1 - v_{ws}) x_{ws},$$

gdzie v_{ws} jest wagą przypisaną estymatorowi klasycznemu. Wielkość v_{ws} jest liczbą z przedziału (0;1). Jej optymalna wartość zależy od błędów średniokwadratowych estymatorów występujących w powyższym wzorze. Jest rzeczą oczywistą, że $MSE(t_{ws})$ oraz $MSE(x_{ws})$ są nieznane i powinny być oszacowane na podstawie próby. Ponieważ próby dla małych obszarów są niewielkie, to oceny wymienionych wyżej błędów średniokwadratowych obarczone byłyby dużymi błędami losowymi, co sprawiłoby, że wyznaczona waga v_{ws} mogłaby znacznie różnić się od optymalnej. Ponadto dla każdego szacowanego parametru należałoby taką wagę obliczać indywidualnie. Spowodowałoby to niesamowite zwiększenie pracochłonności obliczeń. Dlatego też zdecydowaliśmy się przyjąć $v_{ws}=0,5$. Na zakończenie tego akapitu chcielibyśmy dodać, że estymator złożony wydaje się lepszy od syntetycznego, mimo większego błędu średniokwadratowego. Spowodowane jest to tym, że estymator złożony poza informacjami z innych źródeł niż próba, wykorzystuje jednak wiedzę dostarczaną

przez próbę o wartości parametru dla małego obszaru. Tym samym należy oczekiwać, że obciążenie estymatora złożonego będzie znacznie niższe w porównaniu z estymatorem syntetycznym, który w sposób automatyczny przenosi strukturę z badania pełnego dostarczającego informacji o interesującym nas parametrze na małe obszary.

W przypadku powiatów będą stosowane różne metody estymacji w zależności od wielkości powiatu mierzonej liczbą jednostek pierwszego stopnia wylosowanej do próby w danym roku. Poniżej prezentujemy tablicę zawierającą informacje o liczbie jps dla powiatów.

Tablica 3.1. Liczby jps zbadanych z powiatów w latach 1995—2002

Woj.	Podr.	Pow.	Nazwa podregionu (powiatu)	1995	1996	1997	1998	1999	2000	2001	2002	Min.	Max.
02	01		Podreg. Jeleniog.-wałbrzyski	494	512	573	616	239	364	370	378	239	616
02	01	01	Powiat bolesławiecki	35	36	32	27	11	15	16	31	11	36
02	01	02	Powiat dzierzoniowski	46	37	41	42	23	37	29	29	23	46
02	01	05	Powiat jaworski	14	20	24	19	9	7	11	16	7	24
02	01	06	Powiat jeleniogórski	15	14	24	18	9	18	27	22	9	27
02	01	07	Powiat kamiennogórski	13	15	29	37	7	9	5	13	5	37
02	01	08	Powiat kłodzki	39	59	88	106	32	53	55	55	32	106
02	01	10	Powiat lubański	24	23	27	42	9	10	16	10	9	42
02	01	12	Powiat lwówecki	16	12	10	14	4	10	10	16	4	16
02	01	17	Powiat strzeliński	14	15	16	15	7	20	10	4	4	20
02	01	19	Powiat świdnicki	63	62	67	66	29	34	36	53	29	67
02	01	21	Powiat wałbrzyski	31	28	18	22	14	9	19	30	9	31
02	01	24	Powiat ząbkowicki	19	18	21	34	9	26	22	14	9	34
02	01	25	Powiat zgorzelecki	33	30	36	46	11	24	30	26	11	46
02	01	26	Powiat złotoryjski	12	19	23	13	10	15	10	7	7	23
02	01	61	Powiat m. Jelenia Góra	44	47	39	44	20	31	31	26	20	47
02	01	63	Powiat m. Wałbrzych	76	77	78	71	35	46	43	26	26	78
02	02		Podreg. Legnicki	195	191	231	252	83	141	149	142	83	252
02	02	03	Powiat głogowski	29	35	42	43	13	30	31	32	13	43
02	02	04	Powiat górowski	4	3	10	19	5	6	9	11	3	19
02	02	09	Powiat legnicki	19	17	12	14	9	16	17	19	9	19
02	02	11	Powiat lubiąski	48	44	58	55	19	22	22	24	19	58
02	02	16	Powiat polkowicki	20	16	28	38	9	20	22	9	9	38
02	02	22	Powiat wołoski	17	15	22	26	7	12	9	9	7	26
02	02	62	Powiat m. Legnica	58	61	59	57	21	35	39	38	21	61
02	03		Podreg. wrocławski	115	116	154	170	65	103	95	86	65	170
02	03	13	Powiat milicki	15	7	2	3	6	12	15	7	2	15
02	03	14	Powiat oleśnicki	31	29	37	40	18	32	27	14	14	40
02	03	15	Powiat oławski	18	22	35	41	14	11	6	12	6	41
02	03	18	Powiat średzki	13	15	10	7	6	19	16	8	6	19
02	03	20	Powiat trzebnicki	17	17	34	49	13	9	13	26	9	49
02	03	23	Powiat wrocławski	21	26	36	30	8	20	18	19	8	36
02	04		Podreg. m Wrocław	307	319	313	319	184	366	376	370	184	376
02	04	64	Powiat m. Wrocław	307	319	313	319	184	366	376	370	184	376
04	05		Podreg. Bydgoski	384	388	424	454	202	413	401	398	202	454
04	05	03	Powiat bydgoski	32	22	13	25	10	26	32	33	10	33
04	05	07	Powiat inowrocławski	70	84	88	86	31	52	63	61	31	88
04	05	09	Powiat mogileński	11	17	21	12	5	14	12	3	3	21
04	05	10	Powiat nakielski	22	17	19	32	12	47	35	20	12	47
04	05	13	Powiat sępoleński	12	13	21	17	10	17	15	17	10	21
04	05	14	Powiat świecki	18	19	28	42	18	31	24	28	18	42
04	05	16	Powiat tucholski	18	16	18	17	10	15	12	12	10	18
04	05	19	Powiat żniński	21	20	28	32	10	16	17	25	10	32
04	05	61	Powiat m. Bydgoszcz	180	180	188	191	96	195	191	199	96	199
04	06		Podreg. toruńsko-włocławski	392	396	431	461	190	383	380	397	190	461
04	06	01	Powiat aleksandrowski	17	20	22	15	12	17	23	30	12	30

Tablica 3.1. Liczby jps zbadanych z powiatów w latach 1995—2002 (cd.)

Woj.	Podr.	Pow.	Nazwa podregionu (powiatu)	1995	1996	1997	1998	1999	2000	2001	2002	Min.	Max.
04	06	02	Powiat brodnicki	19	27	35	34	12	21	26	24	12	35
04	06	04	Powiat chełmiński	12	13	9	15	8	19	17	15	8	19
04	06	05	Powiat golubsko-dobrzyński	14	13	13	18	6	15	14	23	6	23
04	06	06	Powiat grudziądzki	18	9	8	26	0	8	2	4	0	26
04	06	08	Powiat lipnowski	15	12	26	34	10	23	27	25	10	34
04	06	11	Powiat radziejowski	16	16	24	29	9	22	11	12	9	29
04	06	12	Powiat rypiński	26	24	16	21	8	14	13	7	7	26
04	06	15	Powiat toruński	19	23	38	23	13	19	13	15	13	38
04	06	17	Powiat wąbrzeski	9	10	16	13	4	11	18	22	4	22
04	06	18	Powiat włocławski	23	24	19	24	8	16	28	26	8	28
04	06	62	Powiat m. Grudziądz	50	46	46	48	24	49	43	54	24	54
04	06	63	Powiat m. Toruń	100	103	100	103	49	100	95	100	49	103
04	06	64	Powiat m. Włocławek	54	56	59	58	27	49	50	40	27	59
06	07		Podreg. białkopodlaski	111	112	133	141	54	102	105	110	54	141
06	07	01	Powiat biały	38	37	49	60	20	38	36	36	20	60
06	07	13	Powiat parczewski	6	6	9	4	0	8	8	13	0	13
06	07	15	Powiat radzyński	18	21	25	24	15	30	24	18	15	30
06	07	19	Powiat włodawski	18	16	20	23	10	12	12	16	10	23
06	07	61	Powiat m. Biała Podlaska	31	32	30	30	9	14	25	27	9	32
06	08		Podreg. chełmsko-zamojski	227	230	274	294	111	242	243	211	111	294
06	08	02	Powiat biłgorajski	40	38	45	40	18	25	30	25	18	45
06	08	03	Powiat chełmski	22	24	33	25	17	42	36	41	17	42
06	08	04	Powiat hrubieszowski	30	29	40	34	5	15	22	23	5	40
06	08	06	Powiat krasnostawski	22	27	36	40	10	21	17	24	10	40
06	08	18	Powiat tomaszowski	28	34	33	53	9	25	25	26	9	53
06	08	20	Powiat zamojski	24	17	27	42	20	52	61	36	17	61
06	08	62	Powiat m. Chełm	30	32	29	30	14	29	24	20	14	32
06	08	64	Powiat m. Zamość	31	29	31	30	18	33	28	16	16	33
06	09		Podreg. lubelski	425	420	487	534	232	457	462	476	232	534
06	09	05	Powiat janowski	8	6	19	36	8	12	12	13	6	36
06	09	07	Powiat kraśnicki	31	33	36	42	17	45	26	24	17	45
06	09	08	Powiat lubartowski	22	35	47	34	14	32	36	39	14	47
06	09	09	Powiat lubelski	27	28	40	46	22	41	50	55	22	55
06	09	10	Powiat łęczyński	19	11	16	20	12	10	16	18	10	20
06	09	11	Powiat łukowski	35	31	42	33	16	34	21	19	16	42
06	09	12	Powiat opolski	14	7	13	13	8	9	18	20	7	20
06	09	14	Powiat puławski	48	55	65	79	21	36	33	44	21	79
06	09	16	Powiat rycki	24	23	20	35	9	20	24	20	9	35
06	09	17	Powiat świdnicki	29	26	26	23	16	28	28	35	16	35
06	09	63	Powiat m. Lublin	168	165	163	173	89	190	198	189	89	198
08	10		Podreg. gorzowski	139	126	128	148	95	220	212	221	95	221
08	10	01	Powiat gorzowski	21	14	13	24	13	27	25	33	13	33
08	10	03	Powiat międzyrzecki	27	22	21	23	14	36	25	33	14	36
08	10	05	Powiat słubicki	10	6	7	15	13	25	19	21	6	25
08	10	06	Powiat strzelecko-drezdenecki	15	18	22	20	12	35	32	28	12	35
08	10	07	Powiat sulęciński	7	5	7	6	5	22	26	15	5	26
08	10	61	Powiat m. Gorzów Wielkopolski	59	61	58	60	38	75	85	91	38	91
08	11		Podregion 11 - zielonogórski	225	220	248	284	154	368	370	370	154	370
08	11	02	Powiat krośniński	17	14	18	39	16	27	31	38	14	39
08	11	04	Powiat nowosolski	44	42	43	40	25	65	45	42	25	65
08	11	08	Powiat świebodziński	16	15	26	29	10	29	34	35	10	35
08	11	09	Powiat zielonogórski	27	27	26	28	14	46	48	66	14	66
08	11	10	Powiat żagański	40	39	40	41	25	49	54	50	25	54
08	11	11	Powiat żarski	23	26	36	45	21	45	51	50	21	51
08	11	12	Powiat wschowski	0	0	0	0	9	23	29	19	0	29
08	11	62	Powiat m. Zielona Góra	58	57	59	62	34	84	78	70	34	84
10	12		Podreg. łódzki	328	344	444	483	175	305	304	318	175	483
10	12	02	Powiat kutnowski	38	44	52	53	23	42	50	45	23	53
10	12	03	Powiat łaski	39	24	23	26	6	9	14	27	6	39
10	12	04	Powiat łęczycki	10	9	9	13	7	15	18	20	7	20
10	12	06	Powiat łódzki wschodni	21	23	40	33	13	22	13	20	13	40
10	12	08	Powiat pabianicki	60	50	59	62	23	45	51	47	23	62
10	12	11	Powiat poddębicki	9	18	27	14	7	14	14	11	7	27
10	12	14	Powiat sieradzki	26	36	47	47	32	56	43	37	26	56

Tablica 3.1. Liczby jps zbadanych z powiatów w latach 1995—2002 (cd.)

Woj.	Podr.	Pow.	Nazwa podregionu (powiatu)	1995	1996	1997	1998	1999	2000	2001	2002	Min.	Max.
10	12	17	Powiat wieluński	20	21	37	62	10	20	15	19	10	62
10	12	18	Powiat wieruszowski	10	13	24	36	0	6	7	8	0	36
10	12	19	Powiat zduńskowolski	26	31	34	35	13	12	17	22	12	35
10	12	20	Powiat zgierski	69	75	92	102	36	55	51	47	36	102
10	12	21	Powiat brzeziński	0	0	0	0	5	9	11	15	0	15
10	13		Podreg. piotrkowsko-skierniew.	325	338	395	437	126	303	299	280	126	437
10	13	01	Powiat bełchatowski	43	37	48	64	23	44	42	41	23	64
10	13	05	Powiat łowicki	25	30	42	48	6	24	29	22	6	48
10	13	07	Powiat opoczyński	14	27	48	55	12	22	24	28	12	55
10	13	09	Powiat pajęczański	22	22	18	14	0	17	18	24	0	24
10	13	10	Powiat piotrkowski	26	25	33	37	0	30	31	22	0	37
10	13	12	Powiat radomszczański	43	50	54	58	20	33	39	44	20	58
10	13	13	Powiat rawski	11	9	10	21	12	18	19	11	9	21
10	13	15	Powiat skierniewicki	23	23	23	19	0	11	11	11	0	23
10	13	16	Powiat tomaszowski	58	47	48	50	18	45	48	47	18	58
10	13	62	Powiat m. Piotrków Trybunalski	42	44	42	41	22	29	21	17	17	44
10	13	63	Powiat m. Skierniewice	18	24	29	30	13	30	17	13	13	30
10	14		Podreg. m Łódź	484	494	487	500	219	388	388	394	219	500
10	14	61	Powiat m. Łódź	484	494	487	500	219	388	388	394	219	500
12	15		Podreg. krakowsko-tarnowski	392	421	515	571	211	350	327	331	211	571
12	15	01	Powiat bocheński	24	35	40	34	17	21	22	21	17	40
12	15	02	Powiat brzeski	18	14	27	24	7	17	18	18	7	27
12	15	03	Powiat chrzanowski	23	46	68	64	30	42	31	35	23	68
12	15	04	Powiat dąbrowski	28	32	31	19	9	6	13	15	6	32
12	15	06	Powiat krakowski	59	55	66	106	31	63	57	49	31	106
12	15	08	Powiat miechowski	16	18	18	20	14	14	12	16	12	20
12	15	12	Powiat olkuski	53	45	51	57	20	25	22	18	18	57
12	15	13	Powiat oświęcimski	41	44	54	66	20	35	32	41	20	66
12	15	14	Powiat proszowicki	12	14	21	17	0	17	11	8	0	21
12	15	16	Powiat tarnowski	36	36	55	64	23	48	55	53	23	64
12	15	19	Powiat wielicki	24	23	26	41	17	25	20	22	17	41
12	15	63	Powiat m. Tarnów	58	59	58	59	23	37	34	35	23	59
12	16		Podreg. nowosądecki	260	269	322	359	133	261	279	281	133	359
12	16	05	Powiat gorlicki	27	30	30	36	14	27	36	31	14	36
12	16	07	Powiat limanowski	33	34	42	38	11	21	30	28	11	42
12	16	09	Powiat myślenicki	21	26	35	21	15	38	17	25	15	38
12	16	10	Powiat nowosądecki	35	34	45	46	15	31	25	34	15	46
12	16	11	Powiat nowotarski	36	39	61	75	22	33	53	50	22	75
12	16	15	Powiat suski	27	19	11	23	12	18	16	25	11	27
12	16	17	Powiat tatrzański	19	20	21	31	7	11	26	23	7	31
12	16	18	Powiat wadowicki	32	37	48	58	22	53	55	37	22	58
12	16	62	Powiat m. Nowy Sącz	30	30	29	31	15	29	21	28	15	31
12	17		Podreg. M. Kraków	356	359	367	370	191	392	388	369	191	392
12	17	61	Powiat m. Kraków	356	359	367	370	191	392	388	369	191	392
14	18		Podreg. ciechanowsko-płocki	202	205	263	308	82	135	132	123	82	308
14	18	02	Powiat ciechanowski	22	23	24	36	13	14	20	19	13	36
14	18	04	Powiat gostyński	14	18	28	51	6	11	17	13	6	51
14	18	13	Powiat mławski	29	31	43	52	13	14	18	17	13	52
14	18	19	Powiat płocki	26	26	30	30	0	26	14	20	0	30
14	18	20	Powiat płoński	32	21	40	46	9	18	16	12	9	46
14	18	27	Powiat sierpecki	10	15	28	25	10	9	5	6	5	28
14	18	37	Powiat żuromiński	13	14	11	6	5	11	8	3	3	14
14	18	62	Powiat m. Płock	56	57	59	62	26	32	34	33	26	62
14	19		Podreg. ostrołęcko-siedlecki	190	224	286	337	87	118	143	145	87	337
14	19	10	Powiat łosicki	4	5	6	20	7	9	7	4	4	20
14	19	11	Powiat makowski	8	14	11	12	4	7	12	19	4	19
14	19	15	Powiat ostrołęcki	9	12	16	21	0	2	8	9	0	21
14	19	16	Powiat ostrowski	26	25	45	58	13	11	12	16	11	58
14	19	22	Powiat przasnyski	9	16	24	25	6	9	11	10	6	25
14	19	24	Powiat pułtuski	12	6	11	18	10	6	9	11	6	18
14	19	26	Powiat siedlecki	20	26	38	41	10	15	23	18	10	41
14	19	29	Powiat sokołowski	26	23	20	27	14	18	17	12	12	27
14	19	33	Powiat węgrowski	13	21	32	22	0	15	14	19	0	32
14	19	35	Powiat wyszkowski	18	22	25	33	7	0	5	9	0	33

Tablica 3.1. Liczby jps zbadanych z powiatów w latach 1995—2002 (cd.)

Woj.	Podr.	Pow.	Nazwa podregionu (powiatu)	1995	1996	1997	1998	1999	2000	2001	2002	Min.	Max.
14	19	61	Powiat m. Ostrołęka	16	22	29	30	8	10	7	6	6	30
14	19	64	Powiat m. Siedlce	29	32	29	30	8	16	18	12	8	32
14	20		Podreg. warszawski	413	406	466	537	157	236	213	213	157	537
14	20	05	Powiat grodziski	31	23	17	25	9	7	11	13	7	31
14	20	06	Powiat grójecki	35	38	41	51	15	32	24	20	15	51
14	20	08	Powiat legionowski	30	32	39	34	10	17	14	16	10	39
14	20	12	Powiat miński	43	39	66	82	20	22	26	30	20	82
14	20	14	Powiat nowodworski	13	19	26	30	16	21	15	16	13	30
14	20	17	Powiat otwocki	35	37	36	47	7	15	9	22	7	47
14	20	18	Powiat piaseczyński	24	31	36	37	12	20	20	14	12	37
14	20	21	Powiat pruszkowski	55	47	47	49	15	22	18	19	15	55
14	20	28	Powiat sochaczewski	30	24	30	38	10	27	25	18	10	38
14	20	32	Powiat warszawski zachodni	28	24	24	32	9	15	13	12	9	32
14	20	34	Powiat wołomiński	59	54	63	72	18	20	18	17	17	72
14	20	38	Powiat żyrardowski	30	38	41	40	16	18	20	16	16	41
14	21		Podreg. radomski	247	238	266	295	101	121	117	121	101	295
14	21	01	Powiat białobrzeski	11	4	9	8	5	4	0	3	0	11
14	21	03	Powiat garwoliński	24	20	17	25	8	16	18	24	8	25
14	21	07	Powiat kozienicki	13	12	21	30	13	9	11	14	9	30
14	21	09	Powiat lipski	19	18	19	31	0	8	5	6	0	31
14	21	23	Powiat przysuski	15	19	21	21	10	11	12	10	10	21
14	21	25	Powiat radomski	36	37	51	58	17	30	20	11	11	58
14	21	30	Powiat sztywniowski	11	12	14	11	2	0	4	4	0	14
14	21	36	Powiat zwoleński	12	9	8	8	4	1	4	7	1	12
14	21	63	Powiat m. Radom	106	107	106	103	42	42	43	42	42	107
14	22		Podreg. m. Warszawa	875	884	869	883	356	548	573	575	356	884
14	22	31	Powiat warszawski	875	884	869	883	356	548	573	575	356	884
16	23		Podreg. opolski	349	363	446	484	242	596	587	589	242	596
16	23	01	Powiat brzeski	29	27	33	31	17	45	49	44	17	49
16	23	02	Powiat głubczycki	11	17	16	7	14	38	27	20	7	38
16	23	03	Powiat kędzierzyńsko-kozielski	46	38	45	43	27	81	74	68	27	81
16	23	04	Powiat kluczborski	17	21	35	30	16	37	39	40	16	40
16	23	05	Powiat krapkowicki	22	26	23	16	10	39	29	34	10	39
16	23	06	Powiat namysłowski	25	15	10	15	10	26	26	24	10	26
16	23	07	Powiat nyski	45	47	66	85	38	84	80	78	38	85
16	23	08	Powiat oleski	12	22	34	39	11	39	30	37	11	39
16	23	09	Powiat opolski	36	42	54	65	32	73	69	67	32	73
16	23	10	Powiat prudnicki	23	29	41	62	15	15	33	36	15	62
16	23	11	Powiat strzelecki	26	17	28	31	17	45	48	48	17	48
16	23	61	Powiat m. Opole	57	62	61	60	35	74	83	93	35	93
18	24		Podreg. rzeszowsko-tarnobrzeski	364	378	433	418	194	448	451	433	194	451
18	24	03	Powiat dębicki	24	29	35	47	20	55	68	51	20	68
18	24	06	Powiat kolbuszowski	24	12	17	20	5	21	19	16	5	24
18	24	08	Powiat leżajski	23	14	8	21	9	27	19	14	8	27
18	24	10	Powiat łańcucki	19	30	45	27	12	20	31	33	12	45
18	24	11	Powiat mielecki	41	47	56	49	20	55	70	69	20	70
18	24	12	Powiat nizański	24	31	25	14	21	39	15	13	13	39
18	24	15	Powiat ropczycko-sędziszowski	18	27	42	35	10	27	21	19	10	42
18	24	16	Powiat rzeszowski	45	40	37	47	22	55	64	48	22	64
18	24	18	Powiat stalowowolski	45	37	37	36	20	35	31	51	20	51
18	24	20	Powiat tarnobrzeski	12	21	33	21	6	17	22	32	6	33
18	24	63	Powiat m. Rzeszów	72	69	72	72	38	74	71	65	38	74
18	24	64	Powiat m. Tarnobrzeg	17	21	26	29	11	23	20	22	11	29
18	25		Podreg. krośniensko-przemyski	266	264	330	399	161	351	352	366	161	399
18	25	01	Powiat bieszczadzki	15	16	34	26	4	11	9	15	4	34
18	25	02	Powiat brzozowski	12	4	6	11	9	25	31	27	4	31
18	25	04	Powiat jarosławski	38	35	40	46	25	37	39	35	25	46
18	25	05	Powiat jasielski	38	38	32	50	21	33	36	55	21	55
18	25	07	Powiat krośniński	21	33	36	50	20	50	38	42	20	50
18	25	09	Powiat lubaczowski	20	15	13	31	4	11	20	12	4	31
18	25	13	Powiat przemyski	9	18	28	31	0	21	23	27	0	31
18	25	14	Powiat przeworski	22	22	33	26	13	32	29	33	13	33
18	25	17	Powiat sanocki	33	28	32	32	17	40	37	32	17	40
18	25	19	Powiat strzyżowski	13	10	23	38	12	16	15	17	10	38

Tablica 3.1. Liczby jps zbadanych z powiatów w latach 1995—2002 (cd.)

Woj.	Podr.	Pow.	Nazwa podregionu (powiatu)	1995	1996	1997	1998	1999	2000	2001	2002	Min.	Max.
18	25	21	Powiat leski	0	0	0	0	0	10	15	10	0	15
18	25	61	Powiat m. Krosno	15	17	25	28	16	22	19	31	15	31
18	25	62	Powiat m. Przemysł	30	28	28	30	20	43	41	30	20	43
20	26		Podreg. białostocko-suwalski	341	335	384	429	203	467	464	473	203	473
20	26	01	Powiat augustowski	27	26	23	27	11	7	9	19	7	27
20	26	02	Powiat białostocki	43	41	50	45	25	71	71	76	25	76
20	26	03	Powiat bielski	22	26	30	27	14	28	29	32	14	32
20	26	05	Powiat hajnowski	20	21	32	35	16	24	18	33	16	35
20	26	08	Powiat moniecki	12	12	21	22	10	22	18	22	10	22
20	26	09	Powiat sejneński	8	2	2	6	0	19	17	4	0	19
20	26	10	Powiat siemiatycki	16	22	20	23	12	30	27	32	12	32
20	26	11	Powiat sokólski	19	18	36	61	16	39	44	30	16	61
20	26	12	Powiat suwalski	11	10	10	19	0	15	9	8	0	19
20	26	61	Powiat m. Białystok	136	132	129	136	85	184	190	192	85	192
20	26	63	Powiat m. Suwałki	27	25	31	28	14	28	32	25	14	32
20	27		Podreg. łomżyński	99	98	112	137	48	123	134	123	48	137
20	27	04	Powiat grajewski	14	10	8	21	11	15	23	31	8	31
20	27	06	Powiat kolneński	7	2	7	13	6	19	21	13	2	21
20	27	07	Powiat łomżyński	20	21	22	17	0	18	18	16	0	22
20	27	13	Powiat wysokomazowiecki	19	21	25	32	9	27	34	28	9	34
20	27	14	Powiat zambrowski	9	15	20	23	6	10	12	13	6	23
20	27	62	Powiat m. Łomża	30	29	30	31	16	34	26	22	16	34
22	28		Podreg. słupski	141	148	184	198	76	155	147	142	76	198
22	28	01	Powiat bytowski	11	12	21	24	10	30	28	24	10	30
22	28	02	Powiat chojnicki	22	22	34	34	13	25	19	23	13	34
22	28	03	Powiat czuchowski	16	18	31	27	11	25	29	22	11	31
22	28	08	Powiat łęborski	24	28	21	20	14	22	28	24	14	28
22	28	12	Powiat słupski	22	21	31	46	9	25	15	21	9	46
22	28	63	Powiat m. Słupsk	46	47	46	47	19	28	28	28	19	47
22	29		Podreg. gdański	244	255	302	320	133	248	260	265	133	320
22	29	04	Powiat gdański	18	16	16	35	14	24	19	17	14	35
22	29	05	Powiat kartuski	13	16	29	25	11	19	12	9	9	29
22	29	06	Powiat kościerski	13	14	24	23	6	14	19	17	6	24
22	29	07	Powiat kwidzyński	27	27	30	26	15	27	28	26	15	30
22	29	09	Powiat malborski	30	38	45	37	10	17	27	39	10	45
22	29	10	Powiat nowodworski	6	6	8	11	5	8	9	8	5	11
22	29	11	Powiat pucki	13	8	18	32	6	14	13	17	6	32
22	29	13	Powiat starogardzki	25	34	42	36	20	39	38	35	20	42
22	29	14	Powiat tczewski	53	51	43	42	18	22	31	27	18	53
22	29	15	Powiat wejherowski	46	45	47	53	20	50	51	54	20	54
22	29	16	Powiat sztumski	0	0	0	0	8	14	13	16	0	16
22	30		Podreg. Gdańsk-Gdynia-Sopot	363	376	378	383	191	377	388	386	191	388
22	30	61	Powiat m. Gdańsk	202	221	225	242	120	232	238	237	120	242
22	30	62	Powiat m. Gdynia	145	131	122	112	59	132	133	128	59	145
22	30	64	Powiat m. Sopot	16	24	31	29	12	13	17	21	12	31
24	31		Podreg. częstochowski	224	219	233	263	96	135	138	143	96	263
24	31	04	Powiat częstochowski	32	39	48	57	23	35	36	24	23	57
24	31	06	Powiat kłobucki	26	23	30	47	14	21	16	19	14	47
24	31	09	Powiat myszkowski	32	26	23	22	8	23	16	20	8	32
24	31	64	Powiat m. Częstochowa	134	131	132	137	51	56	70	80	51	137
24	32		Podreg. bielsko-bialski	214	216	258	262	89	132	156	164	89	262
24	32	02	Powiat bielski	31	31	56	54	17	32	29	33	17	56
24	32	03	Powiat cieszyński	63	57	60	67	23	40	50	44	23	67
24	32	17	Powiat żywiecki	32	41	58	54	18	22	26	27	18	58
24	32	61	Powiat m. Bielsko-Biała	88	87	84	87	31	38	51	60	31	88
24	33		Podreg. centralny śląski	1447	1470	1520	1556	598	939	941	918	598	1556
24	33	01	Powiat będziński	61	79	87	85	29	42	55	61	29	87
24	33	05	Powiat gliwicki	43	39	51	60	13	18	28	20	13	60
24	33	07	Powiat lubliniecki	21	19	34	43	7	15	23	22	7	43
24	33	08	Powiat mikołowski	42	43	47	49	16	24	15	25	15	49
24	33	10	Powiat pszczyński	19	29	29	24	12	20	13	13	12	29
24	33	13	Powiat tarnogórski	42	35	42	58	26	44	40	34	26	58
24	33	14	Powiat bieruńsko-lędziński	12	11	17	26	12	11	5	10	5	26
24	33	16	Powiat zawierciański	61	61	67	59	23	32	32	42	23	67

Tablica 3.1. Liczby jps zbadanych z powiatów w latach 1995—2002 (cd.)

Woj.	Podr.	Pow.	Nazwa podregionu (powiatu)	1995	1996	1997	1998	1999	2000	2001	2002	Min.	Max.
24	33	62	Powiat m. Bytom	103	106	117	120	37	48	57	61	37	120
24	33	63	Powiat m. Chorzów	72	68	74	76	34	62	55	38	34	76
24	33	65	Powiat m. Dąbrowa Górnicza	72	67	76	78	28	49	60	67	28	78
24	33	66	Powiat m. Gliwice	98	109	104	105	38	43	37	51	37	109
24	33	68	Powiat m. Jaworzno	42	45	43	46	17	26	30	30	17	46
24	33	69	Powiat m. Katowice	195	188	190	193	92	157	174	191	92	195
24	33	70	Powiat m. Mysłowice	46	49	45	45	17	34	30	21	17	49
24	33	71	Powiat m. Piekary Śląskie	34	29	30	32	14	22	16	13	13	34
24	33	72	Powiat m. Ruda Śląska	70	84	86	88	36	57	61	50	36	88
24	33	74	Powiat m. Siemianowice Śl.	38	45	45	43	16	22	22	17	16	45
24	33	75	Powiat m. Sosnowiec	148	152	139	131	51	91	76	62	51	152
24	33	76	Powiat m. Świętochłowice	38	28	28	30	14	19	21	27	14	38
24	33	77	Powiat m. Tychy	75	75	63	57	23	44	37	24	23	75
24	33	78	Powiat m. Zabrze	115	109	106	108	43	59	54	39	39	115
24	45		Podreg. rybnicko-jastrzębski	245	256	274	295	107	150	156	178	107	295
24	45	11	Powiat raciborski	28	36	45	48	21	34	34	31	21	48
24	45	12	Powiat rybnicki	23	25	33	31	13	23	19	20	13	33
24	45	15	Powiat wodzisławski	53	62	64	79	23	38	37	37	23	79
24	45	67	Powiat m. Jastrzębie-Zdrój	46	51	42	47	18	18	21	41	18	51
24	45	73	Powiat m. Rybnik	69	52	59	61	23	24	32	42	23	69
24	45	79	Powiat m. Żory	26	30	31	29	9	13	13	7	7	31
26	34		Podreg. Świętokrzyski	435	447	556	615	266	586	594	592	266	615
26	34	01	Powiat buski	14	16	32	38	15	25	25	34	14	38
26	34	02	Powiat jędrzejowski	28	29	35	38	16	38	34	28	16	38
26	34	03	Powiat kazimierski	13	15	20	6	7	14	18	10	6	20
26	34	04	Powiat kielecki	37	41	67	80	19	59	57	52	19	80
26	34	05	Powiat konecki	27	28	36	47	22	32	19	24	19	47
26	34	06	Powiat opatowski	27	28	32	32	17	27	34	30	17	34
26	34	07	Powiat ostrowiecki	48	54	61	49	21	47	54	49	21	61
26	34	08	Powiat pińczowski	13	14	24	26	5	12	11	9	5	26
26	34	09	Powiat sandomierski	22	16	25	42	18	45	38	32	16	45
26	34	10	Powiat skarżyski	26	24	33	40	18	26	36	43	18	43
26	34	11	Powiat starachowicki	49	45	46	53	19	34	37	36	19	53
26	34	12	Powiat staszowski	15	18	25	34	16	38	27	35	15	38
26	34	13	Powiat włoszczowski	15	18	20	24	7	14	19	23	7	24
26	34	61	Powiat m. Kielce	101	101	100	106	66	175	185	187	66	187
28	35		Podreg. elbląski	179	193	209	234	109	204	227	234	109	234
28	35	02	Powiat braniewski	8	15	17	16	14	22	16	16	8	22
28	35	03	Powiat działowski	21	27	29	31	8	18	20	17	8	31
28	35	04	Powiat elbląski	27	19	17	28	8	22	23	27	8	28
28	35	07	Powiat iławski	28	24	30	34	18	19	34	40	18	40
28	35	12	Powiat nowomiejski	7	12	16	10	11	22	19	26	7	26
28	35	15	Powiat ostródzki	27	35	41	52	24	46	49	52	24	52
28	35	61	Powiat m. Elbląg	61	61	59	63	26	55	66	56	26	66
28	36		Podreg. olsztyński	184	197	248	273	120	288	271	283	120	288
28	36	01	Powiat bartoszycki	21	22	33	43	15	30	25	38	15	43
28	36	08	Powiat kętrzyński	22	17	23	37	19	37	27	30	17	37
28	36	09	Powiat lidzbarski	6	13	21	17	3	18	24	15	3	24
28	36	10	Powiat mragowski	13	11	4	13	13	26	26	26	4	26
28	36	11	Powiat nidzicki	19	15	12	11	0	23	23	18	0	23
28	36	14	Powiat olsztyński	20	31	52	44	15	34	42	40	15	52
28	36	17	Powiat szczycieński	13	15	27	35	14	38	24	23	13	38
28	36	62	Powiat m. Olsztyn	70	73	76	73	41	82	80	93	41	93
28	37		Podreg. elcki	77	99	125	126	50	118	99	83	50	126
28	37	05	Powiat elcki	25	31	33	40	18	39	29	22	18	40
28	37	06	Powiat giżycki	27	31	44	44	14	24	15	17	14	44
28	37	13	Powiat olecki	17	17	20	25	5	12	11	5	5	25
28	37	16	Powiat piski	8	20	28	17	6	16	17	16	6	28
28	37	18	Powiat gołdapski	0	0	0	0	3	16	12	11	0	16
28	37	19	Powiat węgorzewski	0	0	0	0	4	11	15	12	0	15
30	38		Podreg. pilski	123	124	154	167	57	82	117	124	57	167
30	38	01	Powiat chodzieski	8	6	11	18	8	9	12	18	6	18
30	38	02	Powiat czarnkowsko-trzcianecki	39	24	38	47	16	15	25	33	15	47
30	38	19	Powiat pilski	40	46	53	50	18	39	44	41	18	53

Tablica 3.1. Liczby jps zbadanych z powiatów w latach 1995—2002 (cd.)

Woj.	Podr.	Pow.	Nazwa podregionu (powiatu)	1995	1996	1997	1998	1999	2000	2001	2002	Min.	Max.
30	38	28	Powiat wągrowiecki	18	25	32	36	10	14	20	17	10	36
30	38	31	Powiat złotowski	18	23	20	16	5	5	16	15	5	23
30	39		Podreg. poznański	282	321	393	396	146	251	300	343	146	396
30	39	03	Powiat gnieźnieński	39	43	45	41	22	33	25	50	22	50
30	39	05	Powiat grodziski	10	14	32	16	5	12	14	16	5	32
30	39	11	Powiat kościański	14	17	23	30	9	16	22	27	9	30
30	39	13	Powiat leszczyński	15	16	19	10	0	7	11	11	0	19
30	39	14	Powiat międzychodzki	7	9	13	14	5	4	13	16	4	16
30	39	15	Powiat nowotomyski	26	31	33	19	8	11	15	18	8	33
30	39	16	Powiat obornicki	14	17	14	15	8	7	16	15	7	17
30	39	21	Powiat poznański	51	59	71	82	28	63	68	86	28	86
30	39	24	Powiat szamotulski	13	14	24	32	10	23	30	27	10	32
30	39	25	Powiat średzki	13	17	24	28	10	12	17	14	10	28
30	39	26	Powiat śremski	23	25	21	21	10	23	25	19	10	25
30	39	29	Powiat wolsztyński	7	15	19	17	3	5	14	17	3	19
30	39	30	Powiat wrzesiński	22	15	25	43	15	17	18	16	15	43
30	39	63	Powiat m. Leszno	28	29	30	28	13	18	12	11	11	30
30	40		Podreg. kaliski	240	244	296	323	105	206	200	225	105	323
30	40	04	Powiat gostyński	25	18	30	31	7	17	16	13	7	31
30	40	06	Powiat jarociński	24	22	37	33	11	10	5	12	5	37
30	40	07	Powiat kaliski	14	22	30	29	0	22	25	29	0	30
30	40	08	Powiat kępiński	18	9	16	11	0	13	15	13	0	18
30	40	12	Powiat krotoszyński	9	15	23	25	16	28	26	26	9	28
30	40	17	Powiat ostrowski	58	50	49	55	18	46	50	59	18	59
30	40	18	Powiat ostrzeszowski	11	17	19	21	11	18	20	14	11	21
30	40	20	Powiat pleszewski	25	23	16	24	8	14	8	6	6	25
30	40	22	Powiat rawicki	15	22	22	36	12	11	5	15	5	36
30	40	61	Powiat m. Kalisz	41	46	54	58	22	27	30	38	22	58
30	41		Podreg. koniński	125	138	180	191	52	112	129	147	52	191
30	41	09	Powiat kolski	21	20	33	46	19	27	38	36	19	46
30	41	10	Powiat koniński	27	29	48	43	0	29	24	32	0	48
30	41	23	Powiat słupecki	19	27	29	19	5	11	13	19	5	29
30	41	27	Powiat turecki	27	26	26	41	14	20	25	26	14	41
30	41	62	Powiat m. Konin	31	36	44	42	14	25	29	34	14	44
30	42		Podreg. m. Poznań	277	284	288	285	158	334	336	337	158	337
30	42	64	Powiat m. Poznań	277	284	288	285	158	334	336	337	158	337
32	43		Podreg. szczeciński	411	433	481	500	251	575	564	556	251	575
32	43	02	Powiat choszczeński	15	21	14	13	8	18	14	15	8	21
32	43	04	Powiat goleniowski	34	30	36	28	7	13	20	23	7	36
32	43	05	Powiat gryficki	25	26	29	41	14	25	22	20	14	41
32	43	06	Powiat gryfiński	30	23	23	39	15	34	22	16	15	39
32	43	07	Powiat kamieński	14	10	13	9	7	16	11	14	7	16
32	43	10	Powiat myśliborski	15	24	38	34	14	32	21	18	14	38
32	43	11	Powiat policki	7	14	24	30	7	30	42	40	7	42
32	43	12	Powiat pyrzycki	12	13	26	19	2	5	15	12	2	26
32	43	14	Powiat stargardzki	50	50	53	57	24	45	29	26	24	57
32	43	18	Powiat łobeski	0	0	0	0	9	12	12	22	0	22
32	43	62	Powiat m. Szczecin	193	204	211	210	135	331	338	331	135	338
32	43	63	Powiat m. Świnoujście	16	18	14	20	9	14	18	19	9	20
32	44		Podreg. koszaliński	209	209	235	286	97	189	212	215	97	286
32	44	01	Powiat białogardzki	19	24	17	22	9	10	15	17	9	24
32	44	03	Powiat drawski	20	18	18	19	8	25	25	17	8	25
32	44	08	Powiat kołobrzeski	23	23	33	39	17	25	24	24	17	39
32	44	09	Powiat koszaliński	15	12	18	32	0	22	28	21	0	32
32	44	13	Powiat sławieński	17	17	12	16	11	16	14	18	11	18
32	44	15	Powiat szczecinecki	25	29	30	18	9	20	23	23	9	30
32	44	16	Powiat świdwiński	15	13	21	36	10	16	20	19	10	36
32	44	17	Powiat walecki	20	15	24	43	10	15	14	23	10	43
32	44	61	Powiat m. Koszalin	55	58	62	61	23	40	49	53	23	62

Z powyższej tablicy wynika, że w wielu powiatach i dla wielu z rozpatrywanych lat liczba wylosowanych jps jest równa zero lub nie przekracza dziesięciu elementów (porównaj przedostatnią kolumnę powyższej tablicy; kolumna ta jest zatytułowana „min.”). Trudno w takich sytuacjach zastosować estymator klasyczny t_w , o którym pisaliśmy na początku punktu 3.1. Estymatora takiego nie sposób zastosować nawet wtedy, gdy liczba jps jest mniejsza od pięćdziesięciu, gdyż względny błąd szacunku byłby zbyt wysoki. Dlatego też postanowiliśmy podzielić powiaty na dwie grupy: małe (do 50 jps) i pozostałe, które umownie będziemy nazywać większymi. Zdajemy sobie sprawę z tego, iż podział ten jest dość arbitralny. Uważamy jednak, że taki podział jest niezbędny i do każdej z wyróżnionych grup powiatów należy zastosować inne metody estymacji parametrów.

Dla małych powiatów może być zastosowany wyłącznie estymator syntetyczny określony wzorem

$$(3.1.4) \quad x_{wsp} = t_w f_{wsp},$$

gdzie f_{wsp} oznacza udział wartości danej zmiennej (w NSP 2002) w p -tym powiecie należącem do s -tego podregionu w stosunku do wartości tej zmiennej dla w -tego województwa. Estymatory (3.1.4) będą charakteryzowały się największymi obciążeniami i największymi błędami średniokwadratowymi. Tego jak wielkie jest obciążenie tych estymatorów nie można określić bez przeprowadzenia bardzo żmudnych badań symulacyjnych. Dla tych estymatorów ich współczynnik zmienności (względny błąd szacunku) będzie mniejszy niż innych rozpatrywanych tu estymatorów i równy współczynnikowi zmienności estymatora z tego poziomu agregacji, z którego były brane dane do „rozszacowywania” ocen dla powiatów zgodnie ze strukturą otrzymaną z NSP 2002r.

W przypadku powiatów większych stosowane będą trzy opisane wyżej estymatory tzn. klasyczny (K), syntetyczny (S) i złożony (Z). Jest rzeczą dyskusyjną jak zmieniać wagi przy konstrukcji estymatora złożonego w miarę wzrastania wielkości powiatu mierzonej liczbą jps. Intuicja podpowiada, że należałoby te wagi zmniejszać w miarę wzrostu wielkości powiatu. Uważamy jednak, że jest to problem na oddzielne badanie symulacyjne. Zapewne od pewnego progu wielkości powiatu można by korzystać z rozwiązań zaproponowanych przez cytowanego już wielokrotnie R. Griffithsa. (1996) (por. uwagi zamieszczone po wzorze (3.1.3)).

W dalszej części opracowania często będziemy odwoływać się do liczb jps zawartych w powiatach, podregionach i województwach. Ponadto będą nas interesowały liczby jps wylosowanych z poszczególnych jednostek administracyjnych.

Tablica 3.2. Numery i nazwy podregionów w województwach oraz liczby jps w roku 2002

Nr woj.	Nr podregionu	Nazwa podregionu	Liczba jps		% zbadan. jps
			wylosow.	zbadan.	
(1)	(2)	(3)	(4)	(5)	(6)
2	1	jeleniogórsko-wałbrzyski	391	378	96,68
2	2	legnicki	147	142	96,60
2	3	wrocławski	86	86	100,00
2	4	Wrocław (m)	416	370	88,94
4	5	bydgoski	411	398	96,84
4	6	toruńsko-włocławski	421	397	94,30
6	7	białskopodlaski	116	110	94,83
6	8	chełmsko-zamojski	213	211	99,06
6	9	lubelski	503	476	94,63
8	10	gorzowski	238	221	92,86
8	11	zielonogórski	386	370	95,85
10	12	łódzki	339	318	93,81
10	13	piotrkowsko-skierniewicki	285	280	98,25
10	14	Łódź (m)	416	394	94,71
12	15	krakowsko-tarnowski	339	331	97,64
12	16	nowosądecki	285	281	98,60
12	17	Kraków (m)	416	369	88,70
14	18	ciechanowsko-płocki	125	123	98,40
14	19	ostrołęcko-siedlecki	147	145	98,64
14	20	warszawski	228	213	93,42
14	21	radomski	123	121	98,37
14	22	Warszawa (m)	625	575	92,00
16	23	opolski	624	589	94,39
18	24	rzyszowsko-tarnobrzeski	454	433	95,37
18	25	krośnieńsko-przemyski	378	366	96,83
20	26	białostocko-suwalski	500	473	94,60
20	27	łomżyński	124	123	99,19
22	28	śląpski	144	142	98,61
22	29	gdański	272	265	97,43
22	30	Gdańsk-Gdynia-Sopot	416	386	92,79
24	31	częstochowski	147	143	97,28
24	32	bielsko-bialski	175	164	93,71
24	33	centralny śląski	949	918	96,73
24	45	rybnicko-jastrzębski	185	178	96,22
26	34	świętokrzyski	624	592	94,87
28	35	elbląski	240	234	97,50
28	36	olsztyński	294	283	96,26
28	37	ełcki	90	83	92,22
30	38	pilski	130	124	95,38
30	39	poznański	354	343	96,89
30	40	kaliski	231	225	97,40
30	41	koniński	151	147	97,35
30	42	Poznań (m)	382	337	88,22
32	43	szczeeciński	613	556	90,70
32	44	koszaliński	219	215	98,17
Σ	Suma		14352	13628	94,96

Tablica 3.3. Liczby podregionów i powiatów w województwach oraz liczby jps w roku 2002

Nr woj.	Nazwa województwa	Liczba podregionów	Liczba powiatów	Liczba jps		% zbadan. jps
				wylosow.	zbadan.	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
2	dolnośląskie	4	29	1040	976	93,85
4	kujawsko-pomorskie	2	23	832	795	95,55
6	lubelskie	3	24	832	797	95,79
8	lubuskie	2	14	624	591	94,71
10	Łódzkie	3	24	1040	992	95,38
12	małopolskie	3	22	1040	981	94,33
14	mazowieckie	5	42	1248	1177	94,31
16	opolskie	1	12	624	589	94,39
18	podkarpackie	2	25	832	799	96,03
20	podlaskie	2	17	624	596	95,51
22	pomorskie	3	20	832	793	95,31
24	Śląskie	4	36	1456	1403	96,36
26	świętokrzyskie	1	14	624	592	94,87
28	warmińsko-mazurskie	3	21	624	600	96,15
30	wielkopolskie	5	35	1248	1176	94,23
32	zachodnio-pomorskie	2	21	832	771	92,67
Σ	Suma	45	379	14352	13628	94,96

Z tablic 3.2 i 3.3 wynika, że w wielu podregionach i województwach odsetek niezbadanych, a wylosowanych jps jest znaczny. W skali kraju osiąga on nieco ponad 5%. Problem ten nie byłby groźny, gdyby występował tylko na pierwszym stopniu losowania. Niestety okazało się, że w wylosowanych i zbadanych jps wiele mieszkań nie zostało zbadanych z różnych przyczyn. Nie posiadamy dokładnych danych na ten temat. Wiadomo jednak, że kompletność badania w IV kwartale 2002 nie przekraczała 80%.

Korzystanie z danych NSP 2002 też nie gwarantuje kompletności badania, mimo iż spis powszechny był badaniem pełnym. Pewne światło na ten problem rzuca następująca tablica.

Tablica 3.4.

Lp.	Województwo	Podregion	Ogółem	Pracujący	Bezrobotni	Bierni zawodowo	O nieustalonym statusie zatrudnienia	%niezb.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1	Polska		31288428	13218344	3558154	13456155	1055775	3,37
2	02		2423065	949816	326414	1079582	67253	2,78
3	02	01	1108765	395044	168668	516036	29017	2,62
4	02	02	409088	166835	56912	172954	12387	3,03
5	02	03	348623	147507	46941	147405	6770	1,94
6	02	04	556589	240430	53893	243187	19079	3,43
7	04		1680344	699103	221344	730537	29360	1,75
8	04	05	826910	346270	104509	361457	14674	1,77

Lp.	Województwo	Podregion	Ogółem	Pracujący	Bezrobotni	Bierni zawodowo	O nieustalonym statusie zatrudnienia	%niezb.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
9	04	06	853434	352833	116835	369080	14686	1,72
10	06		1779912	788606	179055	776845	35406	1,99
11	06	07	249084	109819	27202	109252	2811	1,13
12	06	08	537616	236126	55606	234815	11069	2,06
13	06	09	993212	442661	96247	432778	21526	2,17
14	08		820822	319438	118760	362517	20107	2,45
15	08	10	309501	123398	44506	133989	7608	2,46
16	08	11	511321	196040	74254	228528	12499	2,44
17	10		2183206	961290	243672	937511	40733	1,87
18	10	12	775862	349253	85340	333579	4690	0,60
19	10	13	715851	333865	79120	295811	7055	0,99
20	10	14	691493	278172	79212	308121	25988	3,76
21	12		2607547	1097673	259540	1152079	98255	3,77
22	12	15	1108939	479391	109136	487209	33203	2,99
23	12	16	845442	345468	92711	376795	30468	3,60
24	12	17	653166	272814	57693	288075	34584	5,29
25	14		4244785	1970284	416025	1707924	150552	3,55
26	14	18	506159	224373	64834	211560	5392	1,07
27	14	19	597872	280254	63402	247456	6760	1,13
28	14	20	1055974	519540	101018	417742	17674	1,67
29	14	21	589229	257761	78872	245011	7585	1,29
30	14	22	1495551	688356	107899	586155	113141	7,57
31	16		878172	337055	92646	356479	91992	10,48
32	16	23	878172	337055	92646	356479	91992	10,48
33	18		1668108	678238	184312	726321	79237	4,75
34	18	24	916729	381248	95919	387202	52360	5,71
35	18	25	751379	296990	88393	339119	26877	3,58
36	20		975010	418474	98983	402165	55388	5,68
37	20	26	729465	312434	73863	304809	38359	5,26
38	20	27	245545	106040	25120	97356	17029	6,94
39	22		1763261	743162	213526	747557	59016	3,35
40	22	28	378597	148457	60164	161384	8592	2,27
41	22	29	737193	311174	94809	313243	17967	2,44
42	22	30	647471	283531	58553	272930	32457	5,01
43	24		3955063	1578148	413606	1771372	191937	4,85
44	24	31	452923	191357	53988	200553	7025	1,55
45	24	32	524449	226846	50869	234193	12541	2,39
46	24	33	2448289	956213	259455	1091139	141482	5,78
47	24	45	529402	203732	49294	245487	30889	5,83
48	26		1061920	454879	130132	453916	22993	2,17
49	26	34	1061920	454879	130132	453916	22993	2,17
50	28		1142388	437231	171539	496269	37349	3,27
51	28	35	423802	163429	66687	185289	8397	1,98
52	28	36	494991	193708	67908	214766	18609	3,76
53	28	37	223595	80094	36944	96214	10343	4,63
54	30		2714024	1240794	284065	1152974	36191	1,33
55	30	38	321652	136269	41529	140919	2935	0,91
56	30	39	904570	426867	90667	380717	6319	0,70
57	30	40	641302	294364	70231	271703	5004	0,78
58	30	41	346952	153734	42106	148415	2697	0,78

Lp.	Województwo	Podregion	Ogółem	Pracujący	Bezrobotni	Bierni zawodowo	O nieustalonym statusie zatrudnienia	%niezb.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
59	30	42	499548	229560	39532	211220	19236	3,85
60	32		1390801	544153	204535	602107	40006	2,88
61	32	43	910659	362662	125220	394127	28650	3,15
62	32	44	480142	181491	79315	207980	11356	2,37

Źródło: obliczenia własne na podstawie danych NSP 2002 r.

Analizując liczby z kolumny 9 widzimy, że w NSP 2002 największy odsetek niezbadanych na okoliczność aktywności ekonomicznej ludności wystąpił w województwie opolskim i przekroczył 10%. Zastanawiające jest również to, że w Warszawie (podregion dwudziesty drugi) liczba niezbadanych (113.141) przekroczyła liczbę bezrobotnych (107.899). W wielu województwach relacja między niezbadanymi a bezrobotnymi przedstawia się również bardzo niekorzystnie, aczkolwiek nie aż tak tragicznie jak w Warszawie. Częste są przypadki, gdy liczba niezbadanych stanowi 1/3 liczby bezrobotnych. To oczywiście może budzić wątpliwości, co do wiarygodności danych z NSP 2002.

Biorąc pod uwagę znaczne odsetki osób niezbadanych w BAELu i NSP, trudno jest kategorycznie stwierdzić, które z tych badań należy uznać za wiarygodne. Problem ten wystąpi z całą ostrością wtedy, gdy będziemy próbować zbadać możliwości wykorzystania danych z NSP do bieżących szacunków w BAELu. Dodatkowym problemem przy porównywaniu danych z obu tych badań jest to, że dane z NSP dotyczą określonego dnia (pierwszego dnia rozpoczęcia NSP 2002), podczas gdy dane z BAELu dotyczą całego roku, przy czym w każdym kwartale badana była różna populacja (na początku każdego kwartału włączane były do próby osoby, które osiągnęły wiek 15 lat). Zakładając nawet, że dane z BAELu i NSP nie są obciążone żadnymi błędami (losowymi bądź nielosowymi), to i tak średnia z czterech kwartałów danych BAELowskich (średnia liczona z danych dla różnych populacji) nie musiałaby się zgadzać z danymi z NSP otrzymanymi dla określonego dnia. Przy okazji należy również wspomnieć o sezonowości bezrobocia, co w porównaniach wieloletnich może mieć dość istotne znaczenie (termin rozpoczęcia prac budowlanych, zbiór płodów rolnych itp.)

3.2. Metoda szacowania precyzji estymatorów

W ostatnim akapicie punktu 2.4 wspomnieliśmy o tym, jak były szacowane wariancje estymatorów do tej pory. Stosowane były dwie metody. Pierwszą z nich była metoda zależnych podprób, przy czym podpróby dotyczyły tylko jps. Następna metoda polegała na line-

aryzacji estymatora w szereg Taylora i ograniczenie się do dwóch pierwszych składników tego rozwinięcia. Metody te opisane są dość szczegółowo, między innymi, w pracy K.M. Woltera (1985). Wspomniane metody są mało efektywne, co wykazały liczne badania symulacyjne. Ponadto w przypadku metody linearyzacji dochodzi bardzo poważny problem liczenia pochodnych estymatora względem wszystkich elementów losowych (w omawianym przypadku jest ich dużo ze względu na losowość wag finalnych). Dlatego też postanowiliśmy zmienić sposób szacowania precyzji. Alternatywą mogłaby być metoda zrównoważonych, wielokrotnych powtórzeń (ang. skrót BBR), której twórcą był P.J. McCarthy (1969). Jej opis można znaleźć również w cytowanej wyżej monografii K.M. Woltera. Metoda ta wymaga tworzenia macierzy Hadamarda dość wysokiego stopnia, co jest bardzo pracochłonne, a ponadto nie jest znacząco efektywniejsza od metod przedstawionych poprzednio (por. M.R. Frankel (1971)). Modyfikacji tej metody dokonali J.N.K. Rao i J. Shao (1996 i 1999), ale naszym zdaniem nie daje ona zadowalających efektów.

Dlatego też do wyznaczenia ocen precyzji badanych estymatorów zastosowaliśmy jedną ze znanych metod przybliżonych, opartą na podpróbkach i zasadzie „bootstrap”. Teoretyczne podstawy metody bootstrap dla prób prostych zostały zapoczątkowane pracą B.Efrona (1979). Wraz ze wzrostem możliwości obliczeniowych kolejnych generacji komputerów metody oparte na podpróbkach zyskują coraz szersze uznanie ze względu na możliwość jednolitego traktowania bardzo różnorodnych parametrów i wyeliminowanie konieczności wyrowadzania dla różnych sytuacji skomplikowanych analitycznych wzorów. Zastosowanie zasady bootstrap do najczęściej występujących w praktyce metody reprezentacyjnej nieprostych, złożonych schematów losowania próby wymaga odpowiednich modyfikacji. Szczegółowy przegląd tej tematyki można znaleźć w monografii J. Shao i D. Tu (1995).

W przypadku złożonego, dwustopniowego schematu losowania próby, jaki występuje w badaniu BAEL szacowanie wariancji odbywa się dla danych uprzednio zagregowanych na poziom danej podpopulacji, zawierającej jednostki losowania pierwszego stopnia podzielone na pewną liczbę warstw. Zastosowany wariant metody bootstrap realizowany jest oddzielnie w każdej warstwie do uzyskania oszacowania odpowiedniego składnika wariancji badanego estymatora. Niezależnie w każdej z warstw zawierającej n_h jednostek, zgodnie z metodą zaproponowaną w pracy P.J. McCarthy i C.B. Snowden (1985) (patrz też praca: J.N.K. Rao i C.F.J. Wu (1988)), losowana jest ze zbioru wszystkich jednostek pierwszego stopnia danej warstwy podpróbka o licznosci $n_h - 1$ metodą losowania prostego ze zwracaniem. Losowanie

takie powtarzane jest B razy; za każdym razem wyznaczamy dla uzyskanej b -tej podpróbki ($b=1,2,\dots,B$) zmodyfikowane wagi

$$(3.2.1) \quad w_i(b) = w_i \frac{n_h}{n_h - 1} m_i(b) ,$$

gdzie w_i jest wagą oryginalną jednostki i , $w_i(b)$ — wagą dla podpróbki, $m_i(b)$ — podaje ile razy jednostka i -ta została wybrana do podpróbki b . Dla danej podpróbki wyznaczamy z użyciem zmodyfikowanych wag oszacowanie interesującego nas parametru, uzyskując wielkość \hat{t}_b^* . Po wykonaniu B iteracji oszacowaniem wariancji badanego parametru \hat{t} (w rozważanej warstwie h) jest zgodnie z zasadą bootstrap wyrażenie

$$(3.2.2) \quad \hat{V}(\hat{t}) = \frac{1}{B-1} \sum_{b=1}^B (\hat{t}_b^* - \hat{t})^2 .$$

Aby uzyskać oszacowanie wariancji dla rozważanej podpopulacji należy dla estymatorów wysumować oszacowania wariancji uzyskane niezależnie we wszystkich warstwach analizowanej podpopulacji.

W przypadku szacowania wariancji ilorazów (współczynników aktywności zawodowej, wskaźników zatrudnienia i stóp bezrobocia) należy wykorzystując wagi określone wzorem (3.2.1) oszacować oddzielnie licznik i mianownik dla poszczególnych podprób, obliczyć ich iloraz, a następnie skorzystać ze wzoru (3.2.2).

Jako wskaźnik precyzji danego estymatora w tablicach podane są oceny współczynników zmienności estymatorów tzn. wielkości

$$(3.2.3) \quad CV(\hat{t}) = \frac{\sqrt{\hat{V}(\hat{t})}}{\hat{t}}$$

Zalecana w literaturze przedmiotu wartość parametru B powinna być rzędu kilkuset. W omawianych tutaj obliczeniach przyjęto $B=500$.

Obecnie zajmujemy się omówieniem precyzji uzyskanych szacunków za pomocą trzech opisanych wcześniej estymatorów. Za miarę precyzji przyjęliśmy współczynnik zmienności estymatora określony wzorem (3.2.3). Współczynniki te zostały obliczone dla wszystkich ocen parametrów zamieszczonych w tablicach wynikowych dotyczących lat 1995—2002. Dla województw i podregionów oraz wybranych powiatów¹⁾ zostały zamieszczone wraz z ocenami parametrów w tablicach podanych w aneksie. Wspomniane tablice zawierają kilkadzie-

¹⁾ Dotyczy to powiatów, dla których liczebność próby w roku 2002 wynosiła co najmniej 50 jednostek pierwszego stopnia.

siąg tysięcy liczb i w związku z tym ich analiza jest praktycznie niewykonalna. Dlatego też, posłużyliśmy się metodą syntetyczną polegającą na obliczaniu decyli dla poszczególnych tablic, a w ich ramach dla województw i podregionów. Wartości tych decyli zawiera poniższa tablica.

Tablica 3.5. Decyle współczynników zmienności ocen parametrów dla województw i podregionów (lata 1995—2002)

Wyszczególnienie			decyle								
			d1	d2	d3	d4	d5	d6	d7	d8	d9
Tablica 1	Województwo	K	2.4	2.7	2.9	3.2	3.5	3.8	4.4	6.3	8.7
		S	0.7	0.7	0.7	0.8	0.8	0.8	0.9	1.7	2.1
		Z	1.4	1.5	1.6	1.7	1.9	2.0	2.2	3.6	4.7
	Podregion	K	3.9	4.4	4.8	5.4	5.9	6.5	7.7	10.6	14.6
		S	2.3	2.5	2.7	2.9	3.2	3.5	4.1	6.0	8.1
		Z	3.0	3.3	3.5	3.7	4.0	4.4	5.0	7.7	10.3
Tablica 2	Województwo	K	4.2	4.8	5.3	5.8	6.3	6.9	7.8	9.0	11.0
		S	1.1	1.2	1.3	1.4	1.5	1.7	1.8	2.1	2.8
		Z	2.3	2.6	2.9	3.1	3.4	3.8	4.1	4.9	6.0
	Podregion	K	6.7	7.7	8.7	9.7	10.8	11.9	13.5	15.7	19.8
		S	3.9	4.5	5.0	5.4	5.9	6.4	7.2	8.5	10.3
		Z	4.9	5.7	6.3	6.9	7.5	8.2	9.2	10.7	13.2
Tablica 3	Województwo	K	2.6	3.0	3.3	3.6	3.9	4.3	4.7	5.2	6.0
		S	0.7	0.8	0.8	0.9	1.0	1.0	1.1	1.2	1.4
		Z	1.5	1.6	1.8	2.0	2.1	2.3	2.6	2.8	3.2
	Podregion	K	4.2	4.9	5.5	6.1	6.6	7.3	8.0	9.0	10.6
		S	2.4	2.8	3.0	3.4	3.7	4.0	4.4	4.9	5.6
		Z	3.1	3.5	3.9	4.3	4.7	5.1	5.5	6.2	7.1
Tablica 4	Województwo	K	3.0	3.6	4.1	4.6	5.2	5.8	6.7	7.6	8.9
		S	0.8	0.9	1.0	1.1	1.3	1.4	1.7	1.9	2.1
		Z	1.7	2.0	2.2	2.5	2.8	3.2	3.7	4.1	4.7
	Podregion	K	4.9	5.9	6.8	7.7	8.7	9.8	11.3	13.2	16.2
		S	2.8	3.3	3.8	4.3	4.9	5.5	6.3	7.2	8.3
		Z	3.6	4.3	4.9	5.5	6.2	7.0	8.0	9.1	10.6
Tablica 5	Województwo	K	5.3	6.3	7.5	8.6	9.9	11.5	13.8	19.1	30.9
		S	1.4	1.6	1.8	2.2	2.5	2.8	3.3	4.0	7.5
		Z	2.9	3.5	4.1	4.7	5.4	6.2	7.3	10.1	16.7
	Podregion	K	8.4	10.5	12.5	14.5	16.7	19.6	24.2	32.1	50.2
		S	5.0	6.0	7.1	8.2	9.3	10.9	13.0	17.9	28.8
		Z	6.3	7.7	9.0	10.4	11.9	13.7	16.5	22.2	35.6
Tablica 6	Województwo	K	2.9	3.4	4.0	4.6	5.4	6.4	7.7	10.1	14.5
		S	0.7	0.8	0.9	1.2	1.4	1.6	1.8	2.4	3.5
		Z	1.6	1.8	2.1	2.5	2.9	3.4	4.1	5.5	7.6
	Podregion	K	4.7	5.6	6.7	7.8	9.2	10.8	13.4	17.3	25.7
		S	2.7	3.2	3.7	4.3	5.0	5.9	7.2	9.4	13.1
		Z	3.4	4.1	4.7	5.5	6.4	7.6	9.1	12.1	17.5

Z informacji podanych w powyższej tabelicy wynika, że najmniej precyzyjne są estymatory klasyczne (K), najlepsze zaś są estymatory syntetyczne (S). Precyzja estymatorów złożonych (Z) zawiera się pomiędzy precyzjami estymatorów S i K. Można by wnioskować z tego, że estymatory syntetyczne są najefektywniejsze. Nie jest to jednak prawdą, gdyż są one obciążone (obciążenia tego nie da się w żaden sposób ustalić). Gdyby możliwe było przy obliczaniu współczynnika zmienności uwzględnienie obciążenia, to mogłoby się okazać, że estymatory te nie są lepsze od estymatorów złożonych. Dlatego też, preferujemy estymatory złożone, które wykorzystują informacje pochodzące z dwóch źródeł. Jednym z nich jest badanie aktywności ekonomicznej ludności, zaś drugim NSP 2002 r.

Ponadto, z danych zawartych w powyższej tabelicy wynika generalny wniosek, że precyzja dla województw jest znacznie wyższa od precyzji dla podregionów. Jest to oczywiste, gdyż liczebność próby (liczona liczbą jednostek pierwszego stopnia) dla podregionów jest znacznie niższa niż dla województw (z wyjątkiem województw opolskiego i świętokrzyskiego, które są jednocześnie podregionami).

Rozważmy precyzję dla grup określających aktywność ekonomiczną. Najbardziej precyzyjne oceny uzyskujemy dla pracujących wg płci i wieku (tablica 1). Nieco gorszą precyzję otrzymujemy dla biernych zawodowo wg płci i wieku (tablica 3). Najniższą precyzję uzyskujemy dla danych odnoszących się do bezrobotnych według płci i wykształcenia (tablica 5). Należy przypuszczać, że wynika to z tego, że bezrobotni są grupą znacznie mniej liczną niż pozostałe. Bezrobotni stanowili w roku 2002 ok. 11% ludności w wieku 15 lat i więcej.

Dodatkowo dla roku 2002 prezentowane są w poniższej tabelicy decyle współczynników zmienności estymatorów interesujących nas parametrów. Z porównania danych z tablic 3.5 i 3.5a wynika, że precyzja dla 2002 roku jest zbliżona do precyzji dla całego rozpatrywanego okresu.

Tablica 3.5a. Decyle współczynników zmienności ocen parametrów dla województw i podregionów (rok 2002).

Wyszczególnienie		Decyle									
		d1	d2	d3	d4	D5	d6	d7	d8	d9	
Tablica 1	Województwo	K	2,7	3,0	3,1	3,4	3,6	3,8	4,2	7,5	9,7
		S	0,8	0,8	0,8	0,8	0,9	0,9	0,9	2,0	2,4
		Z	1,5	1,6	1,7	1,8	2,0	2,1	2,2	4,1	5,2
	Podregion	K	4,2	4,5	5,1	5,7	6,1	6,9	8,0	11,4	15,7
		S	2,5	2,7	2,9	3,0	3,3	3,7	4,0	6,9	8,6
		Z	3,3	3,5	3,7	3,9	4,3	4,6	5,0	8,6	11,2
Tablica 2	Województwo	K	4,4	4,8	5,2	5,4	5,8	6,3	7,2	8,4	10,2
		S	1,1	1,2	1,4	1,4	1,4	1,6	1,6	2,0	2,6
		Z	2,3	2,6	2,8	2,9	3,1	3,4	3,8	4,5	5,5
	Podregion	K	6,1	7,0	7,8	8,9	9,9	11,0	12,2	14,5	17,8
		S	3,9	4,2	4,6	5,0	5,3	5,8	6,4	7,4	9,3
		Z	4,6	5,3	5,8	6,4	6,8	7,5	8,4	9,6	11,7
Tablica 3	Województwo	K	2,6	3,0	3,2	3,5	3,8	4,1	4,6	4,9	5,8
		S	0,7	0,7	0,7	0,8	0,9	1,0	1,2	1,2	1,4
		Z	1,4	1,6	1,8	1,9	2,0	2,2	2,5	2,7	3,1
	Podregion	K	4,2	4,7	5,3	5,8	6,4	7,1	7,8	8,8	10,2
		S	2,5	2,7	3,0	3,2	3,5	3,8	4,2	4,7	5,3
		Z	3,2	3,5	3,8	4,1	4,5	4,9	5,3	5,9	6,8
Tablica 4	Województwo	K	3,1	3,8	4,3	5,1	5,7	6,3	7,0	7,6	8,8
		S	0,8	1,0	1,1	1,2	1,3	1,6	1,9	2,0	2,2
		Z	1,8	2,1	2,3	2,7	3,0	3,5	3,9	4,2	4,7
	Podregion	K	5,3	6,3	7,3	8,2	9,3	10,6	11,9	14,1	17,8
		S	3,2	3,6	4,1	4,8	5,3	6,0	6,7	7,4	8,6
		Z	3,9	4,6	5,3	5,9	6,7	7,5	8,5	9,6	11,0
Tablica 5	Województwo	K	5,1	6,0	7,1	8,2	9,2	10,4	12,8	15,5	21,8
		S	1,5	1,6	1,8	2,1	2,5	2,7	3,5	4,1	5,8
		Z	2,8	3,4	3,8	4,4	5,1	5,9	6,7	8,7	11,6
	Podregion	K	8,1	10,2	11,9	13,8	15,7	19,1	22,6	28,7	38,9
		S	5,0	5,7	6,7	7,7	8,9	9,9	12,4	15,1	21,2
		Z	6,2	7,3	8,5	9,8	11,4	13,1	15,6	19,6	27,0
Tablica 6	Województwo	K	2,9	3,4	3,9	4,4	5,1	5,9	6,9	9,1	12,8
		S	0,7	0,9	0,9	1,2	1,3	1,5	1,7	2,3	3,3
		Z	1,6	1,9	2,1	2,5	2,8	3,3	3,8	5,1	6,8
	Podregion	K	4,7	5,7	6,6	7,7	8,9	10,3	12,5	15,9	23,0
		S	2,8	3,3	3,8	4,3	4,8	5,7	6,4	8,2	12,1
		Z	3,5	4,1	4,7	5,5	6,3	7,3	8,4	10,8	16,3

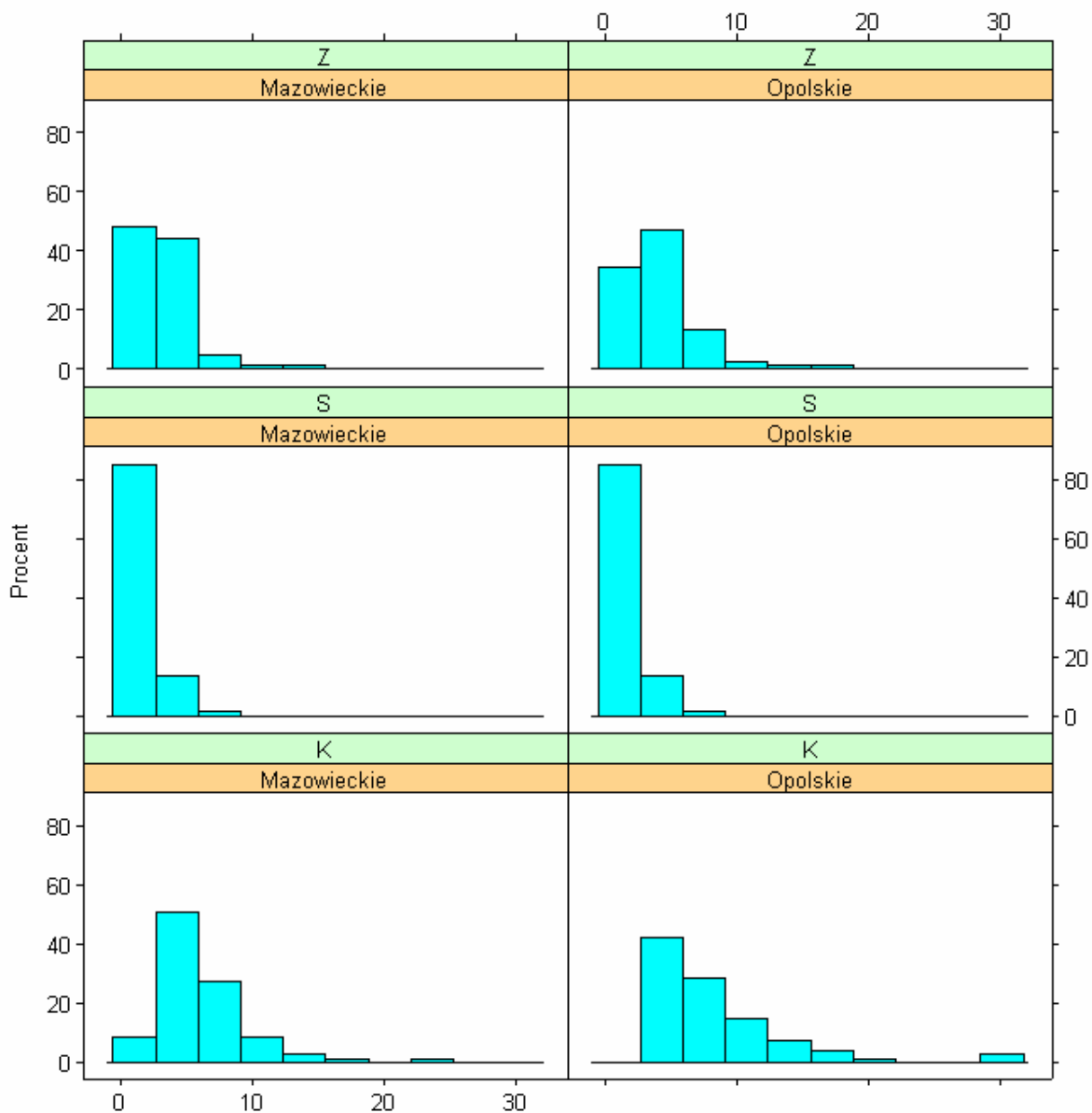
Niezależnie od prezentowanych decyli współczynników zmienności (CV) obliczone zostały częstości w dziesięciu przedziałach klasowych. Dolną granicą pierwszego przedziału jest minimalna wartość CV, natomiast górną wartość maksymalna. Rozstęp został podzielony na 10 równej długości przedziałów klasowych. Częstości występowania wartości CV należących do poszczególnych klas przedstawia poniższa tablica.

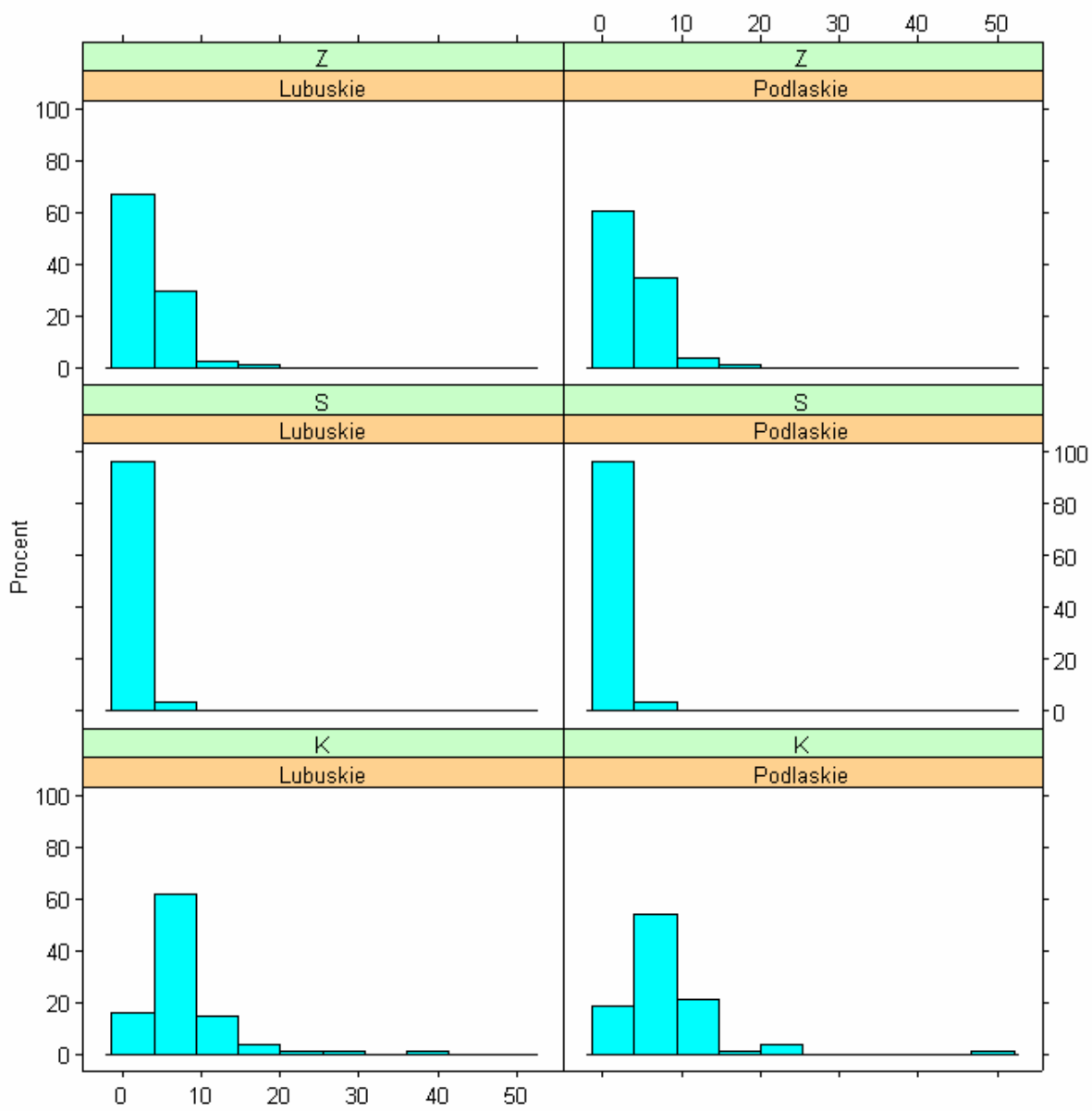
Tablica 3.5b Rozkłady współczynników zmienności ocen parametrów (rok 2002)

Wyszczególnienie				Przedziały klasowe										
				min	max	1	2	3	4	5	6	7	8	9
Tablica 1	Województwo	K	2,4	14,6	53	22	1	3	5	6	4	3	2	2
		S	0,8	2,7	75	0	0	0	0	0	8	0	8	8
		Z	1,3	7,2	45	30	0	1	6	4	7	3	2	3
	Podregion	K	3,2	39,6	58	19	9	6	4	2	1	1	0	0
		S	2,2	14,6	56	19	1	5	8	4	2	1	2	0
		Z	2,6	22,6	63	12	5	8	6	4	1	1	0	0
Tablica 2	Województwo	K	3,2	15,2	10	34	22	10	8	6	4	2	1	2
		S	1,1	2,7	25	23	19	8	0	8	0	0	0	17
		Z	1,8	7,6	13	30	23	13	4	4	9	2	2	2
	Podregion	K	3,9	33,7	19	30	25	13	6	4	3	1	0	0
		S	3,1	15,2	21	32	20	7	8	6	3	1	1	0
		Z	3,2	18,3	13	25	26	13	8	6	3	3	1	1
Tablica 3	Województwo	K	2,2	7,6	11	20	19	16	14	6	6	5	3	2
		S	0,6	1,5	25	8	17	8	8	0	17	0	0	17
		Z	1,2	4,0	11	18	21	14	13	7	7	5	2	3
	Podregion	K	2,7	19,1	14	29	25	16	9	3	2	1	0	0
		S	2,0	7,8	16	21	22	15	10	10	4	3	1	1
		Z	2,2	10,4	6	26	24	17	12	7	5	2	1	1
Tablica 4	Województwo	K	2,4	12,2	13	19	14	15	13	13	7	5	2	0
		S	0,8	2,3	20	0	20	13	0	7	7	7	13	13
		Z	1,4	6,2	12	16	16	12	10	15	10	7	1	2
	Podregion	K	3,1	33,2	19	29	23	12	7	5	2	1	1	0
		S	2,2	12,7	11	21	18	16	16	7	7	2	1	0
		Z	2,5	17,6	11	23	20	17	14	8	4	3	0	0
Tablica 5	Województwo	K	3,1	50,1	38	32	12	9	4	3	1	1	0	0
		S	1,1	7,6	27	20	20	13	0	7	0	7	0	7
		Z	1,8	19,8	26	28	19	9	5	4	4	3	0	1
	Podregion	K	4,0	103,4	41	31	14	7	3	2	1	0	0	1
		S	3,2	47,8	39	31	13	7	5	2	1	1	0	0
		Z	3,3	62,3	35	33	14	8	5	2	1	2	0	0
Tablica 6	Województwo	K	2,1	20,3	30	28	17	5	5	5	4	2	2	1
		S	0,7	3,7	33	13	13	13	7	0	7	0	7	7
		Z	1,2	10,7	31	26	18	5	5	6	4	3	2	1
	Podregion	K	2,4	72,4	53	28	9	4	3	1	1	0	0	0
		S	2,0	20,9	31	29	17	5	5	7	2	3	0	1
		Z	2,0	33,1	35	34	12	5	6	3	3	1	0	0

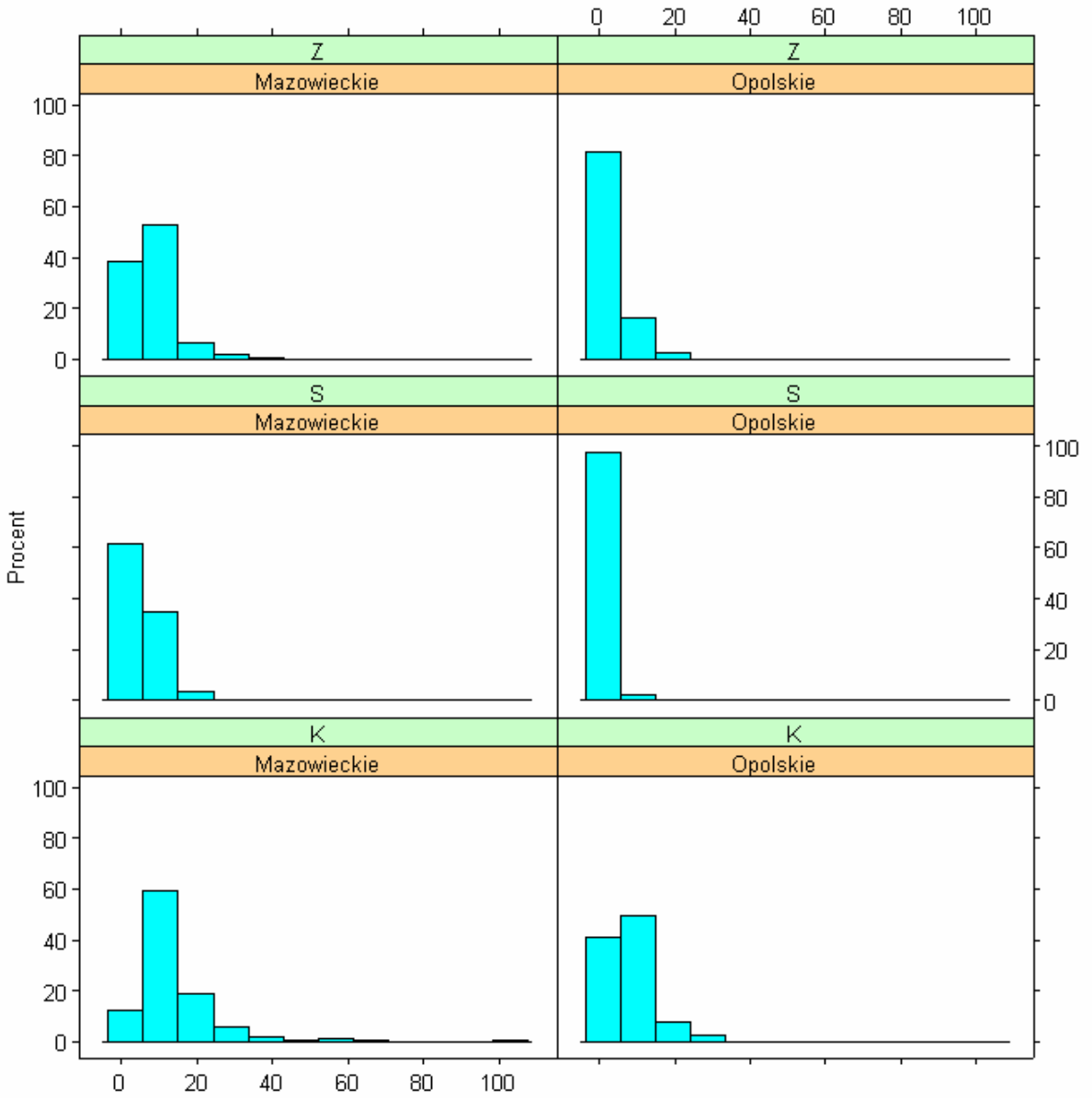
Niezależnie od wyników przedstawionych w powyższej tabelicy przedstawiamy histogramy częstości występowania współczynników zmienności należących do 10 przedziałów klasowych. Histogramy te zostały wykonane dla trzech rozpatrywanych w pracy estymatorów oraz dla czterech wybranych województw. Rys.1 zawiera histogramy współczynników zmienności ocen parametrów dla wybranych czterech województw, zaś drugi dla podregionów w tych województwach. Do prezentacji graficznej wybrano województwa: mazowieckie jako duże pod względem obszaru i ludności, opolskie jako jedno z mniejszych, zaś podlaskie i lubuskie jako „reprezentanty” województw odpowiednio wschodnich i zachodnich.

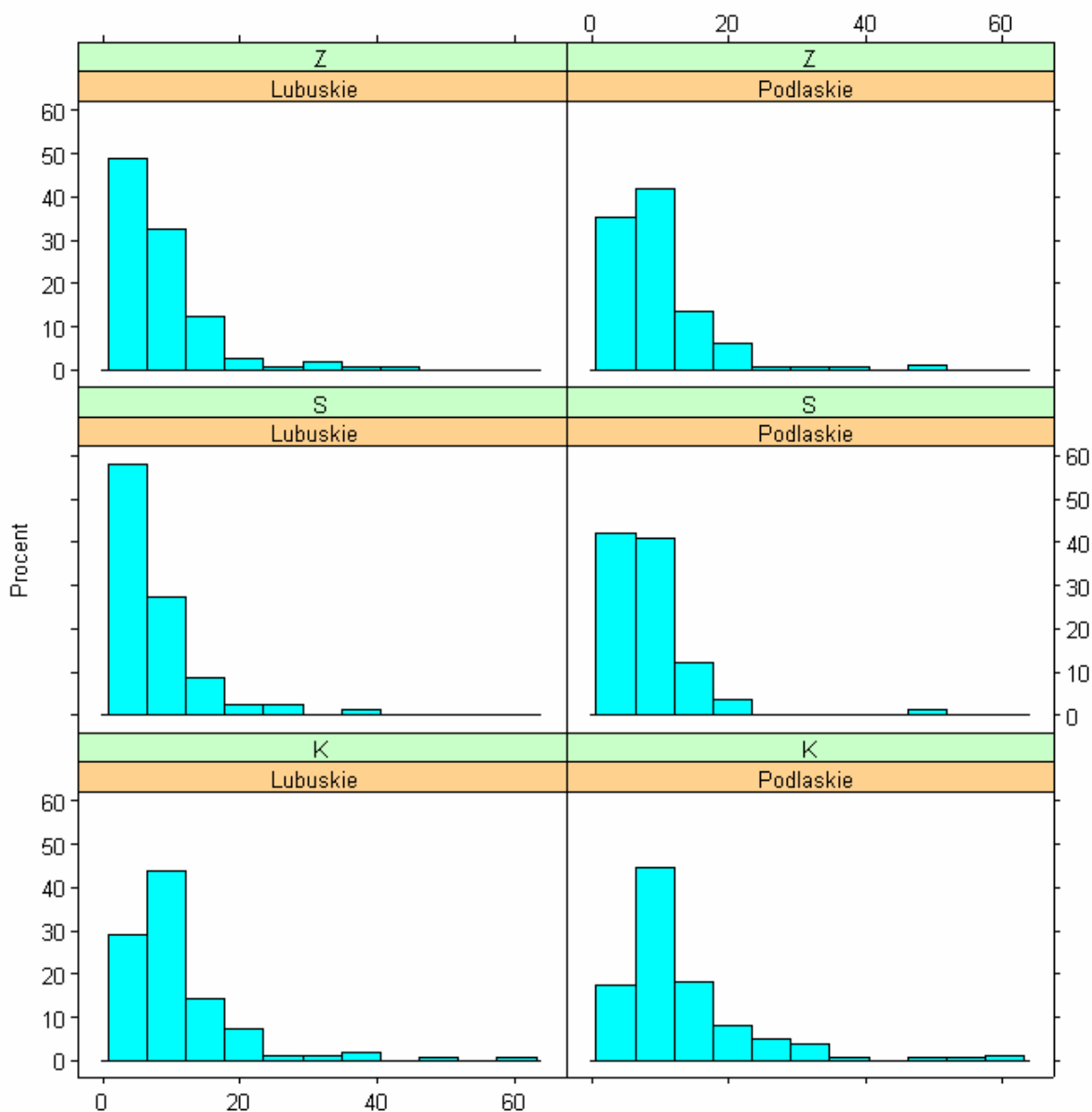
Rys. 1. Rozkłady współczynników zmienności ocen parametrów dla wybranych województw i roku 2002.





Rys. 2. Rozkłady współczynników zmienności ocen parametrów dla podregionów w wybranych województwach w roku 2002





Z tablicy 3.5b oraz rysunków 1 i 2 zawierających histogramy częstości występowania współczynników zmienności o wartościach należących do poszczególnych przedziałów klasowych wynika, że w rozkładach tych współczynników występuje zdecydowana prawostronna asymetria. Najwyższe częstości występują na ogół do pierwszych trzech klas szeregu rozdzielczego. Świadczy to o tym, że przypadki niskiej precyzji szacunków dla województw i podregionów występują bardzo rzadko.

Jak już wcześniej wspomniano, dla powiatów mniejszych można zastosować tylko estymator syntetyczny. Współczynnik zmienności tego estymatora jest równy współczynnikowi zmienności estymatora klasycznego dla województwa. Wynika to z definicji estymatora syntetycznego. Dlatego też nie ma potrzeby ich publikacji.

4. Zgodność estymatorów klasycznych i wykorzystujących dane z NSP 2002

W punkcie 1.1 zwróciliśmy uwagę na to, że dane dotyczące aktywności ekonomicznej ludności pochodzące z NSP 2002 będą wnosić co raz mniej informacji o aktualnych liczbach bezrobotnych, pracujących i zawodowo biernych wraz z upływem czasu. Dlatego też zasadnym wydaje się wprowadzenie współczynnika zgodności danych pochodzących z danych bieżących BAEL otrzymywanymi metodami klasycznymi, a danymi otrzymywanymi metodami wykorzystującymi dane z NSP za pośrednictwem pewnych modeli statystycznych. Naturalne jest pytanie jak długo dane z NSP 2002 można będzie wykorzystywać w celu poprawienia jakości estymacji parametrów dla kolejnych lat.

Rozsądną miarą takiej zgodności jednych i drugich danych dla s -tego subregionu w w -tym województwie wydają się następujące współczynniki:

a) dla estymatora syntetycznego (S)

$$(4.1.1) \quad z_{ws} = \frac{x_{ws} - t_{ws}}{t_{ws}},$$

gdzie:

— t_{ws} jest estymatorem klasycznym (K) określonego parametru w s -tym podregionie w -tego województwa,

— x_{ws} jest określone wzorem (3.1.1),

b) dla estymatora złożonego (Z)

$$(4.2.2) \quad u_{ws} = \frac{y_{ws} - t_{ws}}{t_{ws}} = 0,5z_{ws},$$

gdzie y_{ws} określone jest wzorem (3.1.3). Druga równość we wzorze (4.2.2) wynika z przyjęcia założenia, że waga v_{ws} jest równa 0,5.

Współczynniki dla województw i dużych powiatów określone są analogicznymi wzorami (różnica polega tylko na zmianie indeksów). Poza współczynnikami zgodności określonymi wzorami (4.1.1) i (4.1.2) obliczone zostały współczynniki zgodności określone wzorem

$$(4.1.3) \quad b_{ws} = \frac{x_{ws} - y_{ws}}{y_{ws}}.$$

Wcześniej wspomnieliśmy, że przy konstrukcji estymatora złożonego (por. wzór (3.1.3)) przyjęliśmy wagę v_{ws} równą 0,5. Wykorzystując ten fakt wzór (4.1.3) można zapisać w postaci

$$(4.1.4) \quad b_{ws} = \frac{x_{ws} - t_{ws}}{x_{ws} + t_{ws}}.$$

Oznacza to, że współczynnik zgodności estymatora syntetycznego (S) i złożonego (Z) jest w tym szczególnym przypadku równy ilorazowi różnicy estymatora syntetycznego i klasycznego (K) przez ich sumę. Z porównania wzorów (4.1.1) i (4.1.4) wynika nierówność

$$(4.1.5) \quad b_{ws} < z_{ws}.$$

Inaczej mówiąc wykorzystywanie w równych częściach informacji z BAEL i NSP daje lepsze oceny niż opieranie się tylko na jednym źródle wiedzy o aktywności ekonomicznej ludności. Powyższy wniosek dotyczy sytuacji, gdy $v_{ws}=0,5$. Wspominaliśmy już o doborze wartości wagi v_{ws} w pkt. 3.1, ale chcielibyśmy jeszcze raz zwrócić uwagę na konieczność przeprowadzenia badań symulacyjnych, które dałyby odpowiedź na pytanie jaką wartość tej wagi należy przyjąć i jak ją zmieniać z upływem czasu.

Poniżej podajemy tablicę, której zawartość pozwoli ocenić na ile dane z BAELu i NSP są zgodne dla roku 2002. Podano w niej względne różnice między ocenami parametrów otrzymanych na podstawie BAEL dla 2002 a odpowiednimi wartościami z NSP.

Tablica 4.1. Ocena zgodności danych otrzymanych w BAELu 2002 z danymi zebranymi w NSP 2002 r. (w %%)

Województwa, podregiony (NTS 3)	Ogółem	Ogółem bez niestabilnych	Pracujący	Bezrobotni	Bierni zawodowo
(1)	(2)	(3)	(4)	(5)	(6)
Polska	-0,7	2,7	4,3	-3,6	2,9
Dolnośląskie	-5,9	-3,2	-3,3	-0,8	-3,9
Podregion 1 - jeleniogórsko-wałbrzyski	-4,8	-2,2	-5,0	0,1	-0,9
Podregion 2 – legnicki	9,8	13,2	21,1	13,1	5,6
Podregion 3 – wrocławski	-19,3	-17,7	-20,7	-11,1	-16,8
Podregion 4 - m. Wrocław	-11,3	-8,2	-6,8	-9,4	-9,2
Kujawsko-Pomorskie	8,4	10,3	13,4	-2,0	11,2
Podregion 5 – bydgoski	-0,1	1,7	2,6	-7,0	3,3
Podregion 6 - toruńsko-włocławski	16,7	18,7	24,0	2,4	18,9
Lubelskie	6,5	8,7	17,6	2,8	0,9
Podregion 7 – białkopodlaski	23,2	24,6	25,7	34,6	20,9
Podregion 8 - chełmsko-zamojski	0,4	2,5	37,2	-24,0	-26,2
Podregion 9 – lubelski	5,6	8,0	5,1	9,2	10,6

Województwa, podregiony (NTS 3)	Ogółem	Ogółem bez nieustalonych	Pracujący	Bezrobotni	Bierni zawodowo
(1)	(2)	(3)	(4)	(5)	(6)
Lubuskie	8,0	10,7	11,2	6,5	11,8
Podregion 10 – gorzowski	11,9	14,7	20,9	6,4	11,8
Podregion 11 – zielonogórski	5,7	8,3	5,0	6,6	11,7
Łódzkie	10,8	12,9	12,4	13,1	13,3
Podregion 12 – łódzki	5,9	7,0	5,7	8,0	8,0
Podregion 13 - piotrkowsko-skierniewicki	8,5	9,6	9,8	16,0	7,8
Podregion 14 - m. Łódź	18,6	23,2	24,0	15,8	24,4
Małopolskie	-0,7	3,2	12,0	-8,8	-2,5
Podregion 15 - krakowsko-tarnowski	-4,9	-2,0	8,8	-12,5	-10,2
Podregion 16 – nowosądecki	6,8	10,8	17,7	-1,1	7,5
Podregion 17 - m. Kraków	-3,3	2,1	10,5	-14,3	-2,6
Mazowieckie	-5,8	-2,4	-1,8	-4,6	-2,6
Podregion 18 - ciechanowsko-płocki	14,4	15,6	17,4	14,4	14,0
Podregion 19 - ostrołęcko-siedlecki	24,4	25,9	33,5	27,8	16,7
Podregion 20 – warszawski	-25,7	-24,4	-26,2	-32,6	-20,3
Podregion 21 – radomski	-9,0	-7,8	-7,8	-3,3	-9,3
Podregion 22 - m. Warszawa	-9,5	-2,1	-1,7	-9,7	-1,2
Opolskie	-9,3	1,3	1,7	-9,5	3,7
Podregion 23 – opolski	-9,3	1,3	1,7	-9,5	3,7
Podkarpackie	-3,8	1,0	11,4	-8,9	-6,3
Podregion 24 - rzeszowsko-tarnobrzeski	-9,2	-3,7	6,8	-11,4	-12,1
Podregion 25 - krośnieńsko-przemyski	2,7	6,5	17,2	-6,2	0,4
Podlaskie	-2,8	3,0	7,3	-8,2	1,4
Podregion 26 - białostocko-suwalski	-0,8	4,7	7,9	-1,3	2,8
Podregion 27 – łomżyński	-8,7	-1,9	5,4	-28,5	-2,9
Pomorskie	-7,5	-4,3	-4,3	-8,5	-3,1
Podregion 28 – słupski	-5,9	-3,8	-4,8	11,3	-8,4
Podregion 29 – gdański	-4,5	-2,1	-2,1	-8,9	-0,1
Podregion 30 – Gdańsk-Gdynia-Sopot	-11,8	-7,2	-6,5	-28,5	-3,3
Śląskie	-3,8	1,1	-2,6	-6,3	6,0
Podregion 31 – częstochowski	-6,9	-5,4	-11,3	-1,5	-1,0
Podregion 32 - bielsko-bialski	23,5	26,5	21,0	0,5	37,4
Podregion 33 - centralny śląski	-12,6	-7,2	-9,6	-12,3	-3,9
Podregion 45 - rybnicko-jastrzębski	12,2	19,2	12,5	13,4	25,8
Świętokrzyskie	9,5	12,0	7,7	-12,8	23,3
Podregion 34 – świętokrzyski	9,5	12,0	7,7	-12,8	23,3
Warmińsko-Mazurskie	-2,6	0,7	1,1	-9,8	4,0
Podregion 35 – elbląski	8,0	10,2	8,3	-19,8	22,6
Podregion 36 – olsztyński	0,4	4,3	8,9	-5,9	3,4
Podregion 37 – etcki	-29,4	-26,0	-32,6	0,9	-30,7
Wielkopolskie	-0,1	1,3	1,6	-1,1	1,5
Podregion 38 – pilski	1,3	2,2	1,7	27,1	-4,5
Podregion 39 – poznański	-6,7	-6,0	-7,7	-13,8	-2,3

Województwa, podregiony (NTS 3)	Ogółem	Ogółem bez nieustalonych	Pracujący	Bezrobotni	Bierni zawodowo
(1)	(2)	(3)	(4)	(5)	(6)
Podregion 40 – kaliski	2,8	3,6	2,4	-7,3	7,8
Podregion 41 – koniński	28,1	29,1	31,2	26,7	27,6
Podregion 42 - m. Poznań	-12,4	-8,9	-2,1	-20,2	-14,1
Zachodniopomorskie	1,1	4,1	2,1	-4,3	8,7
Podregion 43 – szczeciński	-3,6	-0,5	-2,6	-5,4	3,1
Podregion 44 – koszaliński	10,0	12,7	11,6	-2,5	19,4

Z przedstawionych w powyższej tabelicy danych wynika, że populacja ludności w wieku 15 lat i więcej badanej w NSP niewiele różni się od populacji rozważanej jako przeciętna z czterech kwartałów w BAEL. Różnica wynosi -0.7%. Jednak dla niektórych województw, a zwłaszcza podregionów różnice te są znaczne. Dotyczy to przede wszystkim dużych aglomeracji miejskich (m.in. Łódź, Warszawa, Poznań, Wrocław oraz podregion Gdańsk-Gdynia-Sopot). Dla poszczególnych kategorii ludności (pracujący, bezrobotni, bierni zawodowo) rozbieżności te zdarzają się większe niż dla ludności ogółem. Omawiane różnice spowodowane są przede wszystkim dwiema przyczynami: po pierwsze, metoda estymacji stosowana w BAEL preferuje wyniki dla Polski, zaś dane dla województw są rzeczą wtórną. Jak już wspomniano wcześniej, odsetek zbadanych mieszkań w BAEL nie przekraczał 80%, zaś w przypadku dużych miast odsetek ten był mniejszy. Z tego powodu, przy uogólnianiu wyników, stosuje się wagi (por. pkt 2.3) uwzględniające kompletność badania w sześciu klasach miejscowości. Wagi te są identyczne dla wszystkich mieszkań w danej klasie. Kompletność badania jest jednak zróżnicowana geograficznie w danej klasie, co prowadzi do przeszacowania lub niedoszacowania wyników dla poszczególnych województw i podregionów. Po drugie, w NSP 2002 r dla ponad 1 mln osób nie został określony ich status na rynku pracy. Przykładowo w woj. opolskim dotyczy to ponad 10%, podczas gdy w woj. wielkopolskim dotyczy to nieco ponad 1%. Ponadto, w niektórych województwach i podregionach liczba takich osób jest znaczna w stosunku do liczby bezrobotnych, a nawet ją przekracza np. w Warszawie.

Obie podane wyżej okoliczności mają wpływ na wartość i znak tych współczynników zgodności, a w konsekwencji negatywnie oddziałują na jakość estymacji dla małych obszarów (por. J. Kordos (2002)).

Dla wszystkich ocen parametrów dla województw i podregionów zostały obliczone współczynniki zgodności i umieszczone w tablicach od 1 do 6 (patrz aneks) dla lat 1995—2002. Współczynniki zgodności obliczone zostały według wzorów (4.1.1) oraz (4.1.3). Nie były liczone natomiast współczynniki zgodności ocen otrzymanych za pomocą

estymatorów klasycznych (K) i złożonych (Z), gdyż jak wynika z równości (4.2.2) współczynniki te są równe połowie współczynników wyrażonych wzorem (4.1.1). Wspomniane wyżej tablice zawierają ponad 20 tys. liczb. Analiza tak dużej ilości współczynników jest niewykonalna. W związku z tym obliczone zostały wartości decyli dla każdej z wymienionych wyżej tablic, województw i podregionów oraz dla dwóch rodzajów współczynników zgodności. Wartości tych decyli zawarte są w poniższej tablicy.

Tablica 4.2. Wartości decyli dla współczynników zgodności dla z lat 1995—2002

			d1	d2	d3	d4	D5	d6	d7	d8	d9
T1	woj	SK	-10,1	-7,5	-5,8	-4,2	-1,3	1,0	3,4	6,4	11,3
T1	podreg	SK	-12,5	-7,4	-4,2	-1,7	0,0	1,6	3,8	6,7	11,2
T1	woj	SZ	-5,3	-3,9	-3,0	-2,1	-0,7	0,5	1,7	3,1	5,4
T1	podreg	SZ	-6,7	-3,9	-2,1	-0,9	0,0	0,8	1,9	3,2	5,3
T2	woj	SK	-14,0	-9,7	-6,4	-3,2	-0,4	2,5	5,3	8,6	14,9
T2	podreg	SK	-20,5	-12,4	-6,8	-3,0	0,0	2,7	7,2	14,3	27,9
T2	woj	SZ	-7,5	-5,1	-3,3	-1,6	-0,2	1,3	2,6	4,1	7,0
T2	podreg	SZ	-11,4	-6,6	-3,5	-1,5	0,0	1,3	3,5	6,7	12,2
T3	woj	SK	-8,6	-5,8	-3,9	-2,0	-0,1	2,0	4,5	6,9	10,3
T3	podreg	SK	-11,3	-7,2	-4,1	-1,7	0,0	1,5	3,9	6,9	13,2
T3	woj	SZ	-4,5	-3,0	-2,0	-1,0	0,0	1,0	2,2	3,3	4,9
T3	podreg	SZ	-6,0	-3,7	-2,1	-0,9	0,0	0,7	1,9	3,4	6,2
T4	woj	SK	-10,4	-7,4	-5,1	-2,9	-0,9	1,2	3,5	6,7	11,8
T4	podreg	SK	-13,0	-8,2	-4,9	-2,0	0,0	2,0	5,0	9,4	16,6
T4	woj	SZ	-5,5	-3,9	-2,6	-1,5	-0,4	0,6	1,7	3,2	5,5
T4	Podreg.	SZ	-7,0	-4,3	-2,5	-1,0	0,0	1,0	2,4	4,5	7,7
T5	Woj.	SK	-19,2	-12,1	-7,6	-3,4	0,0	3,6	7,7	14,5	26,7
T5	Podreg.	SK	-24,3	-14,9	-8,7	-3,4	0,0	3,8	9,6	18,5	36,5
T5	Woj.	SZ	-10,6	-6,5	-3,9	-1,7	0,0	1,8	3,7	6,8	11,8
T5	Podreg.	SZ	-13,7	-7,9	-4,4	-1,5	0,0	2,1	5,0	9,1	17,3
T6	Woj.	SK	-10,3	-6,6	-4,0	-2,0	0,4	2,9	5,5	9,0	14,7
T6	Podreg.	SK	-14,8	-8,9	-4,9	-2,0	0,0	2,6	6,1	10,5	20,8
T6	woj	SZ	-5,4	-3,4	-2,1	-1,0	0,2	1,4	2,7	4,3	6,8
T6	podreg	SZ	-8,0	-4,7	-2,5	-1,0	0,0	1,3	3,0	5,0	9,4

Tablice 2 i 5 zawarte w aneksie dotyczą bezrobotnych. Jest to grupa najmniejsza wśród grup określających aktywność ekonomiczną. Z tego wynika, że zgodność szacunków dla tej grupy jest najmniejsza. Ponadto, dla tej grupy mamy największą dyspersję pomiędzy poszczególnymi rodzajami ocen. Rzutuje to na wartości współczynników zgodności, o czym informują pierwsze i dziewiąte decyle w powyższej tablicy. Dla pozostałych grup aktywności ekonomicznej skrajne decyle mają zbliżone wartości.

Z powyższej tablicy wynika również, że decyle dla współczynników zgodności estymatorów syntetycznych i złożonych (SZ) są znacznie niższe od decyli dla współczynników zgodności estymatorów syntetycznych i klasycznych (SK). Jest to zgodne z podaną wcześniej nierównością (4.1.5). Potwierdza to fakt, że estymator złożony wykorzystujący informację z dwóch źródeł (BAEL i NSP) jest efektywniejszy od estymatora klasycznego opartego wyłącznie na danych z BAEL.

Dla podregionów decyle mają większą wartość niż dla województw, co jest rzeczą oczywistą, gdyż estymatory dla podregionów szacowane były w oparciu o mniejszą próbę.

Analogicznie jak dla współczynników zmienności, również dla współczynników zgodności przeanalizowano dodatkowo oceny parametrów dotyczące roku 2002. Obliczone zostały wartości decyli dla współczynników zgodności (tab.4.2.a) oraz szeregi rozdzielcze dla tych współczynników (tab..4.2.b). Ponadto dla wybranych województw prezentujemy histogramy częstości występowania wartości współczynników zgodności należących do poszczególnych przedziałów klasowych (rys. 3 i 4).

Tablica 4.2a. Wartości decyli dla współczynników zgodności (rok 2002)

Wyszczególnienie			d1	d2	d3	d4	d5	d6	d7	d8	d9
Tab. 1	Wojew.	SK	-8,8	-7,2	-6,2	-3,3	0,9	2,6	5,1	7,2	8,8
		SZ	-4,6	-3,7	-3,2	-1,7	0,4	1,3	2,5	3,5	4,2
	Podregion	SK	-17,8	-9,2	-6,5	-2,0	0,0	2,9	5,2	8,5	14,5
		SZ	-9,8	-4,8	-3,4	-1,0	0,0	1,4	2,5	4,1	6,8
Tab. 2	Wojew.	SK	-9,2	-5,8	-2,8	-0,8	1,8	3,9	5,6	7,1	9,1
		SZ	-4,8	-3,0	-1,4	-0,4	0,9	1,9	2,7	3,4	4,4
	Podregion	SK	-20,2	-10,9	-6,2	-2,9	0,0	1,8	5,8	12,2	25,8
		SZ	-11,2	-5,8	-3,2	-1,5	0,0	0,9	2,8	5,7	11,4
Tab. 3	Wojew.	SK	-10,5	-7,4	-4,5	-2,9	-0,8	1,7	4,5	6,0	9,1
		SZ	-5,5	-3,8	-2,3	-1,5	-0,4	0,8	2,2	2,9	4,3
	Podregion	SK	-15,8	-9,5	-5,7	-1,5	0,0	2,3	5,6	9,8	17,8
		SZ	-8,6	-5,0	-2,9	-0,7	0,0	1,1	2,7	4,7	8,2
Tab. 4	Wojew.	SK	-11,8	-7,9	-6,0	-2,6	0,2	2,6	4,9	7,4	10,7
		SZ	-6,3	-4,1	-3,1	-1,3	0,1	1,3	2,4	3,6	5,1
	Podregion	SK	-15,6	-9,4	-5,8	-2,9	0,0	2,1	5,7	10,5	22,8
		SZ	-8,4	-4,9	-3,0	-1,5	0,0	1,0	2,8	5,0	10,2
Tab. 5	Wojew.	SK	-13,0	-7,3	-3,9	-1,3	0,7	4,0	6,7	9,5	20,6
		SZ	-6,9	-3,8	-2,0	-0,7	0,4	2,0	3,3	4,5	9,3
	Podregion	SK	-22,9	-14,1	-7,4	-3,0	0,0	3,0	7,4	15,1	30,3
		SZ	-12,9	-7,5	-3,8	-1,5	0,0	1,5	3,7	7,1	14,1
Tab. 6	Wojew.	SK	-14,6	-8,7	-6,2	-2,4	-0,2	1,8	4,0	7,2	11,6
		SZ	-7,9	-4,5	-3,2	-1,2	-0,1	0,9	1,9	3,5	5,5
	Podregion	SK	-18,6	-10,5	-6,7	-3,2	0,0	3,0	6,8	11,3	23,7
		SZ	-10,2	-5,5	-3,5	-1,7	0,0	1,5	3,3	5,3	10,6

Z powyższej tablicy wynika, że skrajne decyle (pierwszy i dziewiąty) są co do wartości bezwzględnej mniejsze dla współczynników zgodności pary estymatorów SK niż dla SZ.

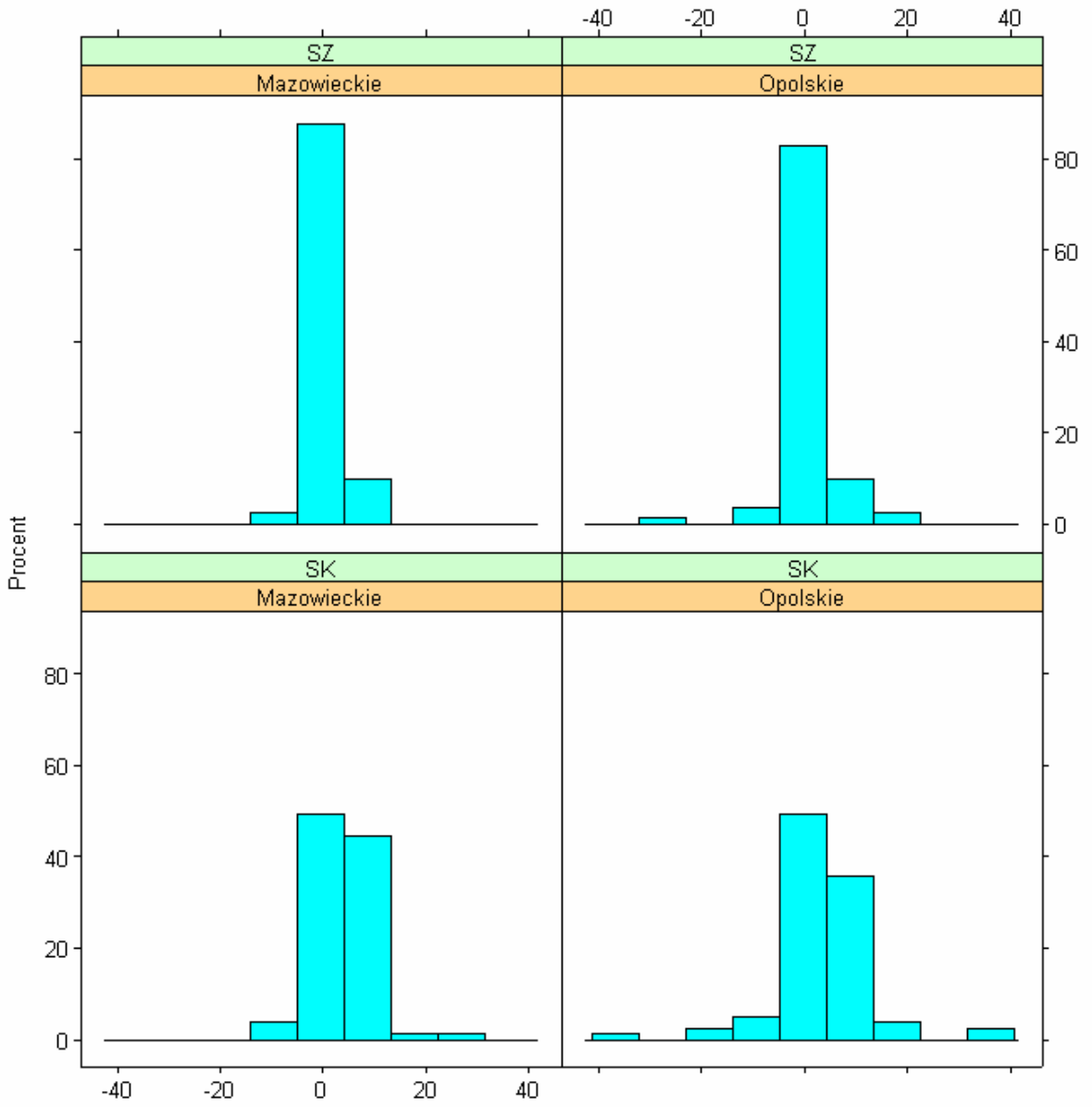
Wniosek ten wynika wprost z definicji analizowanych estymatorów. Ponadto wartość bezwzględna piątego decyła (mediany) we wszystkich przypadkach niewiele różni się od zera.

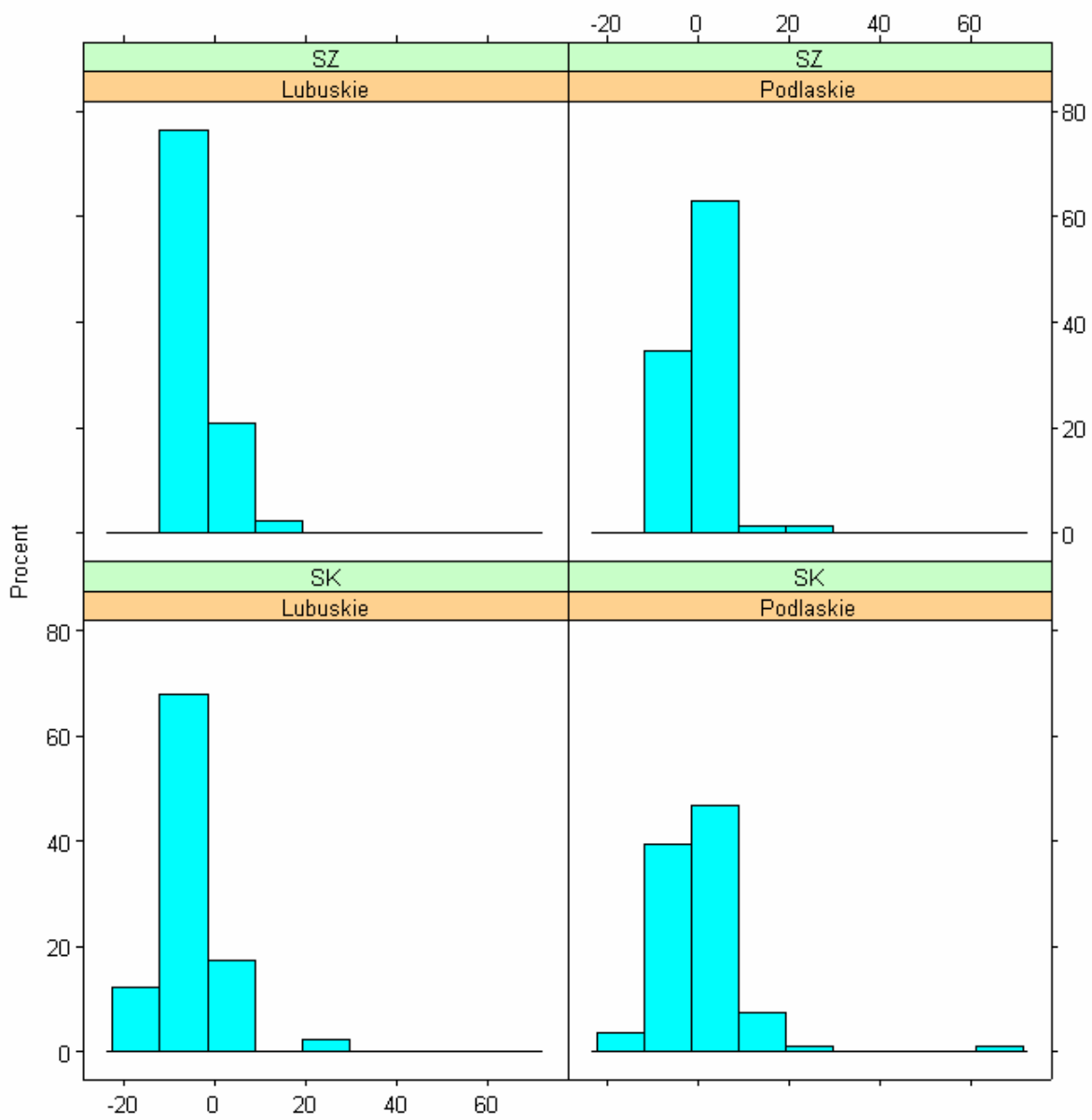
Tablica 4.2b. Rozkłady współczynników zgodności (rok 2002)

Wyszczególnienie			Przedziały klasowe											
			min	max	1	2	3	4	5	6	7	8	9	10
Tablica 1	Województwo	SK	-36,8	34,9	1	1	1	12	31	30	18	4	1	2
		SZ	-22,6	14,9	1	1	1	0	17	28	35	14	2	2
	Podregion	SK	-45,1	82,7	2	7	21	43	19	3	2	1	1	0
		SZ	-29,1	29,3	1	1	5	10	34	37	6	3	2	1
Tablica 2	Województwo	SK	-19,0	30,7	4	7	15	20	26	22	5	1	0	1
		SZ	-10,5	13,3	2	5	11	17	21	27	15	2	0	1
	Podregion	SK	-39,4	105,9	5	16	45	19	8	4	2	0	0	0
		SZ	-24,5	34,6	1	5	11	25	35	11	7	3	1	0
Tablica 3	Województwo	SK	-18,8	20,2	6	3	13	16	19	15	16	9	2	1
		SZ	-10,4	9,2	5	3	7	13	22	15	20	9	5	2
	Podregion	SK	-34,8	75,9	3	11	26	40	13	3	2	2	0	0
		SZ	-21,0	27,5	1	4	9	21	38	16	6	2	2	1
Tablica 4	Województwo	SK	-21,1	38,3	5	11	20	25	23	12	4	1	0	0
		SZ	-11,8	16,1	3	7	16	17	25	21	8	3	0	0
	Podregion	SK	-43,4	116,8	4	12	51	22	6	3	1	0	0	0
		SZ	-27,7	36,9	1	3	6	27	41	13	5	2	1	0
Tablica 5	Województwo	SK	-53,6	79,3	1	1	6	36	41	6	5	0	1	3
		SZ	-36,6	28,4	0	0	0	3	15	43	28	5	3	3
	Podregion	SK	-62,4	261,8	5	53	33	5	1	1	0	0	0	0
		SZ	-45,3	101,0	1	5	35	49	7	2	1	0	0	1
Tablica 6	Województwo	SK	-38,3	29,8	0	1	5	8	21	28	25	8	2	3
		SZ	-23,7	13,0	0	0	2	5	8	22	33	20	7	3
	Podregion	SK	-50,3	393,4	32	64	4	0	0	0	0	0	0	0
		SZ	-33,6	66,4	1	4	24	53	13	4	0	0	0	0

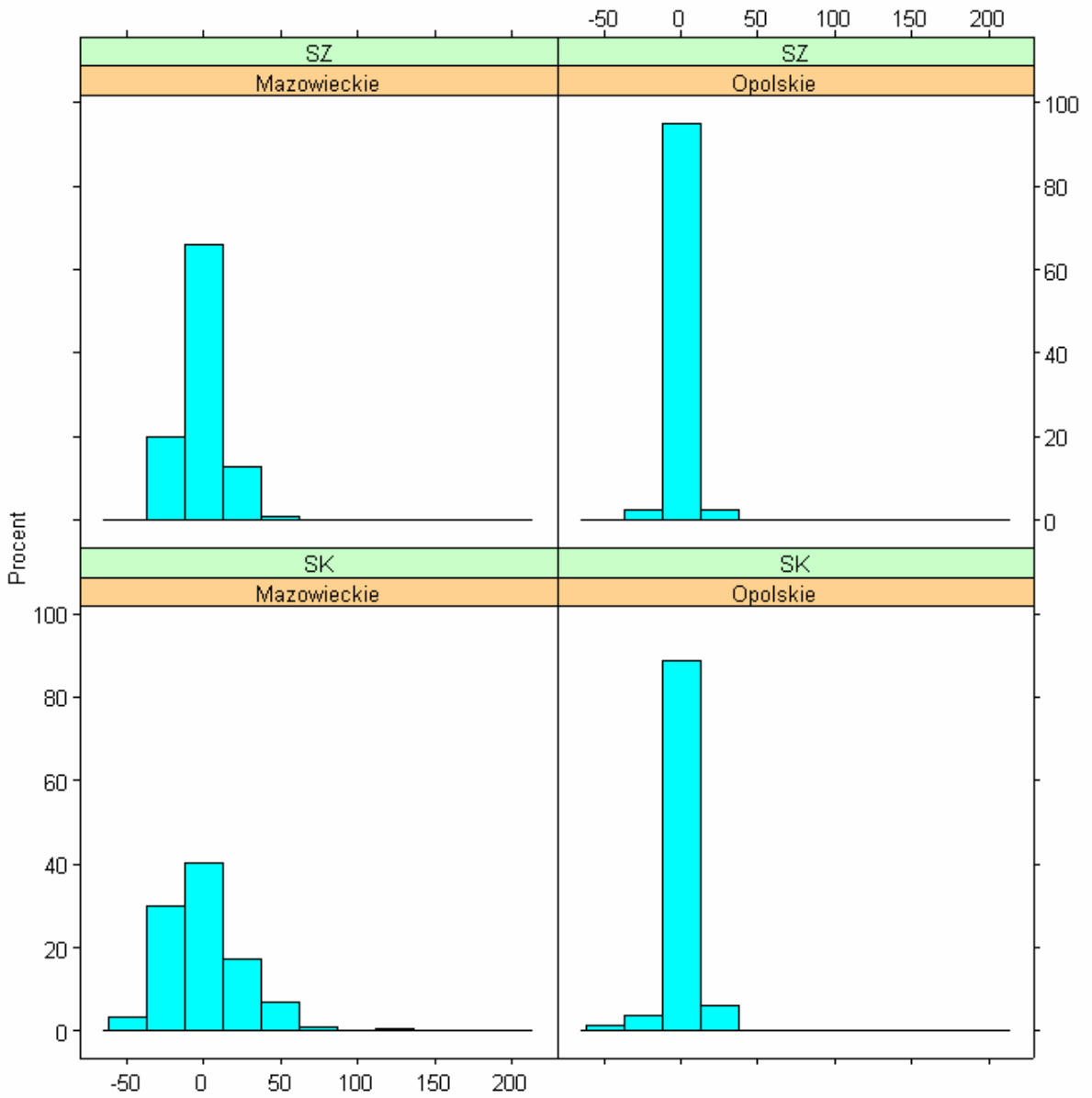
Częstości występowania wartości współczynników zgodności w środkowych przedziałach klasowych (4, 5 i 6) są na ogół najwyższe. Nieliczne wyjątki od powyższego stwierdzenia dotyczą pary estymatorów SK odnoszących się do podregionów i tablic wynikowych 5 i 6. Dla pary estymatorów SZ rozkłady współczynników zgodności dla województw są zgrupowane w pobliżu wartości zero.

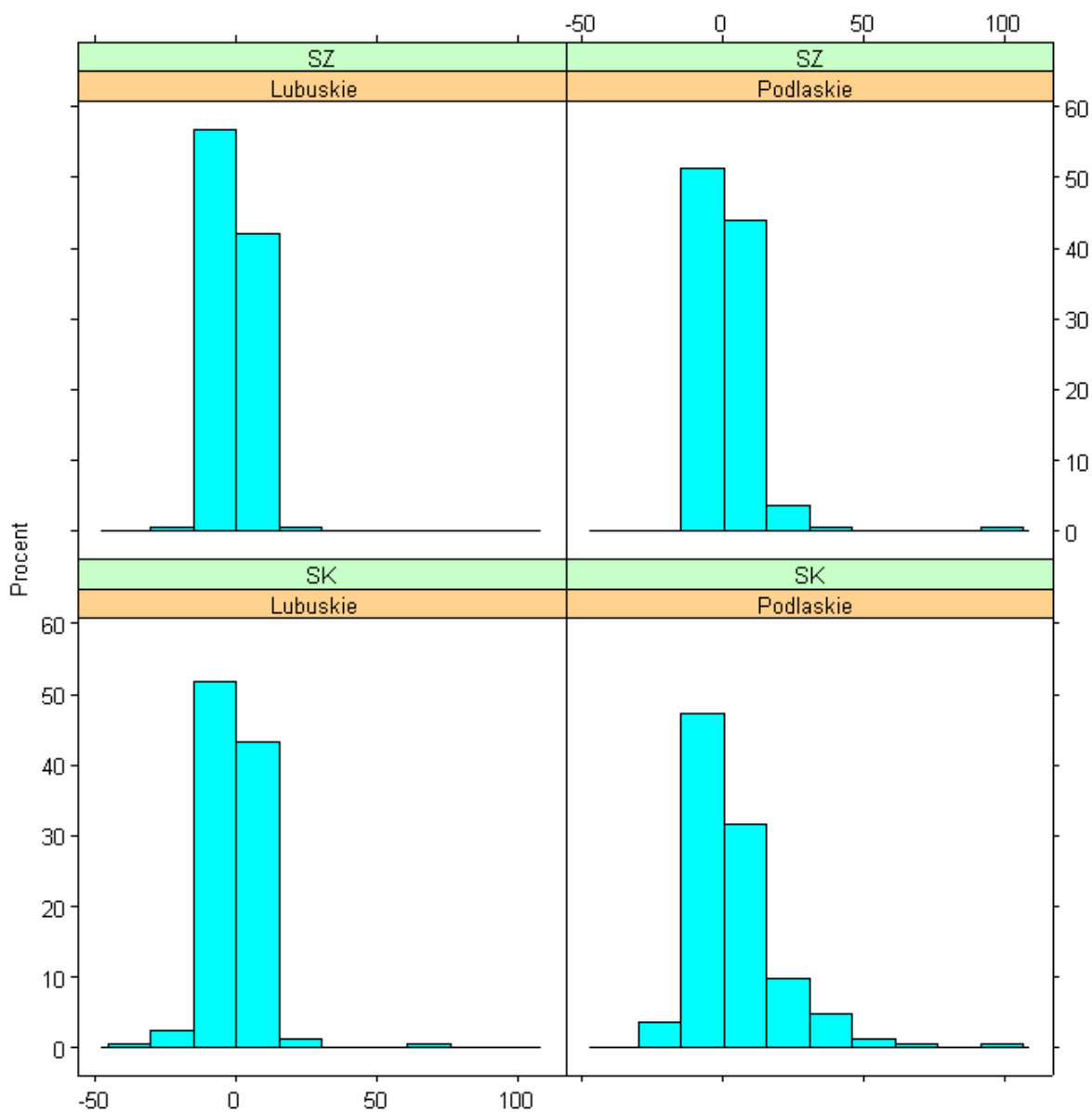
Rys. 3. Rozkłady współczynników zgodności dla wybranych województw w roku 2002.





Rys. 4. Rozkłady współczynników zgodności dla podregionów w wybranych województwach w roku 2002.





Powyższe histogramy potwierdzają wcześniej przedstawioną uwagę, że większość współczynników zgodności skupia się wokół zera. Lewą asymetrię daje się zauważyć dla województwa lubuskiego, zarówno dla estymatorów dla całego województwa jak i dla podregionów.

5. Uwagi końcowe

W badaniu aktywności ekonomicznej ludności prowadzonym w latach 1995 – 2002 przełomowym był rok 1999. W roku tym nastąpiły zmiany w podziale administracyjnym kraju oraz w sposobie doboru próby. Zmiany w schemacie losowania wynikły z wprowadzenia metody ciągłej w badaniu, a ponadto zastosowano inną alokację próby pomiędzy województwa. W miejsce alokacji proporcjonalnej wprowadzono podział próby między województwa zapewniający zbliżone liczebności prób. Zmiany te zostały szczegółowo opisane w pkt 2.4 i sprzyjały one estymacji parametrów dla województw, a nie tylko dla Polski. Jednak za zmianami w schemacie losowania nie wprowadzono zmian w metodzie estymacji. W szczególności nie wprowadzono zmian przy obliczaniu współczynników realizacji zachowując ich ogólnopolski charakter. Przy estymacji parametrów dla całego kraju postępowanie takie jest poprawne. W przypadku szacowania parametrów dla województw uzyskujemy zawyżone lub zaniżone oceny. Powodem tego jest m.in. zróżnicowana kompletność badania pomiędzy województwami w ramach danej klasy miejscowości. Np. w badaniu BAEL 2002 r. nastąpiło przeszacowanie liczby ludności w wieku 15 i więcej lat dla miasta Łodzi o 18.6% przy niedoszacowaniu tej zmiennej dla takich aglomeracji miejskich jak: Poznań (-12.4%), Trójmiasto (-11.8) oraz Wrocław (-11.3%). Anomalie takie nie dotyczą tylko dużych miast. Przykładowo dla podregionu konińskiego przeszacowanie wyniosło 28.1%, ostrołęcko-siedleckiego 24.4%, bielsko-bialskiego 23.5%. Z kolei niedoszacowanie dla podregionów: ełckiego (-29.4%), warszawskiego (-25.7) oraz wrocławskiego (-19.3%). Szczegóły zawiera tablica 4.1.

Innym powodem niedoszacowań bądź przeszacowań parametrów dla NTS2 i NTS3 może być stosowanie poststratyfikacji opartej na 48 ogólnopolskich grupach wieku (12 grup wieku w korelacji z płcią w podziale na miasto i wieś). Taki sposób wykorzystania danych demograficznych nie pozwala na uwzględnienie zróżnicowań regionalnych.

Mimo przedstawionych wyżej rozbieżności wynikających ze stosowanej metody estymacji, różnice w ocenach parametrów otrzymanych w kolejnych latach interesującego nas okresu a danymi z NSP 2002 r są generalnie niezbyt duże. Odzwierciedlają to wskaźniki podobieństwa rozkładów (por. J. Kordos (1973, str 116)). Wskaźniki te określone są wzorem

$$(5.1) \quad W_p = 1 - \frac{1}{2} \sum_{i=1}^k |g_i - f_i|$$

gdzie:

g_i — frakcja danej grupy ludności (np. bezrobotnych) w i -ym województwie (podregionie) wg BAEL 1995 (2002),

f_i — frakcja danej grupy ludności (np. bezrobotnych) w i -ym województwie (podregionie) wg NSP 2002,

$k=16$ dla województw oraz $k=45$ dla podregionów.

Tablica 5.1. Wskaźniki podobieństwa rozkładów z BAEL i NSP 2002 r. wg grup aktywności ekonomicznej ludności dla województw i podregionów

Aktywność ekonomiczna	Województwa		Podregiony	
	1995	2002	1995	2002
Ludność w wieku 15 lat i więcej	0.984	0.976	0.973	0.952
Pracujący	0.968	0.969	0.960	0.949
Bezrobotni	0.964	0.976	0.941	0.947
Bierni zawodowo	0.969	0.973	0.957	0.951

Podane w powyższej tablicy wskaźniki podobieństwa są bliskie jedności, co wskazuje na wysoki poziom zgodności rozważanych rozkładów. Oznacza to, że mimo zmian absolutnych liczb określających aktywność ekonomiczną ludności (wzrost lub spadek) struktura geograficzna (województwa i podregiony) ulega niewielkim zmianom. Oznacza to, że w poszczególnych jednostkach terytorialnych mamy podobne trendy w zakresie aktywności ekonomicznej. Na tej podstawie można stwierdzić, że dane z NSP 2002 r mogą być wykorzystywane do dezagregacji wyników BAEL w przyszłości. Horyzont czasowy wykorzystania danych spisowych w BAEL jest trudny do sprecyzowania. Wydaje się jednak, że możliwe będzie korzystanie z tych danych tak długo jak długo nie nastąpią radykalne zmiany w strukturze geograficznej aktywności ekonomicznej ludności związane z migracją ludności lub dużymi inwestycjami.

W świetle dotychczasowych rozważań wydaje się konieczne podjęcie prac studialnych nad metodami estymacji pod kątem wykorzystania danych zbieranych w badaniu aktywności ekonomicznej ludności do uzyskiwania ocen parametrów dla małych obszarów (województw, podregionów, powiatów). Przede wszystkim wydaje się konieczne stosowanie równoległe dwóch metod estymacji. Jedną z nich wykorzystywano by celem uzyskiwania ocen paramet-

trów dla Polski, drugą zaś dla województw (i być może podregionów). Poniżej przedstawiamy propozycje zmian estymatorów.

Otrzymanie precyzyjnych danych dla województw i podregionów będzie wymagało innej niż dotychczas metody estymacji. Po pierwsze, inaczej niż dotychczas trzeba będzie obliczać współczynniki realizacji (zindywidualizować je dla poszczególnych województw), a po drugie inaczej dokonywać poststratyfikacji. Zasady wykorzystania danych demograficznych trzeba będzie dostosować do specyfiki danego województwa. Wydaje się, że liczba grup wieku ludności dla województw będzie mniejsza niż stosowana dla Polski ogółem. Niezbędne będzie pozyskiwanie w każdym kwartale danych demograficznych o ludności w gospodarstwach domowych (bez gospodarstw zbiorowych) wg województw, podregionów i klas miejscowości. Oceny parametrów dla województw będą podstawą uzyskiwania szacunków dla podregionów i powiatów.

Zastosowanie proponowanych rozwiązań doprowadzi do tego, że suma ocen parametrów dla województw może nie być równa ocenom parametrów dla Polski tj. dane dla Polski nie będą prostą sumą danych wojewódzkich. Sumowanie ocen wojewódzkich mogłoby prowadzić do zmniejszenia precyzji ocen ogólnopolskich w grupach wieku i płci.

Uważamy ponadto, że należy doskonalić proponowane w opracowaniu metody estymacji złożonej dla podregionów i dużych powiatów poprzez szukanie optymalnych wag (we wzorach na estymatory złożone). Drugim problemem, który naszym zdaniem powinien być przedmiotem dalszych rozważań jest rozpatrzenie szerokiej klasy estymatorów służących do dezagregacji danych wojewódzkich z BAEL między podregiony i powiaty. Przy estymacji parametrów dla Polski proponujemy natomiast zastosowanie estymatorów regresyjnych wykorzystujących dużą liczbę zmiennych dodatkowych. Należy rozważyć również inne, niż dotychczas stosowane metody ważenia, a w szczególności takie, które dostarczają wag wspólnych dla osób, gospodarstw domowych i mieszkań.

6. Bibliografia

Cz. Bracha (1993): Properties of total value estimators for domains in two-stage sample surveys. Artykuł w publikacji "Small Area statistics and survey designs" vol. II wydanej przez GUS, Warszawa str.169—183.

Cz. Bracha (1994): Metodologiczne aspekty badania małych obszarów. Z prac ZBSE z. 43, ZBSE GUS i PAN, Warszawa.

Cz. Bracha (1996): Teoretyczne podstawy metody reprezentacyjnej. Wydawnictwo Naukowe PWN, Warszawa.

Cz. Bracha (1998): Metoda reprezentacyjna w badaniu opinii publicznej i marketingu. Wydawnictwo „EFEKT”, Warszawa.

W.G. Cochran (1977): Sampling techniques (wyd. III). JW&S, New York—London.

B. Efron (1979): Bootstrap methods: Another look at the jackknife, *Annals of Statistics* 7, str. 1—26.

M.R. Frankel (1971): Inference from survey sampling: an empirical investigation. University of Michigan. Ann Arbor, Michigan.

W.A. Fuller i R.M. Harter (1987): The multivariate components of variance model for small area estimation. Artykuł w publikacji "Small Area Statistics" wydanej przez R. Platka, J.N.K. Rao, C.E. Särndala i M.P. Singha, JW&S, New York, str. 103—123.

R. Griffiths. (1996): Current population survey small area estimation for congressional districts. Proceeding of the Section On Survey Research Method. American Statistical Association, STR 314 – 319.

J. Jakubowski i Cz. Bracha (2001): Przybliżone szacowanie wariancji w przypadku złożonych schematów losowania. *Studia i Prace z Prac ZBSE, ZBSE GUS i PAN*, Warszawa 2001, zeszyt 273, str. 1—99.

J. Jakubowski i Cz. Bracha (2001): Influence of numbers of grouped balanced half-samples on effectiveness of variance estimation for complex sample surveys. *Statistical in Transition* 5.3, str. 383—404.

J. Kordos (1973): Metody analizy i prognozowania rozkładów płac i dochodów ludności. PWE, Warszawa.

J. Kordos (1999): Problemy estymacji danych dla małych obszarów. *Wiadomości Statystyczne* 44.1, str. 85—101.

- J. Kordos (2000): Teoria i sztuka badań reprezentacyjnych. *Wiadomości Statystyczne* 45.1, str. 1—13.
- J. Kordos (2002): Niektóre aspekty jakości w statystyce małych obszarów. *Wiadomości Statystyczne*, nr 47.11, str. 14 – 28.
- J. Kordos, B. Lednicki, M. Żyra (2002): The household of sample survey in Poland. *Statistics in Transition. Journal of the Polish Statistical Association*. 5.4, str. 555—589.
- J. Kordos i J. Paradysz (2000): Prace badawcze nad zastosowaniem metod estymacji dla małych obszarów w Polsce. *Wiadomości Statystyczne* 45.11, str. 1—22.
- J. Kordos i J. Paradysz (2000): Some experiments in small area estimation in Poland. *Statistics in Transition* 4.4., str. 679—697.
- B. Lednicki i J. Wesołowski (1994): Lokalizacja próby pomiędzy subpopulacje. *Wiadomości Statystyczne* 39.9, str. 2—4.
- P.J. McCarthy (1969): Pseudo-replication: half samples. *Intern. Statist. Review* 37.3, str. 239—364.
- P.J. McCarthy i C.B. Snowden (1985): The bootstrap and finite population sampling. *Vital and Health Statistics*, str. 2—95, Public Health Service Publication 85—1369, U.S. Government Printing Office, Washington, DC.
- U. Norlen i T. Waller (1979): Estimation in a complex - experiences from a survey of buildings with regard to energy usage. *Statistisk Tidskrift* 17.2, str. 111—124.
- R.L. Plackett i P.J. Burman (1946): The design of optimum multifactorial experiments. *Biometrika* 33, str. 305—325.
- R. Platek, J.N.K. Rao, C.E. Särndal i M.P. Singh (wydawcy) (1987): *Small Area Statistics. An International Symposium*. JW&S, New York.
- N.G.N. Prasad i J.N.K. Rao (1990): The estimation of the mean squared error of small-area estimators. *JASA* 85, str. 163—171.
- N.J. Purcell i L. Kish (1979): Estimation for small domains. *Biometrics* 35, str. 365—384.
- N.J. Purcell i L. Kish (1990): Postcensal estimates for local areas (or domains). *Intern. Statist. Review* 48.1, str. 3—18.
- J.N.K. RAO, (2003) *Small Area Estimation*. Willey-Interscience. JW&S, Inc., Publication.
- J.N.K. Rao i J. Shao (1996): On balanced half-sample variance estimation in stratified random sampling. *JASA* 91, str. 343—348.

J.N.K. Rao i J. Shao (1999): Modified balanced repeated replication for complex survey data. *Biometrika* 88.2, str. 403—415.

J.N.K. Rao i C.F.J. Wu (1988): Resampling inference with complex survey data. *Journal of the American Statistical Association*, 83, str. 231—241.

C.E.Särndal, B. Swensson i J. Wretman (1992): *Model Assisted Survey Sampling*. Springer Verlag.

J. Shao i D. Tu (1995): *The jackknife and bootstrap*. New York, Springer-Verlag.

P.J. Szablowski, J. Wesołowski i R. Wieczorkowski (1996): Estymacja w podpopulacjach. *Wiadomości Statystyczne* 41.7, str. 1—13.

A. Szarkowski (1981): Problemy doboru próby do badania warunków bytu. Artykuł w publikacji "Metoda reprezentacyjna w masowych badaniach statystycznych: teoria i praktyka" wydanej w serii "Z Prac ZBSE" z. 122, ZBSE GUS, Warszawa 46—89.

A. Szarkowski (2002): Aktywność ekonomiczna ludności Polski. II kwartał 2002. Uwagi metodologiczne, str. XI—XXII. GUS, Warszawa.

J. Witkowski (2002): Badanie aktywności ekonomicznej ludności w NSP 21002. *Wiadomości Statystyczne* 47.4, str. 56—64.

K.M. Wolter (1985): *Introduction to Variance Estimation*. Springer-Verlag, New York, Berlin, Heidelberg, Tokyo.

M. Żyra (1997): Uczestnictwo gospodarstw domowych w badaniu aktywności ekonomicznej ludności. *Wiadomości Statystyczne* 42.11, str. 15—22.

7. Aneks

7.1. Opis tablic

Zamieszczona w aneksie tablica 0 stanowi zbiorcze zestawienie ocen uzyskanych za pomocą estymatorów klasycznego (K) i złożonego (Z) dla województw i podregionów w latach 1995—2002.

Tablice 1—6 dotyczą roku 2002. Dla każdego parametru w tych tablicach podane zostały trzy oceny otrzymane za pomocą trzech wcześniej rozważanych metod estymacji: klasycznej tj takiej jak w BAEL (K), syntetycznej (S) i złożonej (Z). Pierwsze trzy tablice dotyczą województw i podregionów z jednej strony oraz płci i wieku z drugiej. Następne trzy tablice, również dla województw i podregionów, zawierają oceny kolejno dla pracujących, bezrobotnych i biernych zawodowo, w korelacji z płcią i wykształceniem. Trzy ostatnie tablice odnoszą się do ocen parametrów dla powiatów. Tablice te zawierają wyłącznie oceny uzyskane za pomocą estymatorów syntetycznych. Tablica 7 dotyczy aktywności ekonomicznej ludności w wieku 15 lat i więcej wg płci i miejsca zamieszkania. Ponadto zawierają oceny wskaźników aktywności zawodowej i zatrudnienia oraz stopy bezrobocia. Tablica 8 podaje oceny parametrów wg statusu zatrudnienia, płci i miejsca zamieszkania. Ostatnia dziewiąta tablica prezentuje oceny parametrów dotyczące pracujących według sektorów ekonomicznych, płci i miejsca zamieszkania.

Dane każdej z tablic 1—9 dotyczą jednego roku z okresu 1995—2002. Wszystkie tablice stanowią oddzielne pliki arkusza kalkulacyjnego EXCEL. Pliki z tablicami od 1 do 6 zgrupowane zostały w folderze **TABLICE_W**. Folder **TABLICE_P** zawiera pliki z tablicami 7—9. Folder **TABLICE_W** obejmuje tablice z ocenami parametrów dla województw i podregionów, a folder **TABLICE_P** dotyczy powiatów. W nazwie pliku zawarty jest numer tablicy oraz rok, którego ona dotyczy.

W tablicach 1—6 pod ocenami parametrów zostały podane wartości współczynników zmienności (CV) wyrażone w %. Współczynniki te są miarą precyzji estymatorów. Ponadto, w folderze **TABLICE_W** zamieszczone zostały pliki z tablicami, w których zapisane zostały (w %) współczynniki zgodności pomiędzy poszczególnymi estymatorami. Każdy z tych plików, obok numeru tablicy z ocenami parametrów, której dotyczą te współczynniki, ma w nazwie symbol SZ albo symbol SK. W pierwszym przypadku tablica zawiera współczynniki zgodności ocen uzyskanych przy wykorzystaniu estymatorów syntetycznych i złożonych, zaś

w drugim, w tablicy zapisane zostały współczynniki zgodności ocen pochodzących z estymacji syntetycznej i klasycznej. Każda z tych tablic zawiera współczynniki zgodności odnoszące się do ocen parametrów dla całego rozpatrywanego okresu 1995—2002.

W tablicach 7—9 zostały podane oceny parametrów otrzymane metodą syntetyczną dla wszystkich powiatów, zaś dla powiatów większych (o liczbie jps powyżej 50 w badaniach w roku 2002) również metodą klasyczną i złożoną. Każda z tablic 7—9, zawierająca oceny parametrów dla danego roku, ma swój odpowiednik w postaci tablicy ze współczynnikami CV. W nazwie pliku, poza numerem tablicy i rokiem, którego dotyczy, dodano skrót CV. W folderze TABLICE_P zamieszczone zostały także tablice zawierające współczynniki zgodności pomiędzy ocenami parametrów uzyskanych przy wykorzystaniu różnych estymatorów. Współczynniki te dotyczą tylko tych powiatów, dla których szacowane były parametry różnymi metodami. W nazwie pliku, oprócz numeru tablicy z ocenami parametrów, której dotyczą zapisane współczynniki zgodności, mamy symbol SZ albo KSZ. W pierwszym przypadku mamy do czynienia z tablicami o analogicznej konstrukcji jak tablice z symbolem SZ w folderze TABLICE_W. W tablicach oznaczonych symbolem KSZ zapisane zostały współczynniki zgodności ocen parametrów pochodzących z estymacji klasycznej i syntetycznej (wiersze oznaczone symbolem S) oraz współczynniki zgodności ocen uzyskanych przez zastosowanie estymatorów klasycznych i złożonych (wiersze o symbolu Z). Także i w tym przypadku każda z tablic dotyczy okresu 1995—2002.

7.2. Płyta CD z danymi dla lat 1995—2002