

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

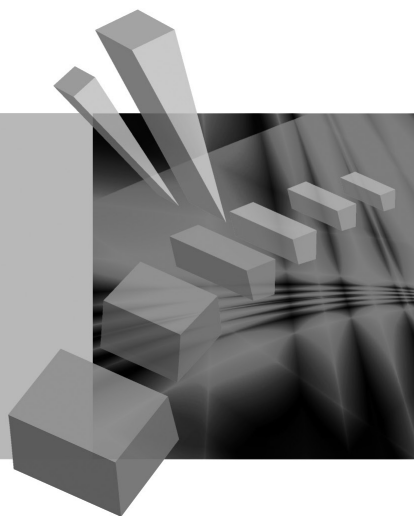
RESEARCH PAPERS

of Wrocław University of Economics

278

Taksonomia 20

Klasyfikacja i analiza danych – teoria i zastosowania



Redaktorzy naukowi

Krzysztof Jajuga

Marek Walesiak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2013

Redaktor Wydawnictwa: Aleksandra Śliwka

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Małgorzata Czupryńska

Projekt okładki: Beata Dębska

Publikacja jest dostępna w Internecie na stronach:

www.ibuk.pl, www.ebscohost.com,

The Central and Eastern European Online Library www.ceeol.com,

a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon

http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php

Informacje o naborze artykułów i zasadach recenzowania znajdują się

na stronie internetowej Wydawnictwa

www.wydawnictwo.ue.wroc.pl

Tytuł dofinansowany ze środków Narodowego Banku Polskiego

oraz ze środków Sekcji Klasyfikacji i Analizy danych PTS

Kopiowanie i powielanie w jakiegokolwiek formie

wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu

Wrocław 2013

ISSN 1899-3192 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)

ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

Wstęp	9
Józef Pocięcha: Wskaźniki finansowe a klasyfikacyjne modele predykcji upadłości firm	15
Eugeniusz Gatnar: Analiza miar adekwatności rezerw walutowych	23
Marek Walesiak: Zagadnienie doboru liczby klas w klasyfikacji spektralnej	33
Joanicjusz Nazarko, Joanna Ejdyś, Anna Kononiuk, Anna M. Olszewska: Analiza strukturalna jako metoda klasyfikacji danych w badaniach foresight	44
Andrzej Bąk: Metody porządkowania liniowego w polskiej taksonomii – pakiet <code>pllord</code>	54
Aleksandra Łuczak, Feliks Wysocki: Zastosowanie mediany przestrzennej Webera i metody TOPSIS w ujęciu pozycyjnym do konstrukcji syntetycznego miernika poziomu życia	63
Ewa Roszkowska: Zastosowanie rozmytej metody TOPSIS do oceny ofert negocjacyjnych	74
Jacek Batóg: Analiza wrażliwości metody ELECTRE III na obserwacje nietypowe i zmianę wartości progowych	85
Jerzy Korzeniewski: Modyfikacja metody HINoV selekcji zmiennych w analizie skupień	93
Małgorzata Markowska, Danuta Strahl: Wykorzystanie referencyjnego systemu granicznego do klasyfikacji europejskiej przestrzeni regionalnej ze względu na filar inteligentnego rozwoju – kreatywne regiony	101
Elżbieta Sobczak: Inteligentne struktury pracujących a efekty strukturalne zmian zatrudnienia w państwach Unii Europejskiej.....	111
Elżbieta Gołata, Grażyna Dehnel: Rozbieżności szacunków NSP 2011 i BAEL.....	120
Iwona Foryś: Wykorzystanie analizy historii zdarzeń do badania powtórnego sprzedaży na lokalnym rynku mieszkaniowym	131
Hanna Dudek, Joanna Landmesser: Wpływ relatywnej deprivacji na subiektywne postrzeganie dochodów.....	142
Grażyna Łaska: Syntaksonomia numeryczna w klasyfikacji, identyfikacji i analizie przemian zbiorowisk roślinnych	151
Magdalena Osińska, Marcin Faldziński, Tomasz Zdanowicz: Analiza zależności między procesami fundamentalnymi a rynkiem kapitałowym w Chinach	161

Andrzej Bąk, Tomasz Bartłomowicz: Mikroekonometryczne modele wielomianowe i ich zastosowanie w analizie preferencji z wykorzystaniem programu R	169
Andrzej Dudek, Bartosz Kwaśniewski: Przetwarzanie równoległe algorytmów analizy skupień w technologii CUDA	180
Michał Trzęsiok: Wycena rynkowej wartości nieruchomości z wykorzystaniem wybranych metod wielowymiarowej analizy statystycznej	188
Joanna Trzęsiok: Wybrane symulacyjne techniki porównywania nieparametrycznych metod regresji.....	197
Artur Mikulec: Kryterium Mojeny i Wisharta w analizie skupień – przypadek skupień o różnych macierzach kowariancji	206
Artur Zaborski: Analiza <i>unfolding</i> z wykorzystaniem modelu grawitacji	216
Justyna Wilk: Identyfikacja obszarów problemowych i wzrostowych w województwie dolnośląskim w zakresie kapitału ludzkiego	225
Karolina Bartos: Analiza ryzyka odejścia studenta z uczelni po uzyskaniu dyplomu licencjata – zastosowanie sieci MLP	236
Ewa Genge: Segmentacja uczestników Industriady z wykorzystaniem analizy klas ukrytych	246
Izabela Kurzawa: Wielomianowy model logitowy jako narzędzie identyfikacji czynników wpływających na sytuację mieszkaniową polskich gospodarstw domowych	254
Marek Lubicz, Maciej Zięba, Konrad Pawelczyk, Adam Rzechonek, Jerzy Kołodziej: Modele eksploracji danych niezbilansowanych – procedury klasyfikacji dla zadania analizy ryzyka operacyjnego.....	262
Aleksandra Łuczak: Zastosowanie rozmytej hierarchicznej analizy w tworzeniu strategii rozwoju jednostek administracyjnych	271
Marcin Pelka: Rozmyta klasyfikacja spektralna <i>c</i> -średnich dla danych symbolicznych interwałowych.....	282
Małgorzata Machowska-Szewczyk: Klasyfikacja obiektów reprezentowanych przez różnego rodzaju cechy symboliczne	290
Ewa Chodakowska: Indeks Malmquista w klasyfikacji podmiotów gospodarczych według zmian ich względnej produktywności działania	300
Beata Bieszk-Stolorz, Iwona Markowicz: Wykorzystanie modeli proporcjonalnego i nieproporcjonalnego hazardu Coxa do badania szansy podjęcia pracy w zależności od rodzaju bezrobocia	311
Marcin Salamaga: Weryfikacja teorii poziomu rozwoju gospodarczego J.H. Dunninga w ujęciu sektorowym w wybranych krajach Unii Europejskiej	321
Justyna Wilk, Michał Bernard Pietrzak, Stanisław Matusik: Sytuacja społeczno-gospodarcza jako determinanta migracji wewnętrznych w Polsce.	330
Hanna Gruchociak: Delimitacja lokalnych rynków pracy w Polsce na podstawie danych z badania przepływów ludności związanych z zatrudnieniem	343

Radosław Pietrzyk: Efektywność inwestycji polskich funduszy inwestycyjnych z tytułu doboru papierów wartościowych i umiejętności wykorzystania trendów rynkowych	351
Sabina Denkowska: Procedury testowań wielokrotnych	362

Summaries

Józef Pocięcha: Financial ratios and classification models of bankruptcy prediction	22
Eugeniusz Gatnar: Analysis of FX reserve adequacy measures	32
Marek Walesiak: Automatic determination of the number of clusters using spectral clustering	43
Joanicjusz Nazarko, Joanna Ejdys, Anna Kononiuk, Anna M. Olszewska: Structural analysis as a method of data classification in foresight research	53
Andrzej Bąk: Linear ordering methods in Polish taxonomy – pllord package	62
Aleksandra Łuczak, Feliks Wysocki: The application of spatial median of Weber and the method TOPSIS in positional formulation for the construction of synthetic measure of standard of living	73
Ewa Roszkowska: Application of the fuzzy TOPSIS method to the estimation of negotiation offers.....	84
Jacek Batóg: Sensitivity analysis of ELECTRE III method for outliers and change of thresholds	92
Jerzy Korzeniewski: Modification of the HINoV method of selecting variables in cluster analysis	100
Małgorzata Markowska, Danuta Strahl: Implementation of reference limit system for the European regional space classification regarding smart growth pillar – creative regions	110
Elżbieta Sobczak: Smart workforce structures versus structural effects of employment changes in the European Union countries	119
Elżbieta Gołata, Grażyna Dehnel: Divergence in National Census 2011 and LFS estimates.....	130
Iwona Foryś: Event history analysis in the resale study on the local housing market	141
Hanna Dudek, Joanna Landmesser: Impact of the relative deprivation on subjective income satisfaction	150
Grażyna Łaska: Numerical syntaxonomy in classification, identification and analysis of changes of secondary communities	160
Magdalena Osińska, Marcin Faldziński, Tomasz Zdanowicz: Analysis of relations between fundamental processes and capital market in China.....	166
Andrzej Bąk, Tomasz Bartłomowicz: Microeconomic polynomial models and their application in the analysis of preferences using R program.....	179

Andrzej Dudek, Bartosz Kwaśniewski: Parallel processing of clustering algorithms in CUDA technology	187
Michał Trzęsiok: Real estate market value estimation based on multivariate statistical analysis	196
Joanna Trzęsiok: On some simulative procedures for comparing nonparametric methods of regression.....	205
Artur Mikulec: Mojena and Wishart criterion in cluster analysis – the case of clusters with different covariance matrices	215
Artur Zaborski: Unfolding analysis by using gravity model	224
Justyna Wilk: Determination of problem and growth areas in Dolnośląskie Voivodship as regards human capital.....	235
Karolina Bartos: Risk analysis of bachelor students' university abandonment – the use of MLP networks	245
Ewa Genge: Clustering of industrial holiday participants with the use of latent class analysis.....	253
Izabela Kurzawa: Multinomial logit model as a tool to identify the factors affecting the housing situation of Polish households.....	261
Marek Lubicz, Maciej Zięba, Konrad Pawelczyk, Adam Rzechonek, Jerzy Kołodziej: Modelling class imbalance problems: comparing classification approaches for surgical risk analysis	270
Aleksandra Łuczak: The application of fuzzy hierarchical analysis to the evaluation of validity of strategic factors in administrative districts.....	281
Marcin Pełka: A spectral fuzzy c-means clustering algorithm for interval-valued symbolic data	289
Małgorzata Machowska-Szewczyk: Clustering algorithms for mixed-feature symbolic objects	299
Ewa Chodakowska: Malmquist index in enterprises classification on the basis of relative productivity changes	310
Beata Bieszk-Stolorz, Iwona Markowicz: Using proportional and non proportional Cox hazard models to research the chances for taking up a job according to the type of unemployment	320
Marcin Salamaga: Verification J.H. Dunning's theory of economic development by economic sectors in some EU countries	329
Justyna Wilk, Michał Bernard Pietrzak, Stanisław Matusik: Socio-economic situation as a determinant of internal migration in Poland	342
Hanna Gruchociak: Delimitation of local labor markets in Poland on the basis of the employment-related population flows research.....	350
Radosław Pietrzyk: Selectivity and timing in Polish mutual funds performance measurement	361
Sabina Denkowska: Multiple testing procedures.....	369

Elżbieta Gołata, Grażyna Dehnel

Uniwersytet Ekonomiczny w Poznaniu

ROZBIEŻNOŚCI SZACUNKÓW NSP 2011 I BAEL

Streszczenie: Celem artykułu była próba określenia przyczyn rozbieżności dotyczących szacunków prezentowanych na podstawie wyników NSP 2011 i Badania Aktywności Ekonomicznej Ludności w odniesieniu do rynku pracy. Omówiona została również kwestia wpływu i znaczenia zaobserwowanych rozbieżności pomiędzy wynikami badań. Zaprezentowano analizę porównawczą bieżącej oceny sytuacji na rynku pracy, wykorzystując do tego celu dane pochodzące ze spisu ludności 2011 oraz wyniki BAEL.

Słowa kluczowe: spis ludności oparty na rejestrach, BAEL, ocena jakości danych.

1. Wstęp

Komunikat prezesa GUS o aktywności ekonomicznej ludności jest jedną z najbardziej oczekiwanych informacji Głównego Urzędu Statystycznego. Według określonego harmonogramu, w każdym miesiącu publikowane są między innymi informacje o bezrobociu rejestrowanym w Polsce, przeciętnym miesięcznym wynagrodzeniu w sektorze przedsiębiorstw, co kwartał o rynku pracy.

Informacje te oczekiwane są z wielu względów. Jednym z kluczowych jest konieczność podejmowania różnego rodzaju strategicznych dla państwa decyzji. Wśród nich podkreśla się zarówno te, które dotyczą kwestii społecznych, określenia wysokości zasiłków dla bezrobotnych, pomocy socjalnej, ich zróżnicowania terytorialnego oraz tak ważnych dla całej gospodarki jak zmiany stóp procentowych. Stąd zrozumiałe było poruszenie wywołane informacją o rynku pracy opracowaną na podstawie wyników ostatniego spisu ludności [GUS 2012d], która okazała się rozbieżna z wynikami opublikowanymi na podstawie Badania Aktywności Ekonomicznej Ludności (BAEL).

Celem artykułu jest próba określenia przyczyn rozbieżności dotyczących prezentowanych na podstawie wyników NSP 2011 i BAEL szacunków oraz wskazanie ich znaczenia. Zaprezentowana zostanie również analiza porównawcza bieżącej oceny sytuacji na rynku pracy na podstawie danych spisu ludności oraz wyników BAEL.

2. Podobieństwa i różnice metodologii NSP 2011 i BAEL w świetle badanej ludności

GUS w *Raporcie z wyników Narodowego Spisu Powszechnego Ludności i Mieszkań 2011* [GUS 2012b, s. 49] poinformował, iż pomimo zastosowania nowej metodologii wyniki NSP 2011 charakteryzowały się dużą zgodnością w porównaniu do bieżących bilansów wyprowadzonych na koniec 2010 r. Przy porównaniu odniesiono się do ludności Polski. Wyniki NSP 2011 wskazały na liczbę ludności Polski większą o 312 tys., tj. o 0,8% ludności spisanej w porównaniu do bilansu na 31 grudnia 2010 r. Zakres rozbieżności powiększa się jednak przy przejściu na niższy poziom agregacji, na przykład przekrój terytorialny lub w ramach kategorii ludności.

Zauważono jednak, że w przypadku innej kategorii, jaką jest liczba ludności aktywnej ekonomicznie, pojawiły się rozbieżności. Według danych BAEL, liczba ludności aktywnej ekonomicznie była o 1 125 tys. osób¹ wyższa, aniżeli wynikało z danych NSP 2011. Zgodnie ze spisem ludność aktywna ekonomicznie wynosiła 16 521 tys. osób, podczas gdy wcześniej opublikowane wyniki BAEL podawały nie tylko wyższą liczbę sięgającą 17 646 tys. osób, ale także odpowiednio wyższą liczbę pracujących. Różnica w szacunku liczby pracujących wynosiła 1 499 tys. osób (NSP 2011 – 14 376 tys. oraz BAEL I kwartał 2011 – 15 875 tys.). Podobne rozbieżności dotyczą między innymi liczby osób bezrobotnych (NSP 2011 – 2 146 tys. oraz BAEL – 1 771 tys.), wskaźnika zatrudnienia, czy stopy bezrobocia.

W maju 2012 r. GUS opublikował komentarz dotyczący „Różnic wyników BAEL i NSP 2011”, w którym wyjaśnił przyczyny rozbieżności w publikowanych informacjach. Można je podzielić na trzy kategorie.

1. Różnice definicyjne w określeniu populacji objętej badaniem

W NSP pytania dotyczące aktywności ekonomicznej zadano tylko osobom zamieszkałym w mieszkaniach (por. [GUS 2012a]). Podobnie w BAEL obserwacja przeprowadzana jest poprzez gospodarstwa domowe. Badaniem objęte są wszystkie osoby w wieku 15 lat i więcej, zamieszkałe w wylosowanych mieszkaniach. Z określenia poddanej badaniu populacji wynika, iż w obu badaniach nie ma możliwości określenia statusu na rynku pracy osób w gospodarstwach zbiorowego zamieszkania oraz bezdomnych.

Od I kwartału 2004 r. BAEL, w przeciwieństwie do NSP 2011, obejmuje również cudzoziemców będących członkami gospodarstw domowych w wylosowanych mieszkaniach, jeżeli są obecni w badanym mieszkaniu przez co najmniej 3 miesiące. Również odmiennie aniżeli w NSP 2011 BAEL nie obejmuje osób będących członkami gospodarstw domowych, jeżeli przebywały za granicą przez co najmniej 3 miesiące².

¹ Wstępne wyniki NSP 2011 opublikowane 23 marca 2011 r. wskazywały, że szacunek aktywnych zawodowo według BAEL był wyższy o 1 397 tys. osób (por. [GUS 2012a]).

² Spis ludności obejmował wszystkie osoby stale zamieszkałe na obszarze Polski bez względu na fakt, czy przebywały w kraju w czasie spisu czy też były za granicą, oraz osoby przebywające czasowo.

2. Różnice wynikające z przyjętych definicji dotyczących statusu na rynku pracy. Zarówno w BAEL, jak i w NSP stosowane są te same definicje ludności aktywnej zawodowo, pracujących i bezrobotnych. Są one zgodne z definicjami przyjętymi na XIII Międzynarodowej Konferencji Statystyków Pracy w październiku 1982 r. i zalecanymi do stosowania przez Międzynarodową Organizację Pracy.

3. Różnice wynikające z przyjętych rozwiązań o charakterze technicznym

W definicji osoby pracującej w BAEL uwzględniani są uczniowie. Ponadto w BAEL do pracujących zaliczono osoby, które nie wykonywały pracy w badanym tygodniu, ale deklarowały jej posiadanie jako pomagający członkowie rodziny. W NSP nie były one zaliczane do pracujących.

Analizując wskazane wyżej kategorie rozbieżności należy zauważyć, że jednym z zasadniczych elementów istotnych w przypadku porównywania wyników badań jest kwestia zgodności definicyjnej. W raporcie opublikowanym przez GUS [GUS 2012b] zwraca się uwagę na fakt, że mamy do czynienia z rozbieżnością definicji charakterystyk rynku pracy. Przykładem może być stopa bezrobocia rejestrowanego oraz pomiaru dokonywanego w ramach BAEL i NSP 2011.

Ponadto uwypuklenia wymagają przyczyny rozbieżności wynikające z przyjętej w opracowaniu wyników NSP 2011 definicji badanej populacji. Szacunki NSP 2011 opracowywane są dla dwóch kategorii ludności: faktycznie zamieszkałej³ oraz rezydującej⁴. Ludność faktyczna uwzględnia osoby przebywające za granicą, natomiast ludność rezydująca dotyczy tylko osób przebywających w kraju przez co najmniej 12 miesięcy. Zgodnie z informacjami publikowanymi przez GUS⁵ na podstawie NSP 2011 ludność faktyczna to 38,501 mln osób, ludność rezydująca zaś jest o 1,257 mln osób mniejsza. Z ujęcia kategorii ludności, do której odnoszone są takie wielkości pracujących czy bezrobotnych, wynikają oczywiste konsekwencje dla wartości takich wskaźników rynku pracy, jak: stopa bezrobocia czy wskaźnik aktywności zawodowej.

3. Rozbieżności szacunków NSP 2011 i BAEL i ich konsekwencje

Rozbieżności dostrzeżono także, porównując wyniki NSP 2011 z BAEL przeprowadzone w pierwszym kwartale 2011 r. Liczba ludności aktywnej ekonomicznie

³ Ludności faktyczna to stali mieszkańcy (osoby zameldowane na pobyt stały lub deklarujące stały pobyt), z wyjątkiem przebywających w kraju, ale poza miejscem zamieszkania, przez okres powyżej 3 miesięcy, wszystkie osoby przebywające za granicą (bez względu na okres nieobecności), osoby przebywające czasowo przez okres powyżej 3 miesięcy, przybyłe z innego miejsca w kraju (nie są uwzględniani przebywający czasowo w Polsce cudzoziemcy).

⁴ Ludność rezydująca to stali mieszkańcy, z wyjątkiem osób przebywających poza miejscem zamieszkania przez co najmniej 12 miesięcy, bez względu na ich miejsce przebywania w kraju czy za granicą, osoby przebywające czasowo przez co najmniej 12 miesięcy, przybyłe z innego miejsca w kraju lub z zagranicy.

⁵ Należy dodać, że dane ze spisu publikowane były do tej pory trzykrotnie: w grudniu 2011 r., w marcu i lipcu 2012 r. i za każdym razem podawana była inna liczba ludności faktycznej. Te zmiany związane są bezpośrednio z korektami, jakich GUS na bieżąco dokonuje.

według spisu była niższa o 1,125 mln osób aniżeli ta, którą podaje BAEL. Kierunek rozbieżności był jednak różny, jeśli weźmie się pod uwagę różne grupy wieku. W młodszych grupach wieku zanotowano przeszacowanie, a w starszych, to jest powyżej 25 roku życia, mieliśmy do czynienia ze znacznym niedoszacowaniem (por. tab. 1).

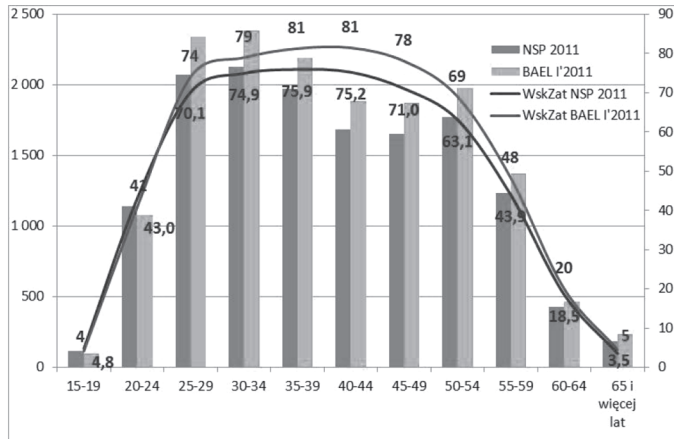
Tabela 1. Ludność w wieku 15 i więcej lat według aktywności ekonomicznej i wieku, BAEL I kwartał 2011, NSP 2011 – ludność faktyczna

Źródło informacji	BAEL I kwartał 2011			NSP 2011			Porównanie BAEL – NSP		
	ogółem	mężczyźni	kobiety	ogółem	mężczyźni	kobiety	ogółem	mężczyźni	kobiety
Wiek	w tys.								
Ogółem	17 646	9 644	8 002	16 521	9 076	7 445	1 125	568	557
15-17	24	17	7	27	17	11	-3	0	-4
18-19	122	75	47	161	96	65	-39	-21	-18
20-24	1 457	861	596	1 583	906	677	-126	-45	-81
25-29	2 658	1 456	1 201	2 426	1 336	1 091	232	121	110
30-34	2 603	1 440	1 163	2 379	1 303	1 076	224	137	87
35-39	2 360	1 248	1 112	2 187	1 174	1 012	173	74	100
40-44	2 019	1 058	960	1 873	985	888	146	73	73
45-49	2 029	1 027	1 002	1 848	953	895	181	74	107
50-54	2 158	1 104	1 054	1 996	1 028	968	162	76	86
55-59	1 481	891	590	1 388	836	552	93	55	39
60-64	499	325	174	469	324	145	30	1	29
64 i więcej	237	141	95	186	119	66	51	22	29
Przedprodukcyjny	24	17	7	27	17	11	-3	0	-4
Produkcyjny	17 211	9 486	7 726	16 164	8 940	7 223	1 048	546	503
Poprodukcyjny	411	141	269	331	119	211	80	22	58

Źródło: [GUS 2011a, s. 13; GUS 2011b, s. 15; GUS 2012a, s. 37; GUS 2012b, s. 143; GUS 2012c, s.15-17].

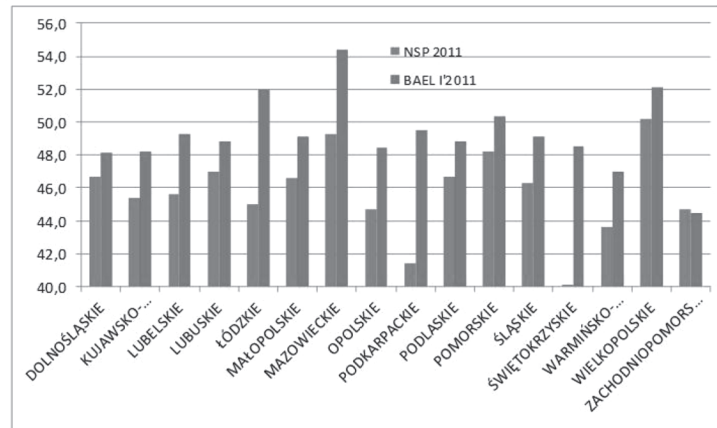
Na wyraźne rozbieżności wskazuje porównanie wyników BAEL i NSP 2011 pod względem struktury wieku liczby zatrudnionych i wskaźnika zatrudnienia (por. rys. 1a). Liczba osób pracujących podawana na podstawie NSP 2011 była niższa, aniżeli wynikało to z BAEL.

Największe różnice wartości wskaźnika zatrudnienia dochodzące do 7 punktów procentowych możemy zaobserwować dla grup wieku od 25 do 50 lat, a więc dla grup najistotniejszych z punktu widzenia rynku pracy. Pewną spójność dostrzec



Rys. 1a. Pracujący i wskaźnik zatrudnienia według wieku, BAEL I'2011-NSP 2011

Źródło: opracowanie własne na podstawie danych GUS.



Rys. 1b. Wskaźnik zatrudnienia w przekroju województw, BAEL I'2011-NSP 2011

Źródło: opracowanie własne na podstawie danych GUS.

można jedynie w najmłodszych i najstarszych grupach wieku, gdzie zaobserwowane różnice nie były duże.

Rozbieżności pomiędzy wynikami BAEL i NSP 2011 pod względem liczby zatrudnionych i wskaźnika zatrudnienia uwypuklają się w przekroju terytorialnym uwzględniającym województwa (por. rys. 1b). Według BAEL, liczba pracujących przykładowo w województwie świętokrzyskim była wyższa aż o 44,5% aniżeli według NSP 2011. W województwie łódzkim z kolei o 37,3%, w lubelskim zaś czy podkarpackim – o 22%. Wyniki publikowane na podstawie BAEL zawierają poza szacunkiem parametrów wielkość błędu estymacji. W przypadku liczby osób pracujących dla całego kraju względny błąd szacunku wynosi 0,8%. W przekroju województw jest on wyższy i wynosi na przykład w województwie świętokrzyskim około 3,8%. Tak więc zakres rozbieżności może być umniejszony o wartość błędu. Nie zmienia to jednak faktu, że i tak wyniki badania NSP 2011 wskazują na braki w liczbie zatrudnionych.

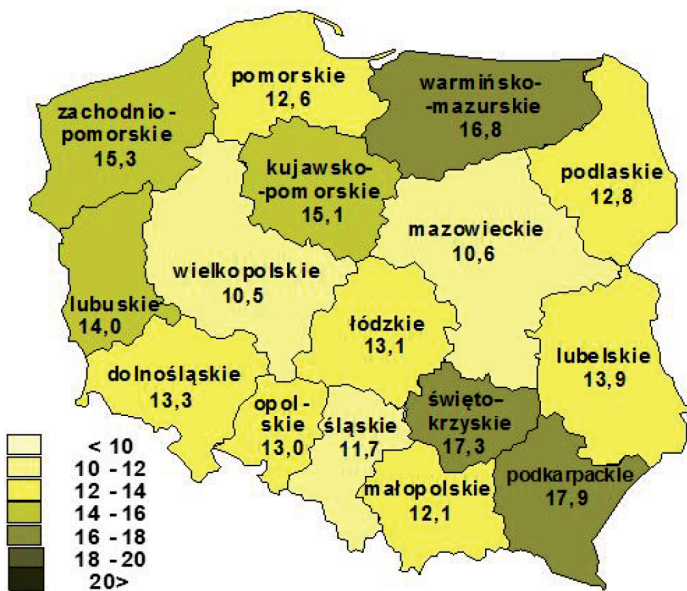
W celu pełniejszego zobrazowania rozbieżności pomiędzy wynikami NSP 2011 i BAEL zestawione zostały różne kategorie ludności wyróżnione w ramach rynku pracy, takie jak: aktywni zawodowo, pracujący, bezrobotni, wskaźnik aktywności ekonomicznej ludności, wskaźnik zatrudnienia oraz stopa bezrobocia, wyznaczone na podstawie dwóch źródeł, por. tab. 2.

Tabela 2. Ludność według statusu na rynku pracy, NSP 2011, BAEL I'2011

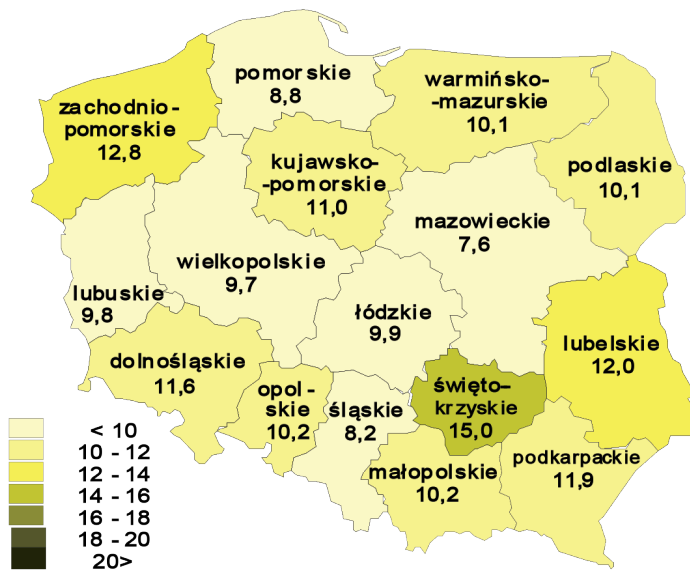
Kategoria ludności Źródło	Aktywni zawodowo	Pracujący	Bezrobotni	Współ. aktywności zawodowej	Wskaźnik zatrudnienia	Stopa bezrobocia
	tys.			%		
NSP 2011	16 521,4	14 375,8	2 145,6	53,3	46,4	13,0
BAEL I'2011	17 646,0	15 875,0	1 771,0	55,6	50,0	10,0
	tys.			pkt proc.		
BAEL I'2011-NSP 2011	1 124,6	1 499,2	374,6	2,3	3,6	-3,0

Źródło: opracowanie własne na podstawie danych GUS.

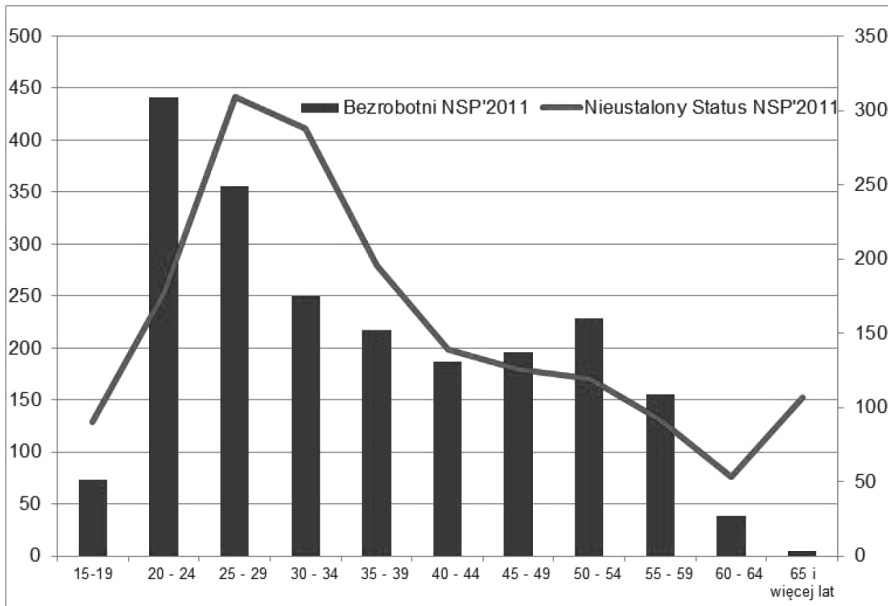
Podjęto próbę wskazania przyczyn rozbieżności. W pierwszej kolejności uwzględniono różne kategorie ludności pracującej i bezrobotnych. Okazało się, że w NSP 2011 zarejestrowano więcej osób bezrobotnych niż w BAEL. Podobne rozbieżności dotyczą również stopy bezrobocia i sięgają 4 punktów procentowych w przekroju grup wiekowych. Jeżeli weźmiemy pod uwagę z kolei zróżnicowanie terytorialne, to nadwyżka bezrobotnych wynosi 22% w województwie świętokrzyskim, a w kilku województwach, takich jak: warmińsko-mazurskie czy zachodniopomorskie, sięga nawet 40% (por. rys. 2a, 2b).



Rys. 2a. Stopa bezrobocia w przekroju województw, NSP 2011
 Źródło: opracowanie własne na podstawie danych GUS.

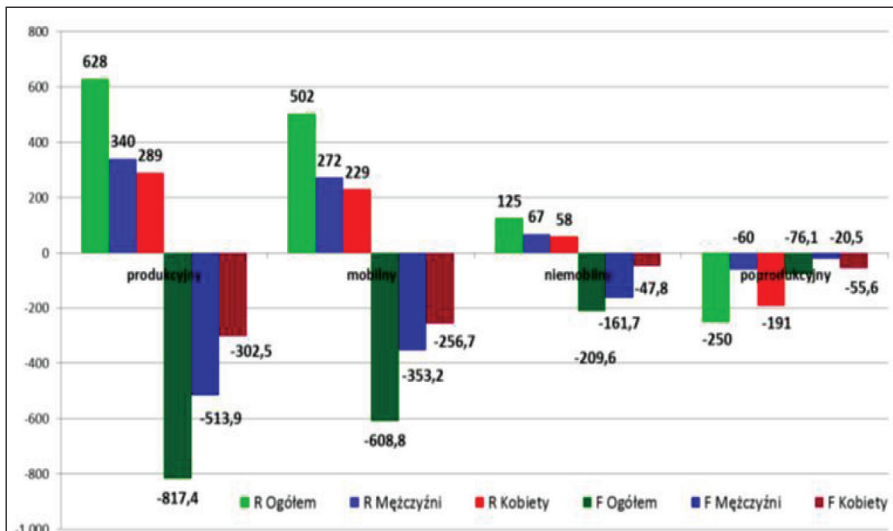


Rys. 2b. Stopa bezrobocia w przekroju województw, BAEL
 Źródło: opracowanie własne na podstawie danych GUS.



Rys. 3a. Ludność bezrobotna i o niestabilnym statusie na rynku pracy, NSP 2011

Źródło: opracowanie własne na podstawie danych GUS.



Rys. 3b. Różnica szacunków ludności rezydującej (R) i faktycznej (F) na podstawie BAEL I'2011 i NSP 2011 według „ekonomicznych” grup wieku

Źródło: opracowanie własne na podstawie danych GUS.

Tabela 3. Podstawowe charakterystyki rynku pracy ogółem, dla różnych kategorii ludności

Wyszczególnienie	Ogółem	Aktywni zawodowo			Bierni zawodowo	Wsp. aktywności zawodowej	Wsk. zatrudnienia	Stopa bezrobocia
		razem	pracujący	bezrobotni				
	w tysiącach					w %		
NSP 2011 Ludność faktyczna w wieku								
Przedprodukcyjnym	7 318							
Produkcyjnym	24 797	16 164	14 031	2 133	7 121	65,2	56,6	13,2
Produkcyjnym mobilnym	15 419	10 608	9 087	1 520	3 664	68,8	58,9	14,3
Produkcyjnym niemobilnym	9 379	5 556	4 944	612	3 458	59,2	52,7	11,0
Poprodukcyjnym	6 730							
NSP 2011 Ludność rezydująca w wieku								
Przedprodukcyjnym	37 244							
Produkcyjnym	23 352					69,2	60,1	13,2
Produkcyjnym mobilnym	14 308					74,1	63,5	14,3
Produkcyjnym niemobilnym	9 044					61,4	54,7	11,0
Poprodukcyjnym	6 686							
BAEL I 2011 Ludność faktyczna w wieku								
Produkcyjnym	23 980	17 211	15 448	1 763	6 769	71,8	64,4	10,2
Produkcyjnym mobilnym	14 810	11 219	9 939	1 278	3 593	75,8	67,1	11,4
Produkcyjnym niemobilnym	9 170	5 992	5 509	485	3 176	65,3	60,1	8,1
Poprodukcyjnym	6 436	411	403	7	6 025	6,4	6,4	1,7

Źródło: opracowanie własne na podstawie danych GUS.

Szukając przyczyn, sięgnięto również po informacje dotyczące ludności o nieustalonym statusie na rynku pracy. Na rysunku 3a przedstawiono strukturę według wieku bezrobotnych o nieustalonym statusie. Różnice w grupach największej aktywności ekonomicznej są zdecydowanie największe. Dodatkowo zaobserwowano umiarkowanie silną korelację dodatnią między liczbą osób o nieustalonym statusie na rynku pracy (wg NSP 2011) a różnicą pomiędzy nadwyżką osób bezrobotnych według NSP 2011 w stosunku do BAEL I'2011 ($\Delta BEZ(NSP-BAEL)$; $NSTATUS$) = 0,4212.

Rozbieżności odniesiono także do dwóch kategorii ludności: rezydującej i faktycznej (por. rys. 3b). Okazuje się, że jeżeli nie zostałyby uwzględniona ludność rezydująca, to informacje z NSP 2011 byłyby odpowiednio niedoszacowane w stosunku do BAEL (por. tab. 3). Z kolei w odniesieniu do ludności faktycznej wartości wskaźników są przeszacowane. W grupie wieku poprodukcyjnego w każdym przypadku różnica w BAEL jest mniejsza niż na podstawie wyników NSP 2011.

Warto na koniec podkreślić, że wyznaczenie współczynników dotyczących aktywności ekonomicznej z uwzględnieniem ludności rezydującej nie należy do prostych zadań. Informacje bowiem dotyczące ludności rezydującej pochodzące z badań nie w pełni pokrywają się ze stanem rzeczywistym. Stąd też podejmowane są próby poszukiwania ludności na podstawie statystyki lustrzanej w innych krajach, do których Polacy emigrują.

4. Wnioski

Spis Ludności 2011 był ogromnym przedsięwzięciem wykorzystującym dane rejestrów administracyjnych i badań reprezentacyjnych, których integracja wymagała zastosowania takich nowoczesnych metod, jak integracja statystyczna oraz łącząca różne źródła estymacja dla małych domen. Opracowanie wieloźródłowego materiału musi uwzględniać mechanizm wzajemnej kontroli, badania zgodności, przeprowadzenia analiz porównawczych, co w efekcie bardziej ją uwiarygadnia.

1. W warunkach korzystania z wielu źródeł pojawia się 'naturalne' niebezpieczeństwo uzyskania niejednoznacznych wyników. Rozbieżnych szacunków nie należy postrzegać w kategoriach wstydlivych błędów, lecz określić ich przyczyny i przy zastosowaniu kalibracji podejmować próby dostarczenia spójnych szacunków. Integralnym elementem metod wykorzystujących różne źródła informacji jest wbudowany w ich algorytm system oceny jakości źródeł.

2. Korzystanie z wielu źródeł informacji wymaga opracowania zintegrowanego podejścia do oceny ich jakości oraz jej poprawy. Poprawa jakości spisów możliwa jest między innymi poprzez wnikliwą i rzetelną ocenę dotychczasowych badań. Wyniki tych analiz winny być wykorzystane do korekty spisów oraz poprawy jakości rejestrów administracyjnych. W badaniu jakości spisu warto korzystać z rekomendacji międzynarodowych oraz wzorować się na doświadczeniach innych krajów.

3. Wykorzystanie różnych źródeł informacji, w tym badań specjalnych, implikuje konieczność uwzględnienia w ocenie spisów także błędów losowych. Szacowanie błędów losowych i nielosowych w przypadku korzystania ze zintegrowanych danych, np. rejestr – badanie specjalne, wymaga wypracowania nowej koncepcji teoretycznej.

Dążąc do poprawy jakości spisów powszechnych w Polsce, warto rozważyć opracowanie strategii wykorzystującej: możliwości nowoczesnej metodologii, dogłębną eksplorację rezultatów analiz demograficznych i wyników niezależnego badania kontrolnego.

Literatura

- GUS, 2011a, *Bezrobocie rejestrowane w I kwartale 2011 r*, Warszawa.
- GUS, 2011b, *Zatrudnienie i wynagrodzenia w gospodarce narodowej w I kwartale 2011 r*, Warszawa.
- GUS, 2012a, *Wyniki Narodowego Spisu Powszechnego Ludności i Mieszkań 2011. Podstawowe informacje o sytuacji demograficzno-społecznej ludności Polski oraz zasobach mieszkaniowych*, Warszawa.
- GUS, 2012b, *Raport z wyników, Narodowy Spis Powszechny Ludności i Mieszkań 2011*, Warszawa.
- GUS, 2012c, *Stan i struktura ludności oraz ruch naturalny w przekroju terytorialnym, stan w dniu 31.12.2010* GUS, Warszawa (dostęp internetowy 18 sierpnia 2012; http://www.stat.gov.pl/gus/5840_655_PLK_HTML.htm).
- GUS, 2012d, *Różnice wyników BAEL i NSP 2011 – komentarz, Materiał na konferencję prasową w dniu 31.05.2012*, GUS, Warszawa (dostęp internetowy 15 sierpnia 2012; http://www.stat.gov.pl/gus/5840_13113_PLK_HTML.htm).

DIVERGENCE IN NATIONAL CENSUS 2011 AND LFS ESTIMATES

Summary: The aim of the article is to determine causes of discrepancies between labour market estimates based on the 2011 National Census and the Labour Force Survey. Another issue addressed in the article is the influence and significance of the observed discrepancies. These considerations are derived from a comparative analysis of the current labour market situation based on data from the 2011 National Census and the Labour Force Survey.

Keywords: register-based census, Labour Force Survey, data quality assessment.