

Cena zł 12,00
(VAT 5%)

Indeks 381306
PL ISSN 0043-518X
e-ISSN 2543-8476

WIADOMOŚCI STATYSTYCZNE

GŁÓWNY
URZĄD
STATYSTYCZNY

POLSKIE
TOWARZYSTWO
STATYSTYCZNE

MIESIĘCZNIK
ROK LXII
WARSZAWA
SIERPIEŃ 2017

Nr **8** (675)



WIADOMOŚCI STATYSTYCZNE

GŁÓWNY
URZĄD
STATYSTYCZNY

POLSKIE
TOWARZYSTWO
STATYSTYCZNE

MIESIĘCZNIK
ROK LXII
WARSZAWA
SIERPIEŃ 2017

Nr **8** (675)

KOLEGIUM REDAKCYJNE

dr Marek Cierpiał-Wolan (redaktor naczelny), dr hab. Andrzej Młodak (zastępca redaktora naczelnego), mgr Renata Bielak, dr Jacek Kowalewski, dr Jan Kubacki, mgr Władysław Wiesław Łagodziński, dr Grażyna Marciniak, dr Stanisław Paradysz, dr hab. Mateusz Pipień, prof. dr hab. Bogdan Stefanowicz, dr Wioletta Wrzaszcz, dr inż. Agnieszka Zgierska

Sekretarz: Alina Świdarska

RADA NAUKOWA

dr Halina Dmochowska (przewodnicząca), dr hab. Bożena Balcerzak-Paradowska, prof. dr hab. Czesław Domański, dr hab. Elżbieta Gołata, prof. dr hab. Semen Matkowski, prof. dr hab. Włodzimierz Okrasa, prof. dr hab. Józef Oleński, prof. dr hab. Tomasz Panek, doc. ing. Iveta Stankovicova, prof. dr hab. Józef Zegar

Sekretarz: Justyna Gustyn

REDAKCJA

al. Niepodległości 208, 00-925 Warszawa, gmach GUS, pok. 353, tel. 22 608 32 25

<http://stat.gov.pl/czasopisma/wiadomosci-statystyczne>

Alina Świdarska (a.swiderska@stat.gov.pl)

Elżbieta Grabowska (e.grabowska@stat.gov.pl)

Wersja internetowa jest wersją pierwotną czasopisma



ZAKŁAD WYDAWNICTW STATYSTYCZNYCH

al. Niepodległości 208, 00-925 Warszawa, tel. 22 608 31 45.

Informacje w sprawach nabywania czasopism tel. 22 608 32 10, 22 608 38 10.

Zbigniew Karpiński (redaktor techniczny), Ewa Krawczyńska (skład i łamanie),

Wydział Korekty pod kierunkiem Bożeny Gorczyzycy, mgr Andrzej Kajkowski (wykresy).

Indeks 381306

Prenumerata realizowana przez RUCH S.A.

Zamówienia na prenumeratę w wersji papierowej i na e-wydania można składać bezpośrednio na stronie www.prenumerata.ruch.com.pl.

Ewentualne pytania prosimy kierować na adres e-mail: prenumerata@ruch.com.pl lub kontaktując się z Centrum Obsługi Klienta „RUCH” pod numerami: 22 693 70 00 lub 801 800 803 — czynne w dni robocze w godzinach 7⁰⁰—17⁰⁰.

Koszt połączenia według taryfy operatora.

SPIS TREŚCI

STUDIA METODOLOGICZNE

- Beata Bieszk-Stolorz* — Zastosowanie metod analizy trwania w badaniu form wychodzenia z bezrobocia 5

STATYSTYKA W PRAKTYCE

- Agnieszka Matulska-Bachura* — Metody pomiaru wskaźnika wolumenu usług na przykładzie działalności związanej z administracyjną obsługą biura 19
- Agnieszka Nocko* — Zróżnicowanie długości życia w zależności od płci i wykształcenia 41
- Marlena Piekut* — Klasyfikacja wiejskich gospodarstw domowych ze względu na kierunki i rodzaje rozchodów 53
- Romana Głowicka-Wołoszyn, Agnieszka Kozera, Feliks Wysocki* — Identyfikacja wewnętrznych uwarunkowań samodzielności i atrakcyjności finansowej gmin wiejskich województwa wielkopolskiego z wykorzystaniem biplotu 74

INFORMACJE. PRZEGLĄDY. RECENZJE

- XXXV Międzynarodowa Konferencja Naukowa *Wielowymiarowa Analiza Statystyczna 2016* 85
- Wydawnictwa GUS — lipiec 2017 r. (oprac. *Justyna Gustyn*) 91
- Do Autorów 93

CONTENTS

METHODOLOGICAL STUDIES

- Beata Bieszk-Stolorz* — Application of survival analysis method for the examination of unemployment leaving forms **5**

STATISTICS IN PRACTICE

- Agnieszka Matulska-Bachura* — Methods for compiling the services volume index by the example of office administrative and support activities **19**
- Agnieszka Nocko* — Differentiation of life expectancy by sex and education **41**
- Marlena Piekut* — Classification of rural households due to directions and types of outgoings **53**
- Romana Głowicka-Wołoszyn, Agnieszka Kozera, Feliks Wysocki* — Internal determinants identification of the autonomy and financial attractiveness of rural gminas in Wielkopolskie voivodship using biplot **74**

INFORMATION. REVIEWS. COMMENTS

- XXXV International Scientific Conference on Multivariate Statistical Analysis 2016 **85**
- Publications of CSO — July 2017 (by *Justyna Gustyn*) **91**
- For the Authors **93**

STUDIA METODOLOGICZNE

Beata BIESZK-STOLORZ

Zastosowanie metod analizy trwania w badaniu form wychodzenia z bezrobocia

Streszczenie. *Celem artykułu jest przedstawienie wybranych metod analizy trwania do oceny prawdopodobieństwa wyjścia z bezrobocia dla różnych rodzajów ryzyka konkurującego. W badaniu wykorzystano funkcję skumulowanej częstości i skumulowane prawdopodobieństwo warunkowe oraz dopełnienie do jedności estymatora Kaplana-Meiera. Za pomocą trzech estymatorów porównano prawdopodobieństwo wyrejestrowania z powodu podjęcia pracy, wykreślenia i pozostałych przyczyn. Analizę przeprowadzono na podstawie danych indywidualnych osób zarejestrowanych w Powiatowym Urzędzie Pracy w Szczecinie.*

Słowa kluczowe: estymator Kaplana-Meiera, funkcja skumulowanej częstości, skumulowane prawdopodobieństwo warunkowe, ryzyko konkurujące, bezrobocie.

JEL: C41, J64

W trakcie analizy czasu trwania zjawisk pojawia się problem występowania zdarzeń konkurujących. Na przykład w zastosowaniach inżynierskich ryzyko konkurujące związane jest z wpływem elementów składowych na cały system. Awaria któregośkolwiek ze składników powoduje awarię całości. Obserwuje się, który składnik spowodował awarię i czas, w którym system zawodzi.

Gooley, Leisenring, Crowley i Storer (1999) definiują ryzyko konkurujące jako zdarzenie, którego wystąpienie wyklucza zajście innego zdarzenia lub fundamentalnie zmienia prawdopodobieństwo tego innego zdarzenia. Należy założyć, że wystąpienie zdarzenia określonego typu nie ma żadnego wpływu na prawdopodobieństwo jakichkolwiek innych zdarzeń (Crowder, 1994, 1996,

1997). Badana jednostka jest narażona na różne rodzaje ryzyka w tym samym czasie, ale przypuszcza się, że ewentualne zdarzenie wynika tylko z jednego z tych czynników, który nazywa się przyczyną niepowodzenia (Aly, Kochar i McKeague, 1994).

Analiza czasu trwania zjawisk związana jest również z występowaniem obserwacji cenzurowanych. Powstają one, gdy czas trwania dla badanej jednostki jest znany jedynie częściowo. W przypadku analizy danych kohortowych najczęściej występuje cenzurowanie prawostronne, które zachodzi wówczas, gdy zdarzenie końcowe nie wystąpiło w okresie obserwacji. Za obserwacje uznawane są również takie sytuacje, w których badana jednostka znika z pola widzenia lub występuje zdarzenie kończące obserwację, wykluczające, by mogło dojść do zdarzenia właściwego (Pepe, 1991), czyli zdarzenie konkurujące.

Kalbfleisch i Prentice (1980, s. 164) zidentyfikowali trzy różne problemy, które pojawiają się w analizie danych z ryzykiem konkurującym:

- 1) szacowanie zależności pomiędzy zmiennymi oraz współczynnika wystąpienia danego rodzaju zdarzenia;
- 2) badanie współzależności pomiędzy typami zdarzeń w ramach określonego zbioru warunków;
- 3) szacowanie ryzyka awarii z określonego powodu, biorąc pod uwagę usunięcie niektórych lub wszystkich pozostałych przyczyn awarii; problem ten jest uważany za klasyczny w analizie ryzyka konkurującego.

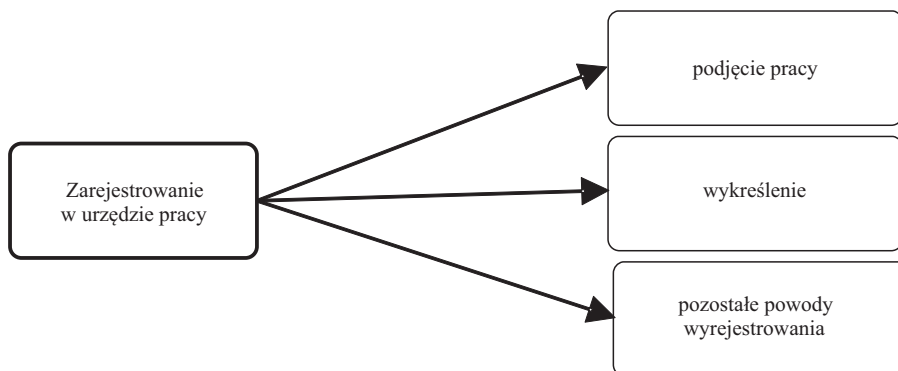
W literaturze obok zdarzeń konkurujących bezwarunkowych rozważa się również warunkowe modele ryzyka konkurującego. W przypadku pierwszego zdarzenia zakłada się jego niezależność. W przypadku tych drugich prawdopodobieństwo przejścia między stanami zależy nie tylko od zmiennych objaśniających, ale również od czasu i typu pobytu we wcześniejszym stanie (Landmesser, 2008b).

Celem badania jest przedstawienie i zastosowanie trzech metod analizy trwania do oceny prawdopodobieństwa wyjścia z bezrobocia dla różnych rodzajów ryzyka konkurującego bezwarunkowego.

DANE WYKORZYSTANE W BADANIU

W badaniu wykorzystano dane indywidualne osób zarejestrowanych w Powiatowym Urzędzie Pracy w Szczecinie (PUP).

W badaniu posłużono się anonimowymi danymi indywidualnymi 22078 osób bezrobotnych zarejestrowanych w 2013 r. w PUP i obserwowanych do końca 2014 r. Zdarzeniem finalizującym obserwację jednostki był moment wyrejestrowania z urzędu. Analizowano czas od zarejestrowania do wyrejestrowania z określonego powodu. W rejestrach urzędów pracy podaje się kilkadziesiąt powodów wyrejestrowania osoby bezrobotnej i tylko kilka z nich związanych jest z podjęciem pracy. Powody te pogrupowano i rozpatrywano trzy grupy zdarzeń konkurujących.

Wykr. 1. GRUPY ZDARZEŃ KONKURUJĄCYCH

Źródło: opracowanie własne na podstawie form wyjścia z bezrobocia, które przyjęto w badaniu.

Część obserwacji nie zakończyła się zdarzeniem, czyli wyrejestrowaniem w analizowanym okresie. Stanowią one obserwacje cenzurowane prawostronnie (powód cenzurowania 1). W zestawieniu podano wielkość grup przyczyn wyrejestrowania.

**ZESTAWIENIE POWODÓW WYREJESTROWANIA
I ODPOWIADAJĄCYCH IM OBSERWACJI CENZUROWANYCH**

Zdarzenia kończące obserwację		Obserwacje cenzurowane			
powód wyrejestrowania	liczba osób	powód cenzurowania 1	liczba osób	powód cenzurowania 2	liczba osób
Praca w sensie ogólnym	9633	zdarzenie nie nastąpiło do końca 2014 r.	1856	zdarzenie nie nastąpiło do końca 2014 r. lub nastąpiło z powodu wykreślenia albo pozostałych powodów	12445
Wykreślenie	8965			zdarzenie nie nastąpiło do końca 2014 r. lub nastąpiło z powodu podjęcia pracy albo pozostałych powodów	13113
Pozostałe	1624			zdarzenie nie nastąpiło do końca 2014 r. lub nastąpiło z powodu podjęcia pracy albo wykreślenia	20454

Źródło: opracowanie własne.

Na każdy z trzech głównych powodów wyrejestrowania miały wpływ powody cząstkowe.

Pracę w sensie ogólnym można podzielić na trzy główne podgrupy: podjęcie pracy lub innego zatrudnienia, rozpoczęcie pracy subsydiowanej przez urząd (m.in. w ramach robót publicznych czy interwencyjnych, utworzonego przez

pracodawcę dodatkowego miejsca pracy z tytułu udzielonej pożyczki czy dofinansowania wynagrodzenia za zatrudnienie bezrobotnego, który ukończył 50 lat) oraz działalność gospodarczą (podjęcie pozarolniczej działalności gospodarczej, przyznanie jednorazowo środków na podjęcie działalności gospodarczej, rozpoczęcie działalności gospodarczej ze środków PFRON).

Wykreślenie jest grupą powodów, w których zarejestrowana osoba bezrobotna nie wykazała się chęcią współpracy z urzędem i została z własnej winy lub na własną prośbę wykreślona z rejestru. Chodzi m.in. o: odmowę bez uzasadnionej przyczyny przyjęcia propozycji odpowiedniego zatrudnienia lub innej pracy zarobkowej, wykonywania prac interwencyjnych lub robót publicznych albo udziału w szkoleniu, stażu, przygotowaniu zawodowym w miejscu pracy; niestawienie się w PUP w wyznaczonym terminie i niepowiadomienie w terminie do 5 dni o uzasadnionej przyczynie tego niestawiennictwa; nieprzedstawienie zaświadczenia o niezdolności do pracy wskutek choroby; odmowę bez uzasadnionej przyczyny udziału w działaniach w ramach Programu Aktywizacja i Integracja (PAI) lub przerwanie z własnej winy udziału w działaniach w ramach PAI; brak gotowości do podjęcia pracy przez okres co najmniej 10 dni; przerwanie z własnej winy szkolenia, stażu, realizacji indywidualnego planu działań oraz wniosek bezrobotnego o wykreślenie z ewidencji.

Pozostałe powody wyrejestrowania, jak wykazały wcześniejsze badania, miały marginalny wpływ na prawdopodobieństwo wyrejestrowania z urzędu, dlatego postanowiono rozpatrywać je razem. Składają się nań m.in.: wyjazd za granicę na okres co najmniej 30 dni, zmiana miejsca zameldowania lub pobytu zainteresowanego poza obszar działania PUP, niezdolność do pracy wskutek choroby lub przebywania w zamkniętym ośrodku odwykowym przez nieprzerwany okres 90 dni, rozpoczęcie szkolenia organizowanego przez podmiot inny niż PUP, zgon, podjęcie nauki w szkole w systemie dziennym, nabycie prawa do emerytury, przyznanie prawa do renty, nabycie prawa do świadczenia rehabilitacyjnego, pobieranie zasiłku stałego, podleganie ubezpieczeniu emerytalno-rentowemu z tytułu stałej pracy jako domownik w gospodarstwie rolnym, pobieranie świadczenia lub dodatku do zasiłku rodzinnego z tytułu samotnego wychowywania dziecka, pobieranie zasiłku dla opiekuna oraz przyznanie prawa do pobierania świadczenia/zasiłku przedemerytalnego.

METODYKA BADANIA

Analiza trwania, powszechnie stosowana w demografii i naukach medycznych do badania czasu trwania życia ludzkiego, może być wykorzystywana do oceny czasu trwania zjawisk ekonomicznych. Można w ten sposób badać: żywotność firm (Markowicz, 2012), aktywność ekonomiczną ludności (Landmesser, 2013), dynamikę ubóstwa w miejskich i wiejskich gospodarstwach domowych (Sączewska-Piotrowska, 2016), ryzyko kredytowe (Matuszyk, 2015; Wysocka, 2015) i czas trwania bezrobocia (Bieszk-Stolorz, 2013; Bieszk-Stolorz i Markowicz, 2012a). Analizie podlega czas trwania jednostki w danym stanie

(okresy: działania firmy, bezrobocia i spłaty kredytu), który jest zmienną losową T . Podstawą tego typu analizy jest funkcja trwania, nazywana również funkcją przeżycia, zdefiniowana następująco:

$$S(t) = P(t > T) = 1 - F(T) \quad (1)$$

gdzie:

T — czas trwania zjawiska,

$F(T)$ — dystrybuanta rozkładu zmiennej losowej T .

Najczęściej stosowany jest tutaj estymator Kaplana-Meiera (Kaplan i Meier, 1958):

$$\hat{S}(t) = \prod_{j: t_j \leq t} \left(1 - \frac{d_j}{n_j} \right) \quad (2)$$

gdzie:

d_j — liczba zdarzeń w momencie t_j ,

n_j — liczba jednostek zagrożonych do momentu t_j .

Funkcja trwania informuje nas o tym, jakie jest prawdopodobieństwo niezajścia zdarzenia co najmniej do czasu t . W demografii i naukach medycznych, gdzie zdarzeniem jest najczęściej zgon, estymacji podlega prawdopodobieństwo przeżycia jednostki. Czasami wygodniej jest analizować dystrybuantę $F(T)$, która wyraża prawdopodobieństwo, że zdarzenie zajdzie najpóźniej do czasu t . Jeżeli zdarzeniem jest podjęcie pracy przez zarejestrowaną w urzędzie osobę bezrobotną, to estymator funkcji trwania informuje o tym, jakie jest prawdopodobieństwo pozostania w rejestrze, a estymator dystrybuanty pozwala określić, jakie jest prawdopodobieństwo podjęcia pracy.

W badaniu wykorzystano trzy sposoby oceny ryzyka konkurującego bezwzajemnego wywodzące się z analizy trwania.

Niech T i C będą zmiennymi losowymi ciągłymi, opisującymi odpowiednio czas do zajścia pewnego zdarzenia i czas do cenzurowania. W przypadku K konkurujących rodzajów ryzyka obserwacji podlegają pary (X, δ) , gdzie $X = \min(T, C)$ i $\delta = 0, 1, \dots, K$. Jeżeli dana obserwacja jest cenzurowana, to $\delta = 0$ oraz $\delta = 1, \dots, k$ dla obserwacji kończących się zdarzeniem (jednym z k konkurujących). W tym kontekście jedno ze zdarzeń k może być uznane za wydarzenie o podstawowym znaczeniu, a wszystkie inne będą rozpatrywane jako zdarzenia konkurujące.

W przypadku zdarzeń konkurujących powszechnie stosowany jest zwykły estymator Kaplana-Meiera (*KME*), jako dopełnienie do 1, określony następująco:

$$1 - KME = 1 - \hat{S}(t) \quad (3)$$

gdzie $\hat{S}(t)$ jest estymatorem Kaplana-Meiera.

W tym przypadku jedno ze zdarzeń k jest uznane za wydarzenie o podstawowym znaczeniu, a wszystkie inne będą rozpatrywane jako obserwacje cenzurowane. Jest to prawdopodobieństwo wystąpienia zdarzenia k przy założeniu, że występuje tylko jeden rodzaj ryzyka ($\delta = 1$). Zastosowanie tego estymatora prowadzi do przeszacowania wartości prawdopodobieństwa (Klein i Moeschberger, 1984; Sherif Bintu, 2004).

Drugi estymator, funkcja skumulowanej częstości występowania (*Cumulative Incidence Function — CIF*) został po raz pierwszy zaproponowany przez Kalbfleischa i Prentice'a (2002) i jest to skumulowane prawdopodobieństwo wystąpienia zdarzenia z powodu k przed czasem t przy założeniu, że jednostka jest narażona na wystąpienie któregośkolwiek ryzyka konkurującego k (Bryant i Dignam, 2004). Jest on zdefiniowany jako (Klein i Moeschberger, 2003, s. 52):

$$CIF_k(t) = P(t \leq T, \delta = k) = \int_0^t S(u) h_k(u) du = \int_0^t S(u) dH_k \quad \text{dla } k = 1, 2, 3, \dots, K \quad (4)$$

gdzie:

$H_k(t)$ — wyspecyfikowana (dla określonego k) funkcja hazardu skumulowanego,
 $S(t)$ — funkcja przeżycia.

Niech $t_1 < t_2 < \dots < t_i < \dots < t_n$ będą momentami zachodzenia zdarzeń. Analogicznie do standardowej skumulowanej funkcji w analizie przeżycia, skumulowana funkcja hazardu $H_k(t)$ dla przyczyny k może być określona przez estymator Nelsona-Aalena:

$$\hat{H}_k(t) = \sum_{j: t_j \leq t} \frac{d_{kj}}{n_j} \quad (5)$$

gdzie:

d_{kj} — liczba zdarzeń z powodu wystąpienia przyczyny k ,

n_j — liczba osób narażonych na ryzyko w czasie t_j .

Ogólna funkcja przeżycia $S(t)$ może być określona na podstawie estymatora Kaplana-Meiera (2). Po połączeniu estymatorów (4) i (5) funkcję skumulowanej

częstości występowania z powodu zaistnienia przyczyny k (Marubini i Valsecchi, 1995) można oszacować jako:

$$\widehat{CIF}_k(t) = \sum_{j: t_{j-1} \leq t} \hat{S}(t_{j-1}) \frac{d_{kj}}{n_j} \quad (6)$$

Funkcja skumulowanej częstości pozwala określić wzorce zajścia zdarzenia z powodu k oraz ocenić, w jakim stopniu każdy powód przyczynia się do całkowitego wyniku.

Jeśli $\sum_{k=1}^K d_{kj} = d_j$, to prawdziwa jest zależność:

$$\sum_{k=1}^K \widehat{CIF}_k(t) = 1 - \hat{S}(t) \quad (7)$$

Jeżeli nie ma zdarzeń konkurujących, to zachodzi równość:

$$\widehat{CIF}(t) = 1 - \hat{S}(t) \quad (8)$$

Trzeci estymator — skumulowane warunkowe prawdopodobieństwo wystąpienia zdarzenia (*Cumulative Probability Conditional — CPC*), to zdarzenie zachodzące z powodu k przed czasem t , pod warunkiem że przed czasem t nie zaszło żadne pozostałe zdarzenie konkurujące określone wzorem (Kleinbaum i Klein, 2005, s. 420):

$$CPC_k(t) = P(t \leq T_k, T_P > t) \quad \text{dla } k = 1, 2, 3, \dots, K \quad (9)$$

gdzie T_p jest czasem do wystąpienia jakiegokolwiek zdarzenia konkurującego.

Estymator CPC_k jest określony następująco (Pepe, 1991; Pepe i Mori, 1993):

$$\widehat{CPC}_k(t) = \frac{\widehat{CIF}_k(t)}{1 - \widehat{CIF}_k(t)} \quad (10)$$

gdzie $\widehat{CIF}_k(t)$ jest skumulowaną funkcją częstości dla wszystkich rodzajów ryzyka konkurującego różnych od k .

Trzy przedstawione wyżej estymatory są związane nierównościami (Klein i Bajorunaite, 2004):

$$\widehat{CIF}_k(t) \leq 1 - KME \leq \widehat{CPC}_k(t) \quad (11)$$

przy czym równości zachodzą wtedy, gdy jest tylko jedno zdarzenie kończące obserwację (brak zdarzeń konkurujących).

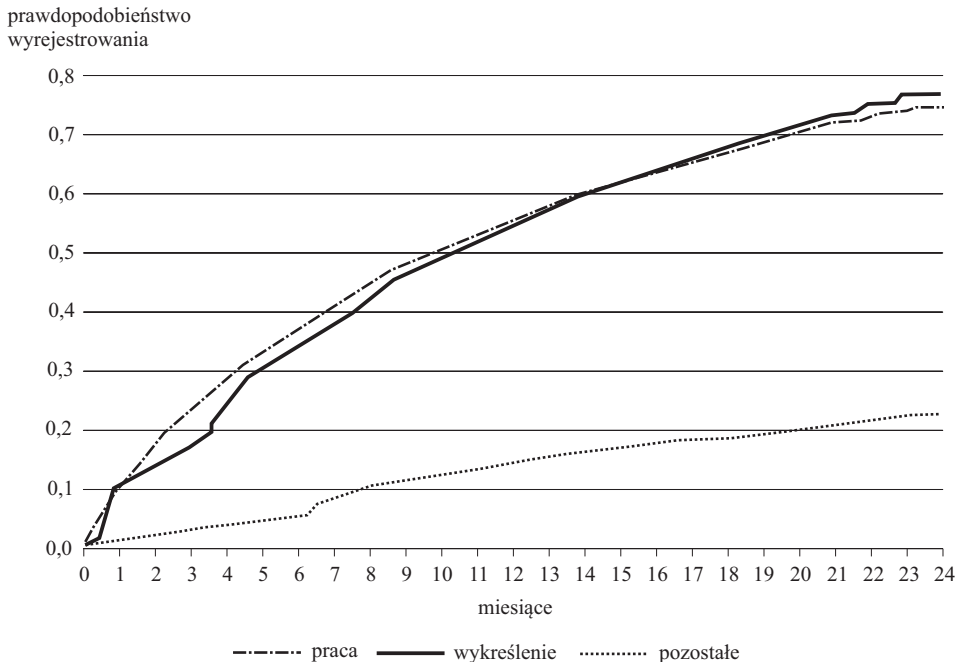
Innym sposobem oceny ryzyka konkurującego jest zastosowanie wielomianowego modelu logitowego. W analizie trwania traktuje się je jako modele ha-

zardu z czasem dyskretnym. Parametry tych modeli pozwalają wyznaczyć szansę względną zajścia zdarzeń konkurujących (Landmesser, 2008a; Bieszk-Stolorz i Markowicz, 2012b).

ANALIZA PRAWDOPODOBIEŃSTWA WYREJESTROWANIA Z URZĘDU PRACY

W pierwszym etapie badania wyznaczono wartości dla estymatora Kaplana-Meiera, określonego wzorem (3). Wyznacza się go przy założeniu, że nie występuje ryzyko konkurujące. W związku z tym obserwacje cenzurowane są tutaj dwojakiego rodzaju — takie, które nie zakończyły się zdarzeniem do końca 2014 r. lub zakończyły się z innego powodu. Każda grupa powodów wyrejestrowania jest analizowana oddzielnie. Wielkość analizowanych grup i odpowiadających im obserwacji cenzurowanych jest zawarta w zestawieniu (powód cenzurowania 2). Estymatory przedstawiono na wyk. 2. Od drugiego do szesnastego miesiąca trwania w bezrobociu prawdopodobieństwo podjęcia pracy było większe niż prawdopodobieństwo wykreślenia. Następnie sytuacja uległa zmianie i od szesnastego miesiąca do końca obserwacji prawdopodobieństwo wykreślenia z rejestru urzędu pracy nieznacznie przewyższyło prawdopodobieństwo podjęcia pracy. Prawdopodobieństwo wyrejestrowania z pozostałych powodów było niskie i na koniec okresu obserwacji wyniosło maksymalnie 0,0228.

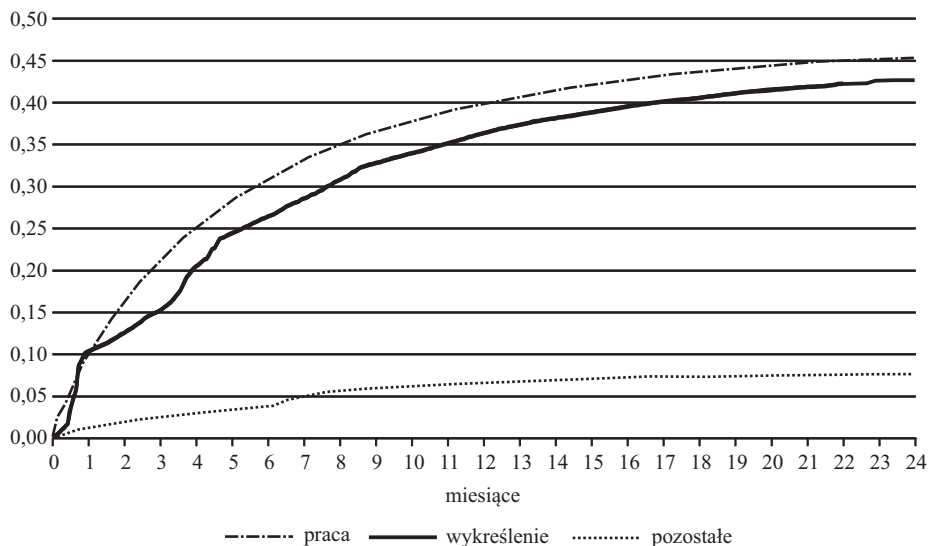
Wykr. 2. DOPEŁNIENIA DO JEDNOŚCI ESTYMATORÓW KAPLANA-MEIERA (1-KME)



W drugim etapie analizy wyznaczono estymatory skumulowanej częstości występowania (CIF_k) dla trzech grup powodów wyrejestrowania z urzędu pracy ($k = 1, 2, 3$), określone wzorem (6). W tym przypadku przyjęto trzy rodzaje zdarzeń kończących obserwację: podjęcie pracy, wykreślenie i pozostałe, będące zdarzeniami konkurującymi. Danymi cenzurowanymi są obserwacje, które nie zakończyły się przed końcem 2014 r. (powód cenzurowania 1). Prawdopodobieństwo wyrejestrowań z powodu podjęcia pracy było wyższe od prawdopodobieństwa wykreślenia od drugiego miesiąca od momentu zarejestrowania (wykr. 3). Wpływ pozostałych powodów był marginalny, a prawdopodobieństwo ich wystąpienia nie przekraczało 0,0765.

Wykr. 3. ESTYMATORY SKUMULOWANEJ FUNKCJI CZĘSTOŚCI (CIF_k)

prawdopodobieństwo
wyrejestrowania



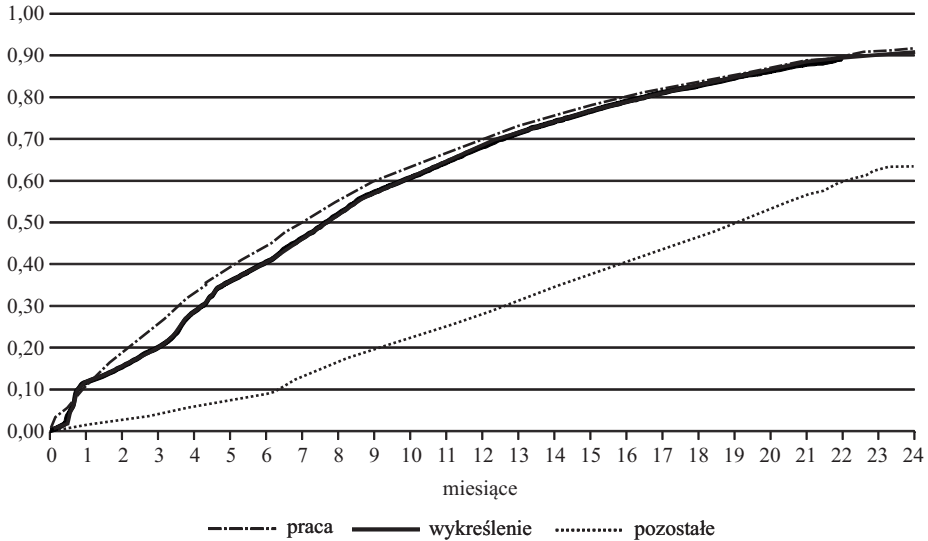
Źródło: jak przy wyk. 2.

Trzeci etap analizy polegał na wyznaczeniu estymatorów skumulowanego warunkowego prawdopodobieństwa wystąpienia zdarzenia (CPC_k). W tym przypadku, podobnie jak w drugim etapie analizy, uwzględniono trzy grupy powodów wyrejestrowania z urzędu pracy ($k = 1, 2, 3$), które stanowią zdarzenia konkurujące. Skorzystano ze wzoru (10). Obserwacje niezakończone przed końcem 2014 r. przyjęto za cenzurowane (powód cenzurowania 1). Wprawdzie od początku drugiego miesiąca od zarejestrowania prawdopodobieństwo warunkowe podjęcia pracy było wyższe od prawdopodobieństwa warunkowego wykreślenia, jednak od trzynastego miesiąca wartości obu estymatorów zbliżyły się do siebie (wykr. 4). Prawdopodobieństwo warunkowe wyrejestrowania z powodu

pozostałych przyczyn było mniejsze niż w przypadku podjęcia pracy i wykreślenia, jednak stosunkowo duże i na koniec obserwacji przyjęło wartość 0,634.

Wykr. 4. ESTYMATORY SKUMULOWANEGO WARUNKOWEGO PRAWDOPODOBIEŃSTWA (CPC_k)

prawdopodobieństwo
wyrejestrowania



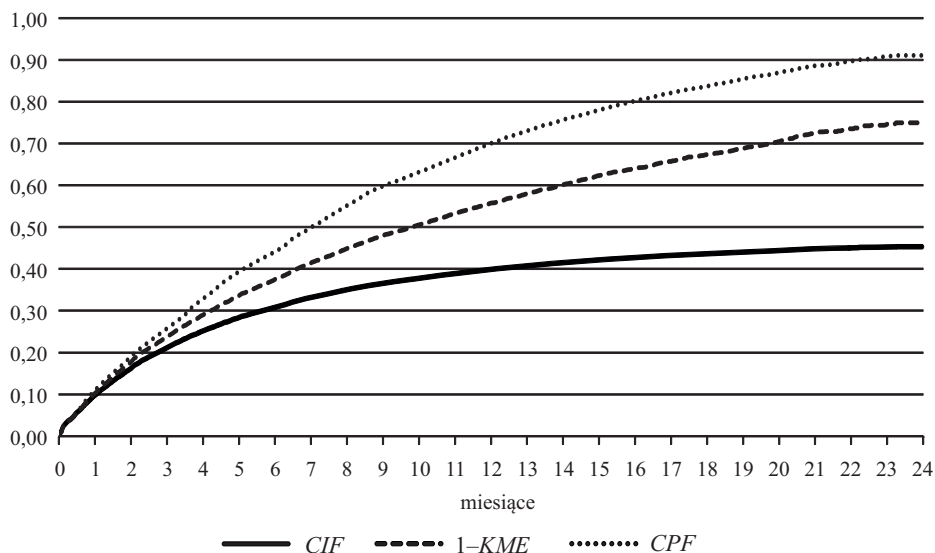
Źródło: jak przy wyk. 2.

Na wyk. 5 zestawiono trzy oszacowane estymatory dla zdarzenia podjęcie pracy. Jest to wizualizacja nierówności (11) opisującej zależność między nimi. Na uwagę zasługuje fakt regularnej krzywizny wszystkich trzech krzywych. Takiej własności nie miały estymatory prawdopodobieństwa wykreślenia i pozostałych powodów, dla których zauważalne są krótkotrwałe skoki wartości.

Szczegółowa analiza powodów wyrejestrowania pozwoliła na ocenę przyczyny zwiększonego wzrostu tych estymatorów. Dla grupy „pozostałe” skok wartości nastąpił w siódmym miesiącu od zarejestrowania. Był spowodowany zwiększoną liczbą wyrejestrowań z powodu przyznania zarejestrowanej osobie bezrobotnej prawa do pobierania świadczenia/zasiłku przedemerytalnego. Dla grupy „wykreślenie” skok w miesiącach pierwszym i czwartym był natomiast związany ze zwiększoną liczbą wyrejestrowań z powodu niestawienia się osoby bezrobotnej w PUP w wyznaczonym terminie. Jak już wcześniej wspomniano, przyjęcie uproszczenia, jakim jest zastosowanie w badaniu estymatora Kaplana-Meiera z dopełnieniem do 1, prowadzi do przeszacowania wartości ryzyka konkurującego CIF_k . W przypadku przedstawionej analizy jest ono bardzo duże i w 24. miesiącu, kończącym obserwację, jest większe o 84%.

Wykr. 5. PORÓWNANIE ESTYMATORÓW RYZYKA KONKURUJĄCEGO DLA ZDARZENIA „PODJĘCIE PRACY”

prawdopodobieństwo
wyrejestrowania



Źródło: opracowanie własne.

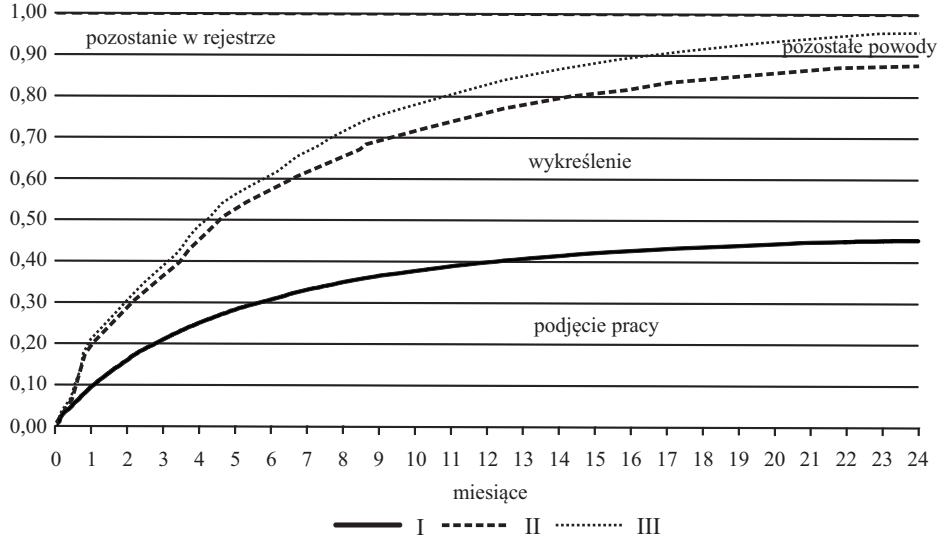
Podsumowanie

Wstępna analiza danych indywidualnych uzyskanych z PUP w Szczecinie wskazuje na to, że wykreślenie było najczęstszą przyczyną wyrejestrowania z urzędu. Dokładniejszych wyników dostarcza badanie przeprowadzone z wykorzystaniem trzech opisanych estymatorów. Estymator skumulowanej funkcji częstości (CIF_k) potwierdził, że praca rzeczywiście była najbardziej prawdopodobnym powodem wyrejestrowania. Pozostałe dwa wskazały, że prawdopodobieństwo wykreślenia było zbliżone do prawdopodobieństwa podjęcia pracy. Otrzymane wyniki sugerują, że chęć zatrudnienia nie była główną przyczyną zarejestrowania się w urzędzie. Pozostałe powody miały znaczenie marginalne, chociaż prawdopodobieństwo warunkowe ich zajęcia było dość duże.

Aby dokonać oceny wyników, zsumowano estymatory skumulowanej funkcji częstości (CIF_k) w sposób przedstawiony na wyk. 6. Gdyby nie było obserwacji cenzurowanych, tj. w przypadku opisanego badania wszystkie osoby z analizowanej kohorty zostałyby wyrejestrowane do końca 2014 r., to suma tych estymatorów dla wszystkich rodzajów ryzyka w 24. miesiącu byłaby równa 1, ale ponieważ jest równa 0,956, to prawdopodobieństwo pozostania w rejestrze w 24. miesiącu, czyli w stanie bezrobocia, osób z analizowanej kohorty jest równe 0,044.

Wykr. 6. ZSUMOWANE ESTYMATORY SKUMULOWANEJ FUNKCJI CZĘSTOŚCI (CIF_k)

skumulowane prawdopodobieństwo
wyrejestrowania



U w a g a. I — praca, II — praca+wykreślenie, III — praca+wykreślenie+pozostałe.
Ź r ó d ł o: opracowanie własne.

Odległości¹ między krzywymi wyrażają prawdopodobieństwo wyrejestrowania z określonego powodu. Odległość między krzywą III a prostą o wartości 1 jest równa prawdopodobieństwu pozostania w rejestrze. Jak widać, zmniejszało się ono w okresie badania. Po 12 miesiącach od zarejestrowania było równe 0,17. Jest to jednocześnie prawdopodobieństwo wejścia w stan długotrwałego bezrobocia. Prawdopodobieństwo pozostania w rejestrze urzędu po 24 miesiącach (koniec okresu obserwacji) było równe 0,04. Wyrejestrowanie z powodu podjęcia pracy po 12 miesiącach od zarejestrowania było równe 0,40, a po 24 miesiącach — 0,45. Analogicznie prawdopodobieństwo wykreślenia było równe 0,36 i 0,43.

Z przedstawionego badania wynika również ważna uwaga metodyczna. Jeżeli występują różne rodzaje zdarzeń kończących obserwację, to warto zastosować modele pozwalające oszacować ryzyko konkurujące. W przypadku czasu trwania bezrobocia rejestrowanego pozwalają one wyznaczyć prawdopodobieństwo podjęcia określonego rodzaju pracy oraz porównać je z innymi przyczynami wyrejestrowania. Potraktowanie innych zdarzeń kończących obserwację jako cenzurowanych może prowadzić do znacznego przeszacowania funkcji skumulowanej częstości.

dr hab. Beata Bieszk-Stolorz — Uniwersytet Szczeciński

¹ Odległość między krzywymi jest mierzona według metryki $d((x_1, y_1), (x_2, y_2)) = |y_2 - y_1|$.

LITERATURA

- Aly, E.A.A., Kochar, S., McKeague, E. (1994). Some Tests for Comparing Cumulative Incidence Functions and Cause-Specific Hazard Rates. *Journal of the American Statistical Association*, Vol. 89, No. 427.
- Bieszk-Stolorz, B. (2013). *Analiza historii zdarzeń w badaniu bezrobocia*. Szczecin: Volumina.pl Daniel Krzanowski.
- Bieszk-Stolorz, B., Markowicz, I. (2012a). *Modele regresji Coxa w analizie bezrobocia*. Warszawa: CeDeWu.
- Bieszk-Stolorz, B., Markowicz, I. (2012b). Wykorzystanie wielomianowego modelu logitowego do oceny szansy podjęcia pracy przez bezrobotnych. W: K. Jajuga, M. Walesiak (red.), *Taksonomia 19. Klasyfikacja i analiza danych — teoria i zastosowania. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, nr 242.
- Bryant, J., Dignam, J.J. (2004). Semiparametric models for cumulative incidence functions. *Biometrics*, Vol. 60, No. 1.
- Crowder, M. (1994). Identifiability Crises in Competing Risks. *International Statistical Review*, Vol. 62, No. 3.
- Crowder, M. (1996). On assessing independence of competing risks when failure times are discrete. *Lifetime Data Analysis*, Vol. 2, No. 2.
- Crowder, M. (1997). A test for independence of competing risks with discrete failure times. *Lifetime Data Analysis*, Vol. 3, No. 3.
- Gooley, T.A., Leisenring, W., Crowley, J., Storer, B.E. (1999). Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Statistics in Medicine*, Vol. 18, No. 6.
- Kalbfleisch, J.D., Prentice, R.L. (2002). *The Statistical Analysis of Failure Time Data. Second Edition*. Hoboken: John Wiley & Sons.
- Kaplan, E.L., Meier, P. (1958). Non-parametric estimation from incomplete observations. *Journal of American Statistical Association*, Vol. 53.
- Klein, J.P., Bajorunaite, R. (2004). *Inference for Competing Risks*. W: N. Balakrishnan, C.R. Rao (red.), *Handbook of Statistics: Advances in Survival Analysis*, Vol. 23. Elsevier.
- Klein, J.P., Moeschberger, M.L. (1984). Asymptotic bias of the product limit estimator under dependent competing risks. *Indian Journal of Productivity, Reliability and Quality Control*, Vol. 9.
- Klein, J.P., Moeschberger, M.L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data. Second Edition*. Nowy Jork: Springer-Verlag.
- Kleinbaum, D., Klein, M. (2005). *Survival Analysis. A Self-Learning Text*. Springer.
- Landmesser, J.M. (2008a). *Aktywność ekonomiczna ludności: klasyfikacja osób za pomocą wielomianowych modeli logitowych oraz jej związek z modelami hazardu dla czasów trwania*. W: K. Jajuga, M. Walesiak (red.), *Taksonomia 15. Klasyfikacja i analiza danych — teoria i zastosowania. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, nr 7.
- Landmesser, J.M. (2008b). *Modele ryzyka konkurencyjnego dla czasu trwania czynności*. W: T. Trzaskalik (red.), *Modelowanie preferencji a ryzyko '08. Prace Naukowe Akademii Ekonomicznej w Katowicach*.
- Landmesser, J. (2013). *Wykorzystanie metod analizy czasu trwania do badania aktywności ekonomicznej ludności w Polsce*. Warszawa: Wydawnictwo SGGW.
- Markowicz, I. (2012). *Statystyczna analiza żywotności firm*. Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.
- Marubini, E., Valsecchi, M. (1995). *Analysing Survival Data from Clinical Trials and Observational Studies*. John Wiley & Sons.
- Matuszyk, A. (2015). *Zastosowanie analizy przetrwania w ocenie ryzyka kredytowego klientów indywidualnych*. Warszawa: CeDeWu.

- Pepe, M.S. (1991). Inference for Events with Dependent Risks in Multiple Endpoint Studies. *Journal of the American Statistical Association*, Vol. 86, No. 415.
- Pepe, M.S., Mori, M. (1993). Kaplan-Meier, marginal or conditional probability curves in summarizing competing risks failure time data? *Statistics in Medicine*, Vol. 12.
- Sączewska-Piotrowska, A. (2016). Dynamika ubóstwa w miejskich i wiejskich gospodarstwach domowych. *Wiadomości Statystyczne*, nr 7, s. 39—59.
- Sherif Bintu, N. (2008). *A comparison of Kaplan-Meier and cumulative incidence estimate in the presence or absence of competing risks in breast cancer data*. Master's Thesis, University of Pittsburgh.
- Wycinka, E. (2015). Modelowanie czasu do zaprzestania spłat rat kredytu lub wcześniejszej spłaty kredytu jako zdarzeń konkurujących. *Problemy Zarządzania*, Vol. 13, nr 3 (55), t. 2.

Summary. *The purpose of this article is to present selected methods of the survival analysis to evaluate the probability of leaving unemployment for the various types of competing risks. Complement to the unity of the Kaplan-Meier estimator, cumulative incidence function and cumulative conditional probability were used in the study. With these three estimators, the probability of deregistering caused by undertaking work, refusal and other causes were compared. The analysis was based on data from the Powiat Labour Office in Szczecin.*

Keywords: Kaplan-Meier estimator, cumulative incidence function, cumulative conditional probability, competing risks, unemployment.

STATYSTYKA W PRAKTYCE

Agnieszka MATULSKA-BACHURA

Metody pomiaru wskaźnika wolumenu usług na przykładzie działalności związanej z administracyjną obsługą biura

Streszczenie. *Celem artykułu jest opis metod opracowywania wskaźnika wolumenu usług, zwanego także wskaźnikiem produkcji usług (Index of Services Production — ISP), oraz ich zastosowania w praktyce polskiej. Przedstawiono wpływ, jaki na szeregi czasowe wskaźników produkcji usług administracyjnej obsługi biura ma sposób pomiaru wartości obrotów (według produktu i według działalności, zależnie od wielkości przedsiębiorstw). W opracowaniu wykorzystano cztery metody obliczeń ISP, a następnie porównano szeregi czasowe wskaźnika uzyskane dla okresu badawczego 2010—2014 przy uwzględnieniu zastosowanych podejść. Na podstawie wyników przeprowadzonych prac ustalono, że każda z metod obliczania wskaźnika wolumenu usług odpowiada na inne potrzeby informacyjne, dlatego też przy wyborze metody liczenia wskaźnika powinno się uwzględniać cel podejmowanych analiz.*

Słowa kluczowe: metody statystyczne, usługi.

JEL: C180, O140

Działalność usługowa przez długi okres pozostawała poza kręgiem zainteresowania statystyki, podczas gdy procesy i zjawiska zachodzące w innych dziedzinach gospodarki narodowej znajdowały odzwierciedlenie w wynikowych informacjach statystycznych. Pomiar realnej wartości świadczonych usług jest utrudniony ze względu na ich niematerialny charakter oraz na dynamiczne zmiany zachodzące na skutek postępującej globalizacji na rynku usług. Jedno-

cześciej rynek ten stanowi jeden z najbardziej heterogenicznych obszarów gospodarki. Obejmuje usługi świadczone zarówno na rzecz gospodarstw domowych (np. sprzątanie czy usługi fryzjerskie), jak również na rzecz przedsiębiorstw czy podmiotów instytucjonalnych (np. usługi architektoniczne, rachunkowo-księgowo itp.). Mają one różnorodny charakter także ze względu na stopień zaawansowania technologii wykorzystywanych przy ich świadczeniu oraz nakłady niezbędnej wiedzy.

Na skutek różnych uwarunkowań praktycznie każda usługa wymaga odmiennego podejścia przy pomiarze jej wolumenu. Metodologiczne aspekty badania działalności usługowej stanowią przedmiot rozważań statystyków na całym świecie. Usiłują oni uzyskać odpowiedź na pytanie, w jaki sposób należy mierzyć wartość świadczonych usług, aby uzyskane wyniki jak najtrafniej odzwierciedlały rzeczywistość gospodarczą. Osiągnięcie tego celu wymaga przeanalizowania zjawisk wpływających na funkcjonowanie rynku danej usługi. Czy badania powinny dotyczyć jedynie przedsiębiorstw usługowych, czy obejmować również podmioty z innych działów gospodarki świadczących usługi w ramach swojej drugorzędnej działalności? Jak powinno się uwzględniać w statystyce usługi wykonywane na własne potrzeby wewnątrz przedsiębiorstwa czy w ramach grupy przedsiębiorstw? Jak na realną wartość świadczonych usług wpływa outsourcing (zlecenie zadań dotychczas realizowanych w przedsiębiorstwie innym podmiotom poprzez wydzielenie ich ze struktury organizacyjnej) czy ich świadczenie w wymianie międzynarodowej?

W artykule podjęto próbę rozwiązania jednego z wymienionych problemów, mianowicie oceny, czy pomiar realnej wartości świadczonych usług powinien dotyczyć całej gospodarki, czy ograniczać się tylko do działalności przedsiębiorstw usługowych. Na przykładzie usług związanych z administracyjną obsługą biura (grupa 82.1 PKWiU) opisano wpływ, jaki na szeregi czasowe wskaźników wolumenu ma pomiar wartości obrotów według produktu, a także uwzględnienie wszystkich badanych przedsiębiorstw, niezależnie od ich wielkości mierzonej liczbą osób pracujących. Na wstępie przedstawiono definicję usług stanowiących przedmiot niniejszych rozważań. Następnie scharakteryzowano rynek omawianych usług. Główną część artykułu poświęcono omówieniu wybranych metod obliczania ISP oraz ocenie ich wpływu na uzyskane wyniki i ich interpretację.

CHARAKTERYSTYKA RYNKU USŁUG ZWIĄZANYCH Z ADMINISTRACYJNĄ OBSŁUGĄ BIURA

Działalność usługowa określona według PKD jako „Działalność związana z administracyjną obsługą biura, włączając działalność wspomagającą” (grupa 82.1 PKD) polega na dostarczaniu różnorodnych usług administracyjnej obsługi biura na podstawie umowy albo jednorazowego zlecenia. Są to także usługi świadczone na rzecz innych podmiotów jako rutynowe usługi wspierające dla

przedsiębiorstw i organizacji, które tradycyjnie wykonują je same. Podmioty zaklasyfikowane do tej grupy nie dostarczają personelu operacyjnego wykonującego cały proces biznesowy.

Grupa 82.1 składa się z dwóch klas — „Działalność usługowa związana z administracyjną obsługą biura” (82.11 PKD) i „Wykonywanie fotokopii, przygotowywanie dokumentów i pozostała specjalistyczna działalność wspomagająca prowadzenie biura” (82.19 PKD). Działalność usługowa związana z administracyjną obsługą biura polega na dostarczaniu różnorodnej kombinacji codziennych usług administracyjnej obsługi biura, takich jak: recepcja, planowanie finansowe, księgowanie rachunków i kont, obsługa personelu, nadawanie poczty itp. Wykonywanie fotokopii, przygotowywanie dokumentów i pozostała specjalistyczna działalność wspomagająca prowadzenie biura obejmuje: przygotowywanie dokumentów, ich edytowanie albo powielanie, przepisywanie i przetwarzanie tekstu, wspomagające usługi sekretarskie, pisanie listów albo sprawozdań, obsługę poczty elektronicznej, sporządzanie list adresowych i obsługę poczty tradycyjnej, tj. sortowanie, adresowanie itp. (bez dostarczania usług drukowania, takich jak np. drukowanie offsetowe). W przypadku kopiowania czy drukowania jest to działalność krótkookresowa.

W 2013 r. w Polsce funkcjonowało ok. 8,7 tys. przedsiębiorstw o podstawowym rodzaju działalności sklasyfikowanym w grupie 82.1 PKD. Stanowiły one 15,2% całkowitej liczby przedsiębiorstw z sekcji N „Działalność w zakresie usług administrowania i działalność wspierająca” i wytwarzały ok. 3% obrotów tej sekcji. W przedsiębiorstwach tych znalazło miejsca pracy 3,2% ogólnej liczby osób pracujących w sekcji N.

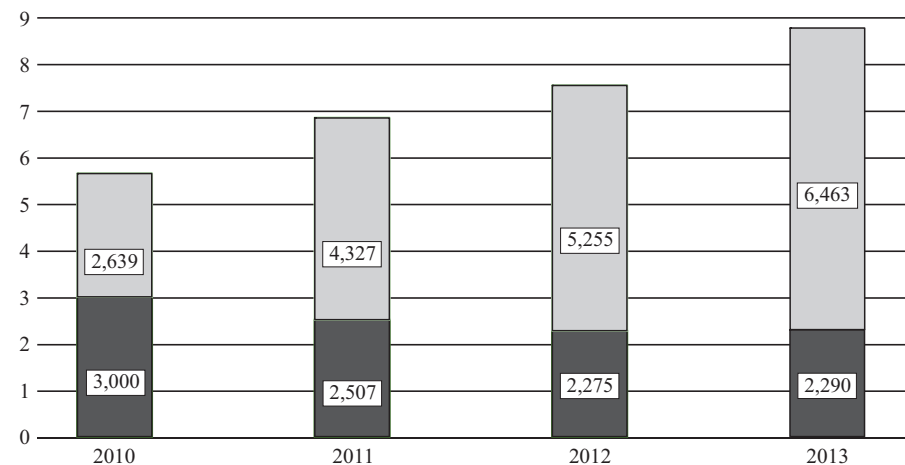
Wśród przedsiębiorstw prowadzących działalność związaną z administracyjną obsługą biura łącznie z działalnością wspomagającą dominującą rolę odgrywają przedsiębiorstwa świadczące usługi związane z administracyjną obsługą biura. W 2013 r. ich udział w liczbie przedsiębiorstw z grupy 82.1 PKD, liczbie pracujących oraz obrotach wyniósł ok. 75% całej grupy. Ponad 99% przedsiębiorstw stanowiły podmioty małe, o liczbie pracujących do 9 osób.

W 2013 r. liczba przedsiębiorstw prowadzących działalność związaną z administracyjną obsługą biura łącznie z usługami wspomagającymi wzrosła w porównaniu z 2010 r. o ok. 50% (wykr. 1), a liczba pracujących — o ok. 68%. Dynamiczny rozwój dotyczył jednak tylko przedsiębiorstw zaliczonych do klasy 82.11 PKD (wzrost liczby przedsiębiorstw i liczby pracujących o prawie 250%). W tym samym czasie liczba przedsiębiorstw działających w klasie 82.19 PKD zmniejszyła się o ok. 25%, a liczba pracujących — o ok. 17%. Spadek nastąpił w przedsiębiorstwach o liczbie pracujących do 9 osób.

Obserwowane zmiany liczby przedsiębiorstw i liczby osób pracujących miały odzwierciedlenie w wartości obrotów w omawianej grupie PKD, które w 2013 r. wyniosły 1,7 mld zł i były o ok. 70% wyższe niż w 2010 r. (wykr. 3). W klasie 82.11 obroty zwiększyły się o ok. 250% w stosunku do 2010 r., podczas gdy w klasie 82.19 spadły o ok. 18%.

Wykr. 1. LICZBA PRZEDSIĘBIORSTW W GRUPIE 82.1 PKD

tys.



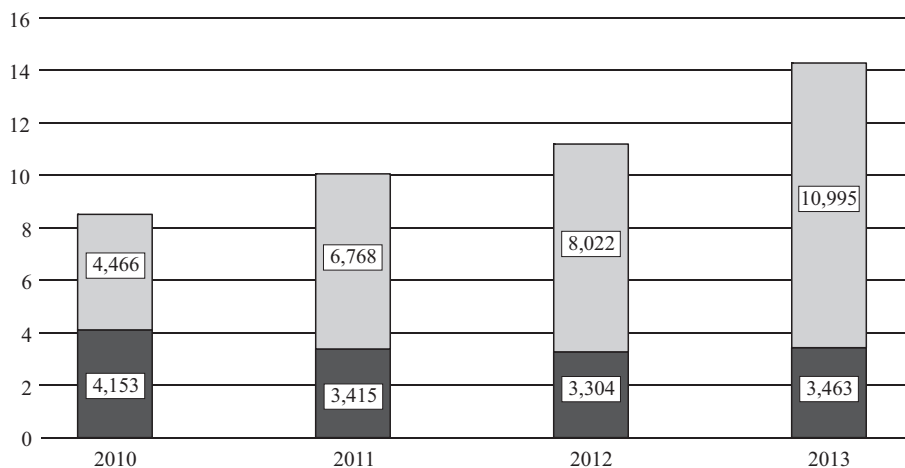
82.11. Działalność związana z administracyjną obsługą biura

82.19. Wykonywanie fotokopii, przygotowywanie dokumentów i pozostała specjalistyczna działalność wspomagająca prowadzenie biura

Źródło: na podstawie danych GUS.

Wykr. 2. LICZBA PRACUJĄCYCH W PRZEDSIĘBIORSTWACH Z GRUPY 82.1 PKD

tys.



82.11. Działalność związana z administracyjną obsługą biura

82.19. Wykonywanie fotokopii, przygotowywanie dokumentów i pozostała specjalistyczna działalność wspomagająca prowadzenie biura

Źródło: jak przy wykr. 1.

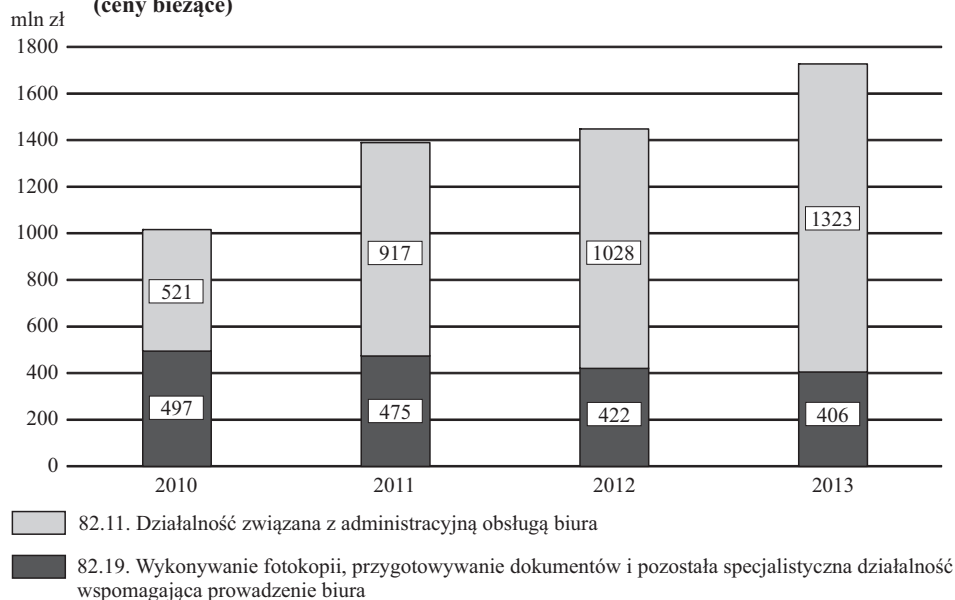
Należy podkreślić, że przedstawiona charakterystyka rynku usług związanych z administracyjną obsługą biura dotyczy przedsiębiorstw, które zgodnie z podstawowym rodzajem prowadzonej działalności są sklasyfikowane w grupie 82.1. Analizując dostępne dane na temat obrotów czy przychodów przedsiębiorstw prowadzących działalność w zakresie administracyjnej obsługi biura i działalności wspomagającej, należy wziąć pod uwagę, że nie odzwierciedlają one wielkości całkowitej produkcji tych usług, lecz wyniki działań przedsiębiorstw z grupy 82.1 PKD, które oprócz świadczenia usług stanowiących efekt ich podstawowej działalności wytwarzają również inne produkty. Jak zatem kształtowałyby się sytuacja na rynku omawianych usług, gdyby w pomiarach zastosować podejście na poziomie produktu, tj. ująć wyłącznie obroty uzyskiwane przez przedsiębiorstwa z tytułu świadczenia usług związanych z administracyjną obsługą biura? Do uzyskania odpowiedzi na to pytanie wykorzystano dane dotyczące przedsiębiorstw o liczbie pracujących powyżej 9 osób w zakresie przychodów ze sprzedaży produktów, towarów i materiałów według rodzaju prowadzonej działalności, które na potrzeby niniejszego opracowania traktowane są jako informacje odzwierciedlające produkcję według produktu.

Zmiana podejścia do pomiaru wartości świadczonych usług ma istotny wpływ na uzyskiwane wyniki. W przypadku przedsiębiorstw z grupy 82.1 o liczbie pracujących powyżej 9 osób roczne przychody netto ze sprzedaży produktów, towarów i materiałów wzrosły w okresie 2010—2013 o ok. 180%, podczas gdy przychody netto ze sprzedaży usług związanych z administracyjną obsługą biura zwiększyły się o 300% (wykr. 4). Całkowita wartość świadczonych usług związanych z administracyjną obsługą biura w gospodarce narodowej wzrosła w analizowanym okresie o 80%.

W 2013 r. przychody przedsiębiorstw niefinansowych o liczbie pracujących powyżej 9 osób uzyskiwane ze sprzedaży usług związanych z administracyjną obsługą biura wyniosły prawie 630 mln zł, z czego 42,5% stanowiły przychody ze sprzedaży usług świadczonych przez przedsiębiorstwa sklasyfikowane w grupie 82.1 PKD. Które zatem przedsiębiorstwa, oprócz tych z grupy 82.1, świadczą omawiane usługi? Obrazuje to wykr. 5.

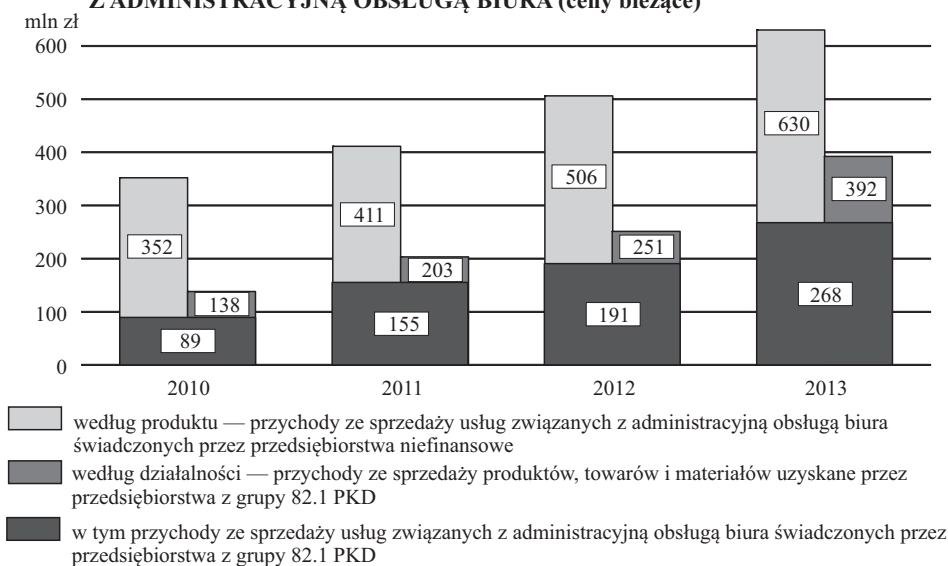
W 2013 r. istotny udział w rynku usług związanych z administracyjną obsługą biura (prawie 20%) miała produkcja przedsiębiorstw prowadzących działalność w zakresie wytwarzania i przetwarzania produktów rafinacji ropy naftowej (grupa 19.2 PKD). Przychody tych przedsiębiorstw z tytułu świadczenia omawianych usług miały jednak nieznaczny udział w całkowitych obrotach grupy 19.2. Ponadto, ze względu na specyfikę działalności prowadzonej przez te przedsiębiorstwa, można przypuszczać, że najprawdopodobniej świadczone je na rzecz innych przedsiębiorstw powiązanych. Usługi związane z administracyjną obsługą biura łącznie z usługami wspomagającymi stanowiły również przedmiot działalności przedsiębiorstw zajmujących się sprzedażą hurtową narzędzi technologii informacyjnej i komunikacyjnej (grupa 46.5 PKD) — prawie 7% rynku omawianych usług, sprzątaniem obiektów (grupa 81.2 PKD) — ponad 5% i w mniejszym zakresie innych.

Wykr. 3. ROCZNA WARTOŚĆ OBROTÓW PRZEDSIĘBIORSTW Z GRUPY 82.1 PKD (ceny bieżące)



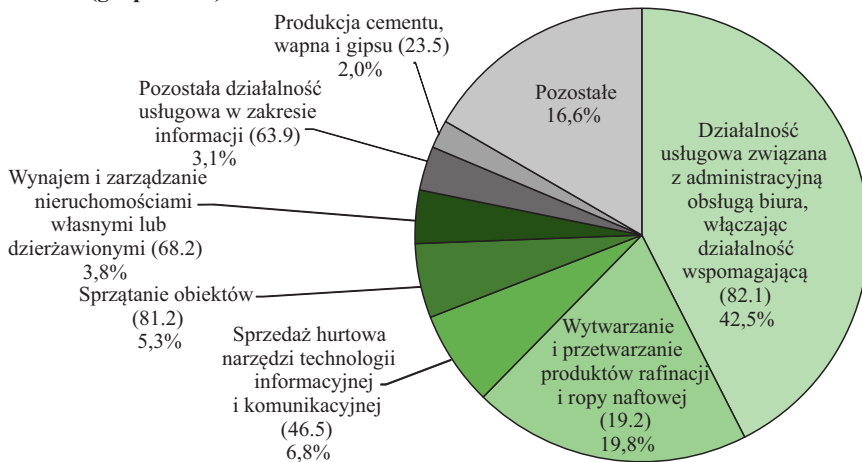
Źródło: jak przy wykr. 1.

Wykr. 4. PRZYCHODY NETTO PRZEDSIĘBIORSTW NIEFINANSOWYCH O LICZBIE PRACUJĄCYCH POWYŻEJ 9 OSÓB ZE SPRZEDAŻY PRODUKTÓW, TOWARÓW I MATERIAŁÓW NA RYNKU USŁUG ZWIĄZANYCH Z ADMINISTRACYJNĄ OBSŁUGĄ BIURA (ceny bieżące)



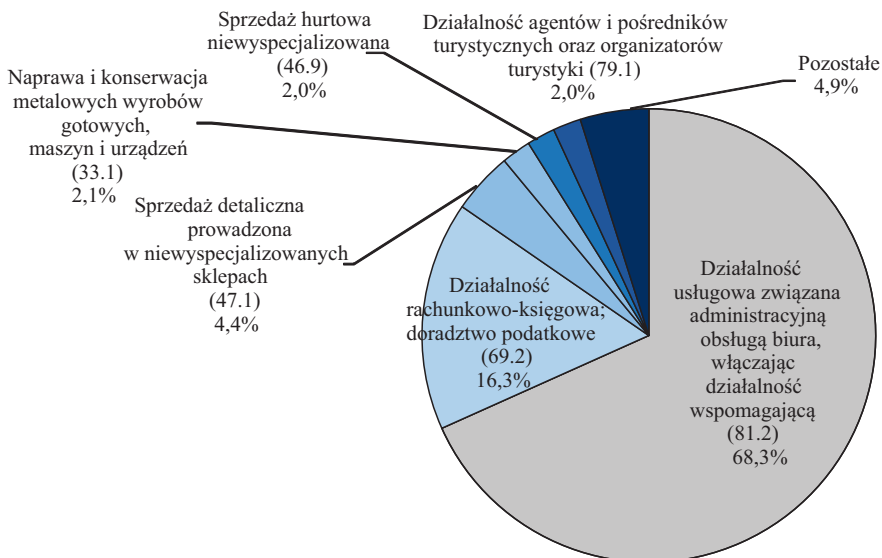
Źródło: obliczenia własne.

Wykr. 5. STRUKTURA PRZYCHODÓW NETTO ZE SPRZEDAŻY USŁUG ZWIĄZANYCH Z ADMINISTRACYJNĄ OBSŁUGĄ BIURA I USŁUG WSPOMAGAJĄCYCH ŚWIADCZONYCH PRZEZ PRZEDSIĘBIORSTWA NIEFINANSOWE O LICZBIE PRACUJĄCYCH POWYŻEJ 9 OSÓB WEDŁUG PODSTAWOWEGO RODZAJU PROWADZONEJ DZIAŁALNOŚCI (grupa PKD) W 2013 R.



Źródło: jak przy wykr. 4.

Wykr. 6. STRUKTURA PRZYCHODÓW NETTO ZE SPRZEDAŻY PRODUKTÓW PRZEDSIĘBIORSTW Z GRUPY PKD 82.1 O LICZBIE PRACUJĄCYCH POWYŻEJ 9 OSÓB WEDŁUG PRODUKTU (PKWiU)^a W 2013 R.



^a Ze względu na dostępność danych przyjęto założenie, że przychody ze sprzedaży produktów, towarów i materiałów wykazywane przez przedsiębiorstwa w ramach prowadzonych rodzajów działalności na poziomie klas PKD odzwierciedlają przychody ze sprzedaży produktów według PKWiU.

Źródło: jak przy wykr. 4.

Z kolei ponad 68% przychodów netto przedsiębiorstw o liczbie pracujących powyżej 9 osób, sklasyfikowanych według podstawowego rodzaju działalności w grupie 82.1, stanowiły w 2013 r. przychody ze sprzedaży usług świadczonych w ramach podstawowej działalności (wykr. 6). Ponad 16% przychodów przedsiębiorstwa te uzyskały z tytułu świadczenia usług rachunkowo-księgowych i audytu (klasa 69.20 PKWiU).

*BADANIA STATYSTYCZNE W ZAKRESIE DZIAŁALNOŚCI
ZWIĄZANEJ Z ADMINISTRACYJNĄ OBSŁUGĄ BIURA
ŁĄCZNIE Z DZIAŁALNOŚCIĄ WSPOMAGAJĄCĄ*

Na podstawie doświadczeń urzędów statystycznych w innych krajach można stwierdzić, że najczęściej wykorzystywanym miernikiem charakteryzującym wielkość produkcji przedsiębiorstw usługowych są obroty w cenach stałych. Dane o obrotach w cenach bieżących uzyskuje się w ramach strukturalnej statystyki przedsiębiorstw (SBS)¹ oraz krótkookresowej statystyki gospodarczej (STS). Dodatkowo informacje na temat działalności przedsiębiorstw z grupy 82.1 PKD można odnaleźć w statystyce rynku pracy, rachunkach narodowych czy statystyce międzynarodowego handlu usługami.

Zgodnie z wymogami SBS pod pojęciem obrotu rozumie się całkowite sumy zafakturowane przez jednostkę podczas okresu sprawozdawczego, co odpowiada sprzedaży rynkowej towarów i usług osobom trzecim. Sprzedaż obejmuje zarówno towary produkowane przez przedsiębiorstwo, jak i zakupione przez detalistę, a także grunt i inne nieruchomości posiadane w celu odsprzedaży. Świadczenie usług polega zazwyczaj na wykonywaniu przez przedsiębiorstwo ustalonego w umowie zadania przez ustalony okres. Wartość obrotów powiększana jest przez wszystkie opłaty i podatki od towarów lub usług fakturowanych przez jednostkę, z wyjątkiem podatków w rodzaju podatku od wartości dodanej (VAT), a także wszystkie inne opłaty (transport, opakowanie itd.) przenoszone na klienta, nawet jeżeli są one osobno wymienione na fakturze. W obrocie nie uwzględnia się natomiast wartości towarów produkowanych na potrzeby własne lub przeznaczanych na cele inwestycyjne. Odlicza się także obniżki cen, rabaty i upusty oraz wartość zwróconych opakowań. W Polsce informacje na temat obrotów opracowywane są na podstawie danych uzyskiwanych od przedsiębiorstw w zakresie różnych kategorii przychodów stanowiących pozycje rachunku zysków i strat².

¹ SBS (Structural Business Statistics) — jeden z działów statystyki wyodrębniony w Europejskim Systemie Statystycznym (ESS), który obejmuje: przemysł, budownictwo, handel i usługi. Udostępnia informacje na poziomie działalności według NACE, charakteryzujące działalność przedsiębiorstw w krajach Unii Europejskiej (UE). Wyniki dostępne są dla wszystkich krajów członkowskich, a jednolita metodologia ich opracowywania zapewnia porównywalność.

² Jest to uregulowane w rozporządzeniu Komisji (WE) nr 250/2009 z dnia 11 marca 2009 r. w sprawie wykonania rozporządzenia (WE) nr 295/2008 Parlamentu Europejskiego i Rady w odniesieniu do definicji cech, technicznego formatu przekazywania danych, wymogów dotyczących podwójnej sprawozdawczości dla NACE wersja 1.1 i NACE wersja 2 oraz odstępstw, które mają zostać przyznane w zakresie statystyk strukturalnych dotyczących przedsiębiorstw.

Należy podkreślić, że obecnie nie istnieją żadne prawne zobowiązania do pomiaru wskaźnika wolumenu usług. W pierwszym etapie modernizacji statystyki gospodarczej ESS planowane jest jednak rozszerzenie obowiązkowego zakresu zmiennych przekazywanych w ramach STS m.in. o ISP. Przyszłe wymogi związane z opracowywaniem miesięcznych ISP zawarte są w projekcie Rozporządzenia Ramowego Integrującego Statystyki Gospodarcze (Framework Regulation Integrating Business Statistics — FRIBS)³. Planowany termin wejścia tej regulacji w życie to 1 stycznia 2019 r. Niemniej jednak ISP obliczany według działalności spełnia tylko podstawowe wymogi informacyjne. Obserwacja zjawisk zachodzących na rynku usług skłania do przeprowadzenia dokładniejszych analiz. Umożliwiłoby je uwzględnienie w metodologii pomiaru ISP obrotów według produktu czy świadczenia usług w wymianie międzynarodowej. W przypadku niektórych usług kierunek zmian uzyskanych szeregów czasowych może być nawet przeciwny do wskazywanego przez wskaźniki obliczone za pomocą obecnie stosowanej metody.

W polskiej statystyce publicznej wybór metody obliczania wskaźnika wolumenu usług — według działalności czy według produktu — uwarunkowany jest głównie dostępnością danych. Informacje o przychodach dotyczą przedsiębiorstw, a nie produktów. Oznacza to, że odnotowywane wartości przychodów obejmują przychody nie tylko z tytułu świadczenia danej usługi, lecz także z tytułu wytwarzania innych produktów. Z drugiej strony wartość przychodów według działalności nie uwzględnia przychodów ze sprzedaży usług uzyskiwanych przez podmioty sklasyfikowane w innych grupach działalności PKD niż omawiana.

Z powodu braku możliwości uzyskania danych bezpośrednio od podmiotów gospodarczych obecnie nie planuje się regularnego obliczania wskaźników produkcji według produktu. Potrzeby informacyjne użytkowników skłaniają jednak do eksperymentalnego opracowywania metod obliczania wskaźnika produkcji według produktu przy wykorzystaniu istniejących źródeł danych. Efekty tych działań przedstawione zostaną w dalszej części artykułu.

Obecna metodologia obliczania wskaźnika wolumenu danej usługi według działalności wymaga zestawienia miesięcznych i rocznych wartości obrotów przedsiębiorstw prowadzących daną działalność gospodarczą w cenach bieżących oraz deflatora odpowiedniego dla danej działalności. Źródłem informacji na temat miesięcznych wartości obrotów w cenach bieżących są dane zbierane za pomocą formularza „Miesięczne sprawozdanie z działalności gospodarczej” (DG-1). Wypełniają go przedsiębiorstwa o liczbie pracujących powyżej 49 osób oraz

³ Prace nad tą regulacją rozpoczęto w Eurostacie w 2011 r. Celem FRIBS jest integracja ESS, umożliwiająca zwiększenie jej efektywności i elastyczności reagowania na zmieniające się potrzeby użytkowników oraz redukcja obciążenia respondentów, a także krajowych urzędów statystycznych. Przygotowywany jeden ramowy akt prawny, wraz z aktami delegowanymi i wykonawczymi uszczegółowującymi wymogi europejskie, ma objąć swym zakresem dotychczas regulowane osobnymi rozporządzeniami: rejestr statystyczny, STS, SBS, statystykę zagranicznych podmiotów zależnych (FATS), Listę Produktów PRODCOM, międzynarodowy handel towarami (ITGS), międzynarodowy handel usługami (ITSS), innowacje (ICT), badania i rozwój (R&D) oraz społeczeństwo informacyjne (ISS) w części dotyczącej przedsiębiorstw.

ok. 10-procentowa próba przedsiębiorstw o liczbie pracujących od 10 do 49 osób. Do obliczenia tej zmiennej wykorzystuje się dane o przychodach netto ze sprzedaży produktów oraz przychodach netto ze sprzedaży towarów i materiałów.

Informacje dotyczące rocznych wartości obrotów w cenach bieżących dla działalności związanej z administracyjną obsługą biura łącznie z działalnością wspomagającą opracowywane są na podstawie danych uzyskiwanych przez statystykę publiczną za pomocą następujących formularzy sprawozdawczych⁴:

- „Roczna ankieta przedsiębiorstw” (SP);
- „Badanie małych przedsiębiorstw” (SP-3);
- „Sprawozdanie z przychodów, kosztów i wyników finansowych oraz aktywów trwałych” (F-01/I-01);
- DG-1.

Za pomocą ankiety SP od przedsiębiorstw o liczbie pracujących powyżej 9 osób uzyskuje się informacje:

- z zakresu sprawozdań finansowych — bilansu oraz rachunku zysków i strat;
- inne charakteryzujące działalność gospodarczą, takie jak: liczba osób pracujących, rodzaj prowadzonej księgowości, udział kapitału zagranicznego, zmiany własnościowe itp.;
- na temat wartości przychodów ze sprzedaży produktów, towarów i materiałów według rodzajów prowadzonej działalności (pierwszo- i drugorzędnej) na poziomie klas PKD. Informacje te są istotne z punktu widzenia opracowywania danych według produktu. Przy założeniu, że klasa PKD odpowiada czwartemu poziomowi PKWiU (cztery znaki), dostępność danych dla przychodów na poziomie klas PKD umożliwi przejście na poziom produktu. W przypadku większości usług (oprócz usług związanych z obsługą działalności gospodarczej) jest to jedyne źródło danych.

Badanie SP-3 obejmuje wszystkie przedsiębiorstwa o liczbie pracujących do 9 osób (z pewnymi wyjątkami) prowadzące działalność gospodarczą w Polsce i zarejestrowane w REGON. Informacje zbiera się rocznie od ok. 5-procentowej próby przedsiębiorstw wylosowanych według ustalonych warstw (grupowania działalności wymagane w SBS, z podziałem na województwa). W badaniu pozyskuje się informacje na temat rodzaju prowadzonej księgowości, liczby osób pracujących, przychodów, kosztów (w tym wynagrodzeń), wartości aktywów trwałych i nakładów na inwestycje. Dane o wartościach obrotów opracowuje się na podstawie danych zbieranych w dziale 4 („Przychody i koszty”) formularza SP-3.

W przypadku działalności związanej z administracyjną obsługą biura łącznie z działalnością wspomagającą roczne dane o obrotach opracowuje się w SBS na poziomie klas 82.11 i 82.19 PKD. Głównym źródłem danych są informacje z SP oraz SP-3. W przypadku niezyskania danych od respondentów wykorzystuje się źródła alternatywne, w szczególności sprawozdania F-01/I-01 i DG-1. Jeśli roczne dane o obrotach poszczególnych podmiotów nie są dostępne w badaniach

⁴ Wzory formularzy są dostępne na stronie <http://form.stat.gov.pl/formularze/formularze.htm>. Badania realizowane przy ich wykorzystaniu obejmują w zasadzie wszystkie podmioty gospodarcze działające w gospodarce narodowej; różnią się częstotliwością zbierania danych i ich szczegółowością.

statystyki publicznej, korzysta się z danych z systemu informacyjnego Ministerstwa Finansów o płatnikach VAT.

Ze względu na brak wskaźnika cen producenta dla działalności sklasyfikowanej w grupie 82.1 PKD, na użytek niniejszego opracowania jako deflator zastosowano wskaźnik cen dóbr i usług konsumpcyjnych ogółem. Zróżnicowany charakter działalności usługowych zaliczanych do administracyjnej obsługi biura utrudnia wskazanie wskaźnika cen producentów, który odzwierciedlałby zmiany w poziomie cen tych usług i mógłby być wykorzystywany do przeliczania wartości w cenach bieżących na wartości w cenach stałych. Trwają prace w zakresie statystyki cen producentów usług, których celem jest udoskonalenie obecnie prowadzonego badania oraz rozszerzenie go o kolejne rodzaje działalności usługowej⁵, m.in. o usługi związane z administracyjną obsługą biura. W ich efekcie docelowo w schemacie obliczania ISP wykorzystany zostanie wskaźnik cen producentów usług.

METODY OBLICZANIA MIESIĘCZNEGO WSKAŹNIKA PRODUKCJI DLA DZIAŁALNOŚCI ZWIĄZANEJ Z ADMINISTRACYJNĄ OBSŁUGĄ BIURA ŁĄCZNIE Z DZIAŁALNOŚCIĄ WSPOMAGAJĄCĄ

Na przykładzie usług związanych z administracyjną obsługą biura przedstawione zostaną cztery metody obliczania ISP, z których każda uwzględni wybrane aspekty pomiaru rynku omawianych usług według:

- A — działalności dla przedsiębiorstw o liczbie pracujących powyżej 9 osób;
- B — działalności dla przedsiębiorstw ogółem;
- C — produktu dla przedsiębiorstw o liczbie pracujących powyżej 9 osób;
- D — produktu dla przedsiębiorstw ogółem.

Dla każdej metody opisano schemat eksperymentalnego obliczenia ISP oraz uzyskano szeregi miesięcznych wskaźników dla lat 2010—2014. Niezależnie od stopnia szczegółowości wybranej metody obliczania ISP w jej schemacie można wyróżnić trzy główne etapy działań:

- obliczenie miesięcznych wartości obrotów w cenach bieżących w poszczególnych miesiącach sprawozdawczych;
- przeliczenie miesięcznych wartości obrotów w cenach bieżących na wartości w cenach stałych przy użyciu wskaźnika cen dóbr i usług konsumpcyjnych ogółem jako deflatora;
- obliczenie wskaźników zmian wartości obrotów w cenach stałych w poszczególnych miesiącach sprawozdawczych w stosunku do średniorocznej wartości obrotów w cenach stałych w roku bazowym t ($t=100$).

Wszelkie modyfikacje metodologiczne badania wolumenu usług polegają na uszczegóławianiu wymienionych działań tak, aby uzyskane szeregi czasowe jak najtrafniej obrazowały rzeczywistość gospodarczą.

⁵ Obecnie — od pierwszego kwartału 2016 r. — w badaniu cen producentów usług zbierane są dane na temat cen dla działalności związanej z administracyjną obsługą biura łącznie z działalnością wspomagającą. Po uzyskaniu wystarczająco długich szeregów czasowych rozpoczęte zostanie regularne obliczanie wskaźników cen.

Pierwszą metodę (A) należy traktować jako wyjściową, która spełnia minimalne wymogi odnośnie do wskaźnika produkcji podane w *Compilation manual for an index of service production*⁶ (OECD, 2007) oraz *Eurostat-OECD Methodological Guide for Developing Producer Price Indices for Services*⁷ (Eurostat, OECD, 2014). Kolejne schematy obliczania wskaźnika stanowią rozwinięcie metody A; uwzględniają pomiar wartości produkcji wszystkich badanych podmiotów i/lub według produktu.

Do eksperymentalnego obliczenia szeregów czasowych ISP wykorzystano dane o:

- miesięcznych wartościach obrotów w cenach bieżących — o sumie przychodów netto ze sprzedaży produktów oraz przychodach netto ze sprzedaży towarów i materiałów — z formularza DG-1 wykorzystywanego w badaniu „Ocena bieżącej działalności gospodarczej przedsiębiorstw” (dotyczy podmiotów o liczbie pracujących powyżej 9 osób);
- rocznych wartościach przychodów netto ze sprzedaży produktów, towarów i materiałów⁸ w cenach bieżących według rodzaju działalności na poziomie klas PKD dla przedsiębiorstw o liczbie pracujących powyżej 9 osób — z formularza SP;
- rocznych wartościach obrotów — z zestawów danych w podziale według klas wielkości przedsiębiorstw mierzonych liczbą osób pracujących, opracowywanych zgodnie z wymogami SBS.

Przyjęto następujące założenia:

- wartości przychodów netto ze sprzedaży produktów, towarów i materiałów według klas działalności uzyskiwane z formularza SP ze względu na spójność definicji mogą być wykorzystane jako przybliżenie obrotów według produktu;
- udział przychodów netto ze sprzedaży danego produktu w przychodach netto ze sprzedaży produktów, towarów i materiałów jest niezmienny w ciągu całego roku sprawozdawczego;
- przedsiębiorstwa o liczbie pracujących do 9 osób prowadzą tylko jeden rodzaj działalności gospodarczej, wskazywanej jako działalność podstawowa; oznacza to, że całkowite obroty przedsiębiorstw sklasyfikowanych w grupie 82.1 PKD generowane są z tytułu świadczenia usług.

W każdej metodzie do uzyskania wartości obrotów w cenach stałych wykorzystano wskaźnik cen dóbr i usług konsumpcyjnych ogółem pochodzący ze statystyki cen detalicznych. ISP obliczano jako zmianę wartości obrotów w ce-

⁶ Podręcznik z zakresu wskaźnika produkcji usług.

⁷ Podręcznik metodologiczny z zakresu rozwoju wskaźników cen producentów usług.

⁸ Roczne wartości przychodów netto ze sprzedaży produktów, towarów i materiałów — pozycje z formularza SP (cz. I, dz. 2 „Rodzaje działalności przedsiębiorstwa”); dla każdego rodzaju działalności, z którego uzyskiwane są przychody, jednostka sprawozdawcza opisuje działalność, podaje symbol PKD na poziomie klasy oraz wykazuje wartość przychodów netto ze sprzedaży. Suma przychodów powinna odpowiadać sumie wartości podanych w cz. II, dz. 2 „Rachunek zysków i strat” formularza, tj. sumie przychodów netto ze sprzedaży produktów, przychodów netto ze sprzedaży towarów i materiałów oraz kosztów wytworzenia produktów na własne potrzeby, pomniejszonych o koszt wytworzenia produktów przekazanych do własnych sklepów [suma=02+07+(05-06)].

nach stałych w danym miesiącu sprawozdawczym w odniesieniu do średniorocznej wartości obrotów w cenach stałych w roku 2010 (2010=100).

Zależnie od stopnia szczegółowości metody schemat obliczania ISP składa się z różnej liczby działań elementarnych (etapów) — od 3 do 13. Przedstawiono w nim wszystkie możliwe działania elementarne (oznaczone liczbami rzymskimi) i działania wykorzystywane w konkretnych metodach (ich kolejność określono liczbami arabskimi). Ciemne pole oznacza niewykonywanie działania w danej metodzie.

SCHEMAT OBLICZANIA ISP WEDŁUG METOD A, B, C I D DLA LAT 2010—2014

Etapy działania	Kolejność składowych działań elementarnych w danej metodzie			
	A	B	C	D
I. Zestawienie miesięcznych wartości obrotów ($ONI_{R,m}$) w cenach bieżących dla przedsiębiorstw o liczbie pracujących powyżej 9 osób zaklasyfikowanych według PKD do grupy 82.1, ze sprawozdania DG-1 ^a (zmienna obrotów w cenach bieżących ($On_{xb_{R,m}}$))	1	1	1	1
II. Zestawienie rocznych danych o przychodach netto ze sprzedaży usług związanych z administracyjną obsługą biura w cenach bieżących (PP_R) według grup PKD z formularza SP dla przedsiębiorstw o liczbie pracujących powyżej 9 osób $PKD_S, \text{ gdzie } \bigwedge_{SYM_{KL}=82.11 \vee 82.19} PP_R = \sum \text{ CZĘŚĆ I, DZIAŁ 2, RUBR. } 3_R$ gdzie: PKD_S — podstawowy rodzaj działalności według grupy PKD, SYM_{KL} — klasa PKWiU wykazywana w formularzu SP w części I, dziale 2, rubryce 2, R — rok sprawozdawczy			2	2
III. Zestawienie rocznych wartości obrotów w cenach bieżących obliczanych w ramach SBS ($SBS\ 12\ 11\ 0$) dla grupy 82.1 w podziale na klasy PKD oraz klasy wielkości mierzone liczbą osób pracujących		2		3
IV. Ustalenie na podstawie zestawienia opracowanego w etapie III struktury przychodów netto ze sprzedaży usług związanych z administracyjną obsługą biura (klasy 82.11 i 82.19 PKD) według działalności podstawowej i wytypowanie rodzajów działalności według PKD, w ramach których udział przychodów netto z tytułu świadczenia usług związanych z administracyjną obsługą biura w całkowitych przychodach netto ze sprzedaży tych usług jest większy niż 1%			3	4
V. Zestawienie wartości przychodów netto ze sprzedaży i zrównanych z nimi (PO_R) dla działalności wytypowanych w etapie IV $\bigwedge_{PKD_S \geq 1\%} PO_R = \sum \text{ CZĘŚĆ II, DZIAŁ 2, W. 01, RUBR. } 1_R$ gdzie $PKD_S \geq 1\%$ — grupa PKD, której udział przychodów netto z tytułu świadczenia usług związanych z administracyjną obsługą biura w całkowitych przychodach netto ze sprzedaży tych usług jest równy bądź większy niż 1%			4	5

^a Do obliczenia tej zmiennej wykorzystuje się dane z formularza DG-1 dotyczące przychodów netto ze sprzedaży produktów oraz przychodów netto ze sprzedaży towarów i materiałów.

SCHEMAT OBLICZANIA ISP WEDŁUG METOD A, B, C I D DLA LAT 2010—2014 (cd.)

Etapy działania	Kolejność składowych działań elementarnych w danej metodzie			
	A	B	C	D
<p>VI. Obliczenie dla każdej grupy PKD wytypowanej w etapie IV udziału przychodów netto z tytułu świadczenia usług związanych z administracyjną obsługą biura ($U82I_R$) w przychodach netto ze sprzedaży i zrównanych z nimi dla tych działalności</p> $\bigwedge_{PKD_S \geq 1\%} U82I_R = \frac{PP_R}{PO_R}$			5	6
<p>VII. Ustalenie dla każdego roku sprawozdawczego na podstawie danych z zestawienia w etapie I udziału obrotów $On_xb_{R,m}$ przedsiębiorstw sklasyfikowanych w grupie 82.1 PKD w poszczególnych miesiącach sprawozdawczych w obrotach rocznych ($UON82I_{R,m}$)</p> $UON82I_{R,m} = \frac{On_xb_{R,m}}{\sum_{m=1}^{12} On_xb_{R,m}}$		3		
<p>VIII. Zestawienie dla każdego roku z okresu 2010—2014 miesięcznych wartości obrotów w cenach bieżących dla przedsiębiorstw zaklasyfikowanych według PKD do rodzajów działalności wytypowanych w etapie IV na podstawie danych z formularza DG-1 ($ONOTHER_{R,m}$)</p> $\bigwedge_{PKD_S \geq 1\%} ONOTHER_{R,m} = On_xb_{R,m}$			6	7
<p>IX. Zestawienie dla każdego roku z okresu 2010—2014 miesięcznych wartości obrotów w cenach bieżących z tytułu świadczenia usług związanych z administracyjną obsługą biura ($ONP_{R,m}$) jako sumy iloczynu miesięcznych wartości obrotów obliczonych w etapie VIII ($ONOTHER_{R,m}$) oraz udziałów obliczonych w etapie VI ($U82I_R$)</p> $\bigwedge_{PKD_S \geq 1\%} ONP_{R,m} = \sum U82I_R \times ONOTHER_{R,m}$			7	8
<p>X. Obliczenie miesięcznych wartości obrotów w cenach bieżących dla przedsiębiorstw o liczbie pracujących do 9 osób, sklasyfikowanych w grupie 82.1 PKD ($ONO_9_{R,m}$) jako iloczynu udziałów określonych w etapie VII i rocznych wartości obrotów przedsiębiorstw o liczbie pracujących do 9 osób ($LP \leq 9$) zestawionych w etapie III. Dla przedsiębiorstw z grupy 82.1 PKD i ($LP \leq 9$)</p> $ONO_9_{R,m} = UON82I_{R,m} \times SBS_{12\ 10\ 0\ R, PKD\ 82.1, LP \leq 9}$ <p>gdzie L oznacza liczbę pracujących, a SBS 12 11 0 — obroty z SBS</p>		4		9
<p>XI. Zestawienie dla każdego roku z okresu 2010—2014 miesięcznych wartości obrotów w cenach bieżących ($ONITOT_{R,m}$) generowanych przez wszystkie przedsiębiorstwa sklasyfikowane w grupie 82.1 PKD jako sumy miesięcznych obrotów przedsiębiorstw o liczbie pracujących powyżej 9 osób (etap I) oraz przedsiębiorstw o liczbie pracujących do 9 osób (etap X). Dla przedsiębiorstw z grupy 82.1 PKD</p> $ONITOT_{R,m} = ONI_{R,m} + ONO_9_{R,m}$		5		10

SCHEMAT OBLICZANIA ISP WEDŁUG METOD A, B, C I D DLA LAT 2010—2014 (dok.)

Etapy działania	Kolejność składowych działań elementarnych w danej metodzie			
	A	B	C	D
<p>XII. Zestawienie dla każdego roku z okresu 2010—2014 miesięcznych wartości obrotów w cenach bieżących generowanych przez przedsiębiorstwa ogółem z tytułu świadczenia usług związanych z administracyjną obsługą biura ($ONPTOT_{R,m}$) jako sumy miesięcznych wartości obrotów obliczonych w etapie IX i X</p> $ONPTOT_{R,m} = ONP_{R,m} + ONO_{9R,m}$				11
<p>XIII. Przeliczenie uzyskanych wartości obrotów w cenach bieżących na wartości w cenach stałych przy wykorzystaniu wskaźnika cen dóbr i usług konsumpcyjnych (CPI). Dla:</p> $ONI_{R,m} : ON\tilde{I}_{R,m} = ONI_{R,m}/CPI_{R,m}$ $ONP_{R,m} : ON\tilde{P}_{R,m} = ONP_{R,m}/CPI_{R,m}$ $ONITOT_{R,m} : ON\tilde{I}TOT_{R,m} = ONITOT_{R,m}/CPI_{R,m}$ $ONPTOT_{R,m} : ON\tilde{I}TOT_{R,m} = ONPTOT_{R,m}/CPI_{R,m}$	2	6	8	12
<p>XIV. Obliczenie ISP jako zmiany wartości obrotów w cenach stałych w danym miesiącu sprawozdawczym w odniesieniu do średniorocznej wartości obrotów w cenach stałych w roku 2010 (2010=100). Dla każdego roku z okresu 2010—2014 przy $R_0=2010$:</p> $ONI_{R,m} : ISP_ONI_{R,m} = \frac{ON\tilde{I}_{R,m}}{\left(\sum_{j=1}^{12} ON\tilde{I}_{R_0,j}\right)/12} \times 100$ $ONP_{R,m} : ISP_ONP_{R,m} = \frac{ON\tilde{P}_{R,m}}{\left(\sum_{j=1}^{12} ON\tilde{P}_{R_0,j}\right)/12} \times 100$ $ONITOT_{R,m} : ISP_ONITOT_{R,m} = \frac{ON\tilde{I}TOT_{R,m}}{\left(\sum_{j=1}^{12} ON\tilde{I}TOT_{R_0,j}\right)/12} \times 100$ $ONPTOT_{R,m} : ISP_ONPTOT_{R,m} = \frac{ON\tilde{I}TOT_{R,m}}{\left(\sum_{j=1}^{12} ON\tilde{I}TOT_{R_0,j}\right)/12} \times 100$	3	7	9	13

Źródło: opracowanie własne.

Pierwszym etapem wszystkich opisanych metod jest zestawienie — uzyskanych ze sprawozdania DG-1 — miesięcznych wartości obrotów w cenach bieżących przedsiębiorstw, o liczbie pracujących powyżej 9 osób sklasyfikowanych w grupie 82.1 PKD. Również dwa ostatnie etapy są wspólne dla wszystkich

metod. Przedostatni polega na przeliczeniu wartości obrotów w cenach bieżących na wartości w cenach stałych, a ostatni — na obliczeniu ISP. Wszelkie modyfikacje wprowadzane do schematów obliczania wskaźników służą uzyskaniu informacji o miesięcznych obrotach w cenach bieżących według produktu i/lub wszystkich przedsiębiorstwach niezależnie od ich wielkości mierzonej liczbą osób pracujących.

Metoda A — ISP liczony według działalności dla przedsiębiorstw średnich i dużych (o liczbie pracujących powyżej 9 osób)

Jest to wyjściowa metoda obliczania danych, zapewniająca wypełnienie zobowiązań przewidywanych w projekcie FRIBS. Wśród jej zalet należy wskazać fakt, że wykorzystywane w niej dane są dostępne w statystyce publicznej i nie trzeba dokonywać dodatkowych przeliczeń w celu uzyskania danych wejściowych. Umożliwia to opracowywanie wskaźników w terminach wskazanych we FRIBS.

Uzyskany ISP przedstawia zmiany wolumenu usług dla przedsiębiorstw o liczbie pracujących powyżej 9 osób prowadzących jako podstawową działalność sklasyfikowaną w grupie 82.1 PKD. Odzwierciedla zmiany wartości obrotów uzyskiwanych przez te przedsiębiorstwa z tytułu świadczenia wszystkich usług czy sprzedaży dóbr, nie uwzględnia natomiast zmian wartości obrotów z tytułu świadczenia usług związanych z administracyjną obsługą biura ani wartości obrotów ze sprzedaży tych usług świadczonych przez przedsiębiorstwa w ramach ich dodatkowej działalności (sklasyfikowanej w innej grupie PKD).

Metoda B — ISP liczony według działalności dla wszystkich przedsiębiorstw

Ta metoda również dotyczy przedsiębiorstw, których podstawowym rodzajem działalności jest działalność związana z administracyjną obsługą biura, z tym że obejmuje obroty przedsiębiorstw ogółem (brane są pod uwagę zarówno przedsiębiorstwa o liczbie pracujących powyżej 9 osób, jak i do 9 osób).

ISP przedstawia zmiany wartości obrotów wszystkich przedsiębiorstw prowadzących działalność w zakresie administracyjnej obsługi biura. Wartości obrotów z tytułu usług związanych z administracyjną obsługą biura świadczonych przez przedsiębiorstwa o podstawowej działalności innej niż 82.1 PKD nie są natomiast uwzględniane.

Obliczenie wskaźnika wymaga roszacowania danych rocznych dla przedsiębiorstw o liczbie pracujących do 9 osób na poszczególne miesiące sprawozdawcze. Problem stanowi termin dostępności danych rocznych dla podmiotów o liczbie pracujących do 9 osób, co uniemożliwia wykonanie obliczeń w terminie wymaganym przez Eurostat. Rozwiązaniem jest przekazywanie danych na temat przedsiębiorstw zatrudniających powyżej 9 osób, a następnie rewizja wskaźników po uzyskaniu dostępu do danych rocznych dla mniejszych przedsiębiorstw.

Metody C i D

Kolejne dwie metody obliczania ISP uwzględniają pomiar wartości świadczonych usług według produktu, tj. obrotów ze sprzedaży usług związanych z administracyjną obsługą biura przez przedsiębiorstwa niefinansowe niezależnie od podstawowego rodzaju prowadzonej działalności gospodarczej. Przesłanką dla takiego podejścia jest fakt, że w 2013 r. niemal 60% tych obrotów generowały przedsiębiorstwa prowadzące jako podstawową działalność inną niż sklasyfikowana w grupie 82.1 PKD (patrz poniższa tablica).

**UDZIAŁ PRZYCHODÓW NETTO ZE SPRZEDAŻY USŁUG ZWIĄZANYCH
Z ADMINISTRACYJNĄ OBSŁUGĄ BIURA W PRZYCHODACH ZE SPRZEDAŻY PRODUKTÓW,
TOWARÓW I MATERIAŁÓW WEDŁUG PODSTAWOWEGO RODZAJU DZIAŁALNOŚCI
PRZEDSIĘBIORSTW ŚWIADCZĄCYCH TE USŁUGI^a w %**

Przedsiębiorstwa o podstawowym rodzaju działalności PKD	2010	2011	2012	2013	2014
19.2. Wytwarzanie i przetwarzanie produktów rafinacji ropy naftowej	34,4	29,4	24,8	19,8	16,8
23.5. Produkcja cementu, wapna i gipsu			2,2	2,0	2,0
35.1. Wytwarzanie, przesyłanie, dystrybucja i handel energią elektryczną	5,3		1,3		0,8
41.2. Roboty budowlane związane ze wznoszeniem budynków mieszkalnych i niemieszkalnych		1,0			
45.3. Sprzedaż hurtowa i detaliczna części i akcesoriów do pojazdów samochodowych, z wyłączeniem motocykli			2,4		
46.3. Sprzedaż hurtowa żywności, napojów i wyrobów tytoniowych				1,6	
46.4. Sprzedaż hurtowa artykułów użytku domowego	1,5			1,2	
46.5. Sprzedaż hurtowa narzędzi technologii informacyjnej i komunikacyjnej	12,1	10,2	8,4	6,8	5,1
46.9. Sprzedaż hurtowa niewyspecjalizowana		2,0			1,8
47.1. Sprzedaż detaliczna prowadzona w niewyspecjalizowanych sklepach		1,8			
47.3. Sprzedaż detaliczna paliw do pojazdów silnikowych na stacjach paliw	1,2				
47.7. Sprzedaż detaliczna pozostałych wyrobów prowadzona w wyspecjalizowanych sklepach	1,3	1,1			
61.1. Działalność w zakresie telekomunikacji przewodowej		3,1	6,4		
62.0. Działalność związana z oprogramowaniem i doradztwem w zakresie informatyki oraz działalność powiązana				1,0	1,3
63.9. Pozostała działalność usługowa w zakresie informacji			1,8	3,1	
68.2. Wynajem i zarządzanie nieruchomościami własnymi lub dzierżawionymi	1,1	1,2		3,8	1,4
68.3. Działalność związana z obsługą rynku nieruchomości wykonywana na zlecenie	4,4				
69.2. Działalność rachunkowo-księgową; doradztwo podatkowe	1,4	1,0	1,5	1,7	2,8
70.2. Doradztwo związane z zarządzaniem	1,3				
73.1. Reklama				2,0	
81.2. Sprzątanie obiektów				5,3	
82.1. Działalność związana z administracyjną obsługą biura, włączając działalność wspomagającą	25,3	37,8	37,7	42,5	57,5
82.9. Działalność komercyjna, gdzie indziej niesklasyfikowana		1,0			
91.0. Działalność bibliotek i archiwów, muzeów oraz pozostała działalność związana z kulturą			1,7		

^a Wykazano rodzaje działalności, których udział obrotów z tytułu świadczenia usług związanych z administracyjną obsługą biura w przychodach ze sprzedaży produktów, towarów i materiałów ogółem w przynajmniej jednym z analizowanych lat był równy lub przekroczył 1%. Ciemne pole oznacza udział obrotów mniejszy niż 1%.

Źródło: obliczenia własne na podstawie danych GUS.

Metoda C — ISP liczony według produktu dla przedsiębiorstw średnich i dużych (o liczbie pracujących powyżej 9 osób)

Ta metoda ma na celu uzyskanie wskaźnika, który odzwierciedla zmiany wartości obrotów według produktu przedsiębiorstw o liczbie pracujących powyżej 9 osób. Uwzględnia się tu obroty przedsiębiorstw niefinansowych niezależnie od podstawowego rodzaju prowadzonej działalności gospodarczej. ISP obrazuje więc zmiany na rynku usług związanych z administracyjną obsługą biura, a konkretnie — miesięczne zmiany wartości obrotów z tytułu świadczenia omawianych usług przez przedsiębiorstwa niefinansowe o liczbie pracujących powyżej 9 osób. Do jego obliczenia konieczne są jednak roczne dane z formularza SP, dostępne dopiero pod koniec roku następującego po roku sprawozdawczym. Uniemożliwia to obliczenie miesięcznych wskaźników w wymaganych terminach. Rozwiązaniem jest stosowanie dla danego roku sprawozdawczego t udziałów obliczonych dla roku sprawozdawczego $t-1$, a następnie, po uzyskaniu danych dla tego roku, rewidowanie szeregów czasowych.

Metoda D — ISP liczony według produktu dla wszystkich przedsiębiorstw

Ta metoda uwzględnia zarówno pomiar produkcji według produktu, jak i obserwację ogółu przedsiębiorstw niefinansowych. Uzyskany wskaźnik odzwierciedla miesięczne zmiany wartości obrotów przedsiębiorstw niefinansowych całej gospodarki narodowej z tytułu świadczenia usług związanych z administracyjną obsługą biura.

Do obliczenia ISP konieczne są roczne dane z formularzy SP i SP-3. Ze względu na ich dostępność dopiero pod koniec roku następującego po roku sprawozdawczym, obliczenie miesięcznych ISP w terminach wymaganych przez Eurostat jest niemożliwe. Podobnie jak w metodach B i C, rozwiązanie polega na wykorzystaniu dla roku sprawozdawczego t danych dla roku $t-1$, a następnie rewidowanie szeregów czasowych.

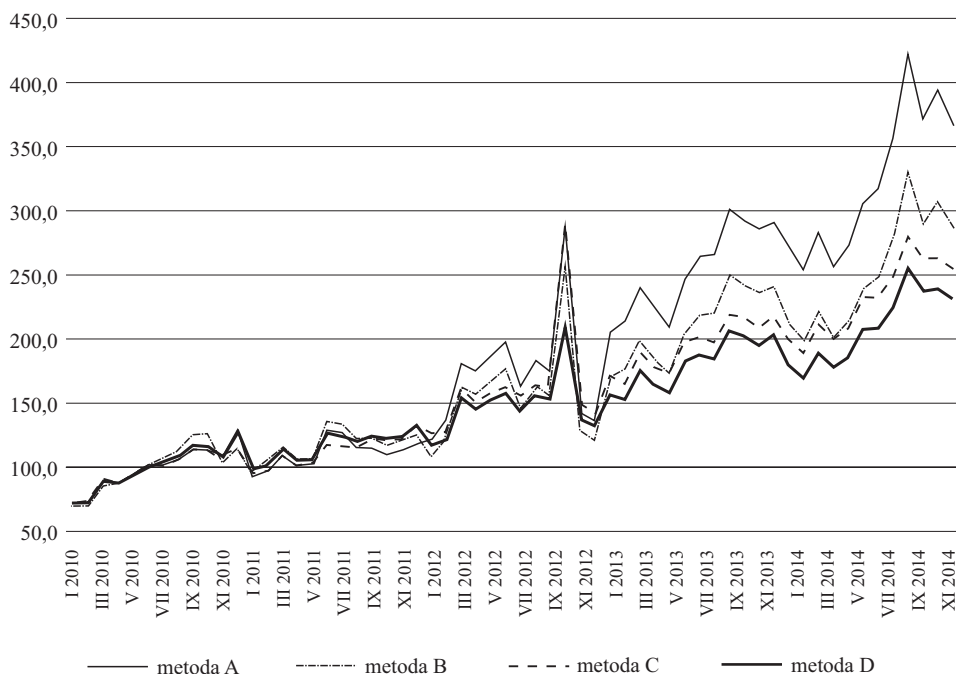
PORÓWNANIE UZYSKANYCH SZEREGÓW CZASOWYCH ISP DLA USŁUG ZWIĄZANYCH Z ADMINISTRACYJNĄ OBSŁUGĄ BIURA

Wyniki eksperymentalnego obliczania wskaźnika wolumenu usług związanych z administracyjną obsługą biura według różnych metod wskazują, że zastosowanie danego schematu implikuje określoną interpretację wyników. Niezależnie jednak od przyjętej metody szeregi wskaźników pokazują ten sam kierunek zmian (wykr. 7). Najbardziej dynamiczny rozwój segmentu omawianych usług można zaobserwować w przypadku metody A, która uwzględnia obroty przedsiębiorstw o liczbie pracujących powyżej 9 osób, prowadzących działalność związaną z administracyjną obsługą biura, a najmniej dynamiczny — w przypadku metody D, czyli pomiaru produkcji całego rynku usług związanych

z administracyjną obsługą biura świadczonych przez ogół przedsiębiorstw niefinansowych.

Wykr. 7. ISP DLA USŁUG ZWIĄZANYCH Z ADMINISTRACYJNĄ OBSŁUGĄ BIURA

2010=100



Źródło: jak przy wykr. 4.

W latach 2010 i 2011 wartość ISP jest zbliżona, ale w kolejnych okresach, szczególnie od stycznia 2013 r. do grudnia 2014 r., występują coraz większe rozbieżności. Na różnice te wpływają poszczególne elementy uwzględniane przy obliczaniu wskaźnika oraz zmiany liczby przedsiębiorstw sklasyfikowanych w grupie 82.1 PKD i liczby osób w nich pracujących. W 2013 r. liczba przedsiębiorstw wzrosła w porównaniu z 2012 r. o 16,1%, a liczba pracujących — o 27,2%. Wzrost odnotowano we wszystkich klasach wielkości przedsiębiorstw mierzonych liczbą pracujących, najbardziej znaczący — w przedsiębiorstwach o liczbie pracujących powyżej 49 osób (liczba przedsiębiorstw zwiększyła się o 75%, a liczba osób pracujących — o prawie 51%).

IPS według działalności obliczany na podstawie danych o obrotach wszystkich przedsiębiorstw o liczbie pracujących do 9 osób prowadzących omawianą działalność (metoda A w zestawieniu z metodą B) wskazuje na wolniejsze tempo wzrostu obrotów tych przedsiębiorstw niż przedsiębiorstw o liczbie pracujących powyżej 9 osób. Ze względu na przyjętą formułę liczenia wartości wskaź-

ników w 2010 r. w poszczególnych szeregach są takie same. W 2011 r. segment przedsiębiorstw o liczbie pracujących do 9 osób rozwijał się szybciej niż przedsiębiorstw większych, natomiast od 2012 r. dynamika obrotów przedsiębiorstw o liczbie pracujących do 9 osób była niższa niż przedsiębiorstw większych. W okresie od stycznia 2012 r. do grudnia 2014 r. wartości obrotów przedsiębiorstw o liczbie pracujących do 9 osób rosły, ale tempo wzrostu było wolniejsze niż w przypadku przedsiębiorstw większych, co spowodowało pogłębianie się różnic w kolejnych okresach sprawozdawczych.

Zmiana metody obliczania ISP dla przedsiębiorstw o liczbie pracujących powyżej 9 osób z liczenia według działalności na liczenie według produktu (metoda A w zestawieniu z metodą C) skutkuje obniżeniem się dynamiki wzrostu produkcji w badanym segmencie usług. W analizowanym okresie odnotowano wyższy wzrost produkcji przedsiębiorstw sklasyfikowanych w grupie 82.1 PKD niż w całym segmencie usług związanych z administracyjną obsługą biura. Przedsiębiorstwa świadczące te usługi jako działalność podstawową uzyskują co prawda przychody ze sprzedaży innych produktów, ale wyższe tempo wzrostu odnotowuje się właśnie w sprzedaży usług świadczonych w ramach ich pierwszorzędnej działalności, co wynika, jak można wnioskować, ze stopnia specjalizacji produkcji i efektów skali.

Podobnie jak w przedsiębiorstwach o liczbie pracujących powyżej 9 osób, także w przypadku wszystkich przedsiębiorstw ISP liczone według działalności (metoda B w zestawieniu z metodą D) wskazuje na wyższą dynamikę zmian niż ISP liczone według produktu. Uwzględnienie we wskaźniku wartości obrotów przedsiębiorstw o liczbie pracujących do 9 osób powoduje zmniejszenie się różnic pomiędzy dynamiką zmian dla przedsiębiorstw o liczbie pracujących do 9 osób oraz dla ogółu przedsiębiorstw, łącznie z najmniejszymi jednostkami.

Produkcja przedsiębiorstw o liczbie pracujących powyżej 9 osób mierzona według produktu (metoda C w zestawieniu z metodą D) charakteryzuje się wyższą dynamiką wzrostu niż produkcja wszystkich przedsiębiorstw. Jest to następstwem włączenia obrotów przedsiębiorstw o liczbie pracujących do 9 osób, które w analizowanym okresie rozwijały się wolniej niż przedsiębiorstwa większe. Najwyższą dynamikę odnotowano dla przedsiębiorstw o liczbie pracujących powyżej 9 osób sklasyfikowanych w grupie 82.1 PKD.

Podsumowanie

Analiza eksperymentalnie obliczonych szeregów wskaźników produkcji usług związanych z administracyjną obsługą biura wskazuje, że wybór metody obliczania wskaźnika ma znaczący wpływ na wyniki, a zatem i sposób ich interpretacji. Główną przesłanką dla zastosowania danej metody mierzenia produkcji powinny być zatem potrzeby informacyjne, które dany wskaźnik ma zaspokoić. Mając alternatywę pomiędzy wskaźnikiem według działalności i według produktu, należy wziąć pod uwagę to, że ISP według produktu pozwala na dokładniejszą ocenę sytuacji na rynku danego produktu, podczas gdy ISP według dzia-

łalności odzwierciedla sytuację przedsiębiorstw prowadzących daną działalność gospodarczą. Te same zależności należy uwzględnić w przypadku każdej zmiany metodologicznej pomiaru wskaźnika wolumenu usług.

Dynamika oraz złożoność zjawisk zachodzących na rynku usług powodują, że obecne metody obliczania ISP poddawane są ciągłym modyfikacjom, aby jak najtrafniej odzwierciedlały zmieniającą się rzeczywistość gospodarczą. Zastosowanie tych metod w praktyce zależy jednak od dostępności danych niezbędnych do kalkulacji wskaźnika. W polskiej statystyce publicznej opracowanie miesięcznego ISP według działalności jest możliwe w terminie 60 dni od zakończenia miesiąca sprawozdawczego, czyli zgodnie z wymaganiami FRIBS. Obliczenie wskaźnika według produktu wiązałoby się natomiast z opóźnieniem, gdyż niezbędne dane dostępne są w listopadzie roku następującego po roku sprawozdawczym. Dla bieżącego roku sprawozdawczego ISP według produktu mógłby być obliczany na podstawie danych rocznych dla roku poprzedniego i rewidowany po uzyskaniu danych za rok bieżący. Podobną procedurę należałoby zastosować w przypadku włączenia do badanej grupy przedsiębiorstw o liczbie pracujących do 9 osób.

Jednocześnie należy podkreślić, że metody obliczania wskaźnika przedstawione na przykładzie usług związanych z administracyjną obsługą biura uwzględniają tylko dwa aspekty pomiaru efektów działalności gospodarczej. Tymczasem rzeczywistość gospodarcza jest dużo bogatsza w zjawiska wpływające na rozwój rynku danego produktu. Istotne byłoby zatem wyodrębnienie wartości obrotów na eksport z całkowitej wartości obrotów uwzględnianych w ISP, wzięcie pod uwagę zmian w poziomach cen usług w międzynarodowym handlu usługami, wpływ outsourcingu, a także uwzględnienie typu usługobiorcy (czy jest nim inne przedsiębiorstwo, czy też gospodarstwo domowe). Opracowanie metodologii liczenia ISP, który uwzględniałby możliwie najwięcej aspektów funkcjonowania rynku usług, stanowi wyzwanie dla statystyków na całym świecie.

Ponadto dynamiczne zmiany obserwowane na rynku usług, które charakteryzują się cyklicznością zdarzeń, stanowią przesłankę dla poszukiwania metod i narzędzi umożliwiających efektywną analizę dostępnych informacji. Zastosowanie w statystyce publicznej metod analiz szeregów czasowych będzie stanowić istotny element w badaniach rozwoju działalności usługowej.

mgr Agnieszka Matulska-Bachura — GUS

LITERATURA

- Baer, A. (2015). *Cross-Cutting Topics: Industry vs. Product*. Pobrano z: www.voorburggroup.org/Documents/2015%20Sydney/1010.pdf (dostęp 30.06.2017).
- Eurostat, OECD. (2014). *Eurostat-OECD Methodological Guide for Developing Producer Price Indices for Services, Second Edition*. Pobrano z: <http://www.oecd.org/std/prices-ppp/eurostat-oecd->

-methodological-guide-for-developing-producer-price-indices-for-services-9789264220676-en.htm (dostęp 30.06.2017).

Maddala, G.S. (2008). *Ekonometria*. Warszawa: Wydawnictwo Naukowe PWN.

Matulska-Bachura, A. (2015). *Mini-presentation on Turnover/Output for Office Administrative and Support Activities (ISIC 8210) in Poland*. Pobrano z: www.voorburggroup.org/Documents/2015%20Sydney/4016.pdf (dostęp 30.06.2017).

OECD. (2007). *Compilation manual for an index of service production*. Pobrano z: www.oecd.org/std/business-stats/37799074.pdf (dostęp 30.06.2017).

Pegler, K., Taylor, C. (2014). *Cross Cutting Topics: CPI use in PPI context*. Pobrano z: <http://www4.statcan.ca/english/voorburg/Documents/2014%20Dublin/Papers/014.pdf> (dostęp 30.06.2017).

Timofiejuk, I., Lasek, M., Pęczkowski, M. (1997). *Miary statystyczne*. Warszawa: GUS.

Summary. *The aim of the article is to describe the methods for compiling the services volume index, named also as the index of services production (ISP), as well as its implementation in Polish practice. The influence, which the method for measuring turnover value (by product or by activity, regarding the size of enterprise) has on time series of production index for office administrative and support activities, was presented. In the research four methods for compiling the ISP were used and then the time series of indices computed for research period 2010—2014 were compared regarding the applied approaches. Based on the results it was settled that each method for computing the services volume index meets different information needs. Therefore, when the method for index computing is chosen, the purpose of undertaken analyses should be considered.*

Keywords: statistical methods, services sector.

Agnieszka NOCKO

Zróźnicowanie długości życia w zależności od płci i wykształcenia

Streszczenie. *W artykule opisano badanie umieralności i trwania życia w zależności od płci i poziomu wykształcenia. Do oceny zjawiska zastosowano wskaźnik częściowego trwania życia osób w wieku od 30 do 69 lat (e_{30-69}). Badanie przeprowadzono dla lat 2002 i 2011. W pierwszym jego etapie oszacowano, metodą Chianga, tablice trwania życia dla osób w wieku 30 lat i więcej według płci oraz czterech grup wykształcenia. Następnie obliczono parametr e_{30-69} . Otrzymane wyniki potwierdziły, że trwanie życia jest zróźnicowane zależnie od płci i wykształcenia. W przypadku osób z wyższym wykształceniem odnotowano dłuższe trwanie życia niż w przypadku osób z wykształceniem niższym. W badanych latach obserwowano wydłużenie trwania życia osób w wieku 30—69 lat z wyjątkiem kobiet z wykształceniem co najwyżej gimnazjalnym. W przypadku kobiet odnotowano mniejsze zróźnicowanie wartości oszacowanego parametru trwania życia pomiędzy najwyższym i najniższym wykształceniem niż u mężczyzn.*

Słowa kluczowe: trwanie życia, wykształcenie, nierówności społeczne, umieralność, Polska.

JEL: I24

Od połowy lat sześćdziesiątych XX w. w wielu zachodnich krajach, a od dziewięćdziesiątych w Polsce i innych krajach Europy Środkowej i Wschodniej obserwuje się systematyczne wydłużanie trwania życia. Jednocześnie notowane są znaczne różnice, jeśli chodzi o stan zdrowia i poziom umieralności. Zarówno międzynarodowe, jak i polskie badania wskazują, że uwarunkowane społecznie, a więc niewynikające z przyczyn biologicznych, nierówności w stanie zdrowia są jedną z najistotniejszych przyczyn nadumieralności (Marek i in., 2012). Jednym z ważnych czynników mających znaczny wpływ na trwanie życia i jego zróźnicowanie w populacji jest wykształcenie.

Zredukowanie nierówności w sferze zdrowia i umieralności stało się jednym z priorytetów polityki społeczno-zdrowotnej w wielu europejskich krajach. W Polsce podstawowym dokumentem poświęconym tej tematyce jest Narodowy Program Zdrowia (NPZ), w którym określono najważniejsze działania na rzecz poprawy zdrowia i związanej z nim jakości życia społeczeństwa. Do celów strategicznych NPZ na lata 2016—2020 zaliczono: wydłużenie życia Polaków, poprawę jakości życia społeczeństwa związaną ze zdrowiem oraz ograniczanie społecznych nierówności w tym zakresie.

Głównym celem artykułu jest analiza umieralności oraz różnic występujących w trwaniu życia w Polsce zależnie od wykształcenia. W opracowaniu do oceny zróżnicowania trwania życia zastosowano wskaźnik częściowego trwania życia osób w wieku od 30 do 69 lat (e_{30-69}). Ze względu na różnice w długości życia mężczyzn i kobiet analizy były prowadzone odrębnie dla każdej płci. Wyniki odnoszą się do lat 2002 i 2011 — wtedy przeprowadzono spisy powszechne i możliwe stało się uzyskanie danych o strukturze ludności według płci, wieku i wykształcenia.

ZRÓŻNICOWANIE UMIERALNOŚCI W EUROPIE W ZALEŻNOŚCI OD WYKSZTAŁCENIA

Wyniki analiz przeprowadzonych przez Europejskie Regionalne Biuro Światowej Organizacji Zdrowia (WHO) wskazują, że życie Europejczyków się wydłuża. Równocześnie zmieniają się rodzaje obciążenia chorobami i zwiększa oddziaływanie czynników, które warunkują zdrowie. Wydłużenie trwania życia nastąpiło głównie dzięki mniejszej częstotliwości występowania niektórych przyczyn zgonów, zmniejszeniu czynników ryzyka oraz poprawie warunków życiowych i społeczno-ekonomicznych (Europejskie Regionalne Biuro WHO, 2013). Według danych Eurostatu¹ średnia długość życia w Unii Europejskiej (UE) w 2015 r. wyniosła 83,3 roku dla kobiet oraz 77,9 dla mężczyzn. Notowane są jednak duże różnice w przeciętnej długości trwania życia pomiędzy krajami. W krajach zachodniej części Europy oraz Skandynawii życie trwa o kilka lat dłużej niż w Europie Środkowej. W porównaniu z krajami Europy Wschodniej różnica ta sięga nawet kilkunastu lat (Rutkowska, 2016). W 2015 r. najdłużej w Europie żyły kobiety mieszkające w Hiszpanii (85,8 roku) oraz Francji (85,5), najkrócej — w Macedonii (77,4) i Serbii (77,9), a spośród państw należących do UE — Bułgarki (78,2) i Rumunki (78,7). Wśród mężczyzn najdłuższe przeciętne trwanie życia odnotowano dla mieszkańców Islandii (81,2) oraz Liechtensteinu (80,9) — biorąc pod uwagę tylko kraje UE, najdłużej żyją Szwedzi (80,4) i Włosi (80,3), najkrócej zaś obywatele Litwy (69,2) i Łotwy (69,7).

Polacy żyją coraz dłużej, jednak nawet o kilka lat krócej niż obywatele krajów plasujących się pod tym względem w czołówce. Dla Polek oszacowana wartość przeciętnego trwania życia w 2015 r. wyniosła 81,6 roku, a dla Polaków — 73,5.

¹ http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_mlexpec&lang=en (dostęp 29.03.2017 r.).

Nierówności pomiędzy poszczególnymi krajami w zakresie przeciętnego trwania życia są większe w przypadku mężczyzn niż kobiet.

Społeczne nierówności w kwestii zdrowia, które przejawiają się m.in. w zróźnicowaniu oczekiwanego trwania życia, są bezpośrednio lub pośrednio spowodowane przez czynniki społeczne, ekonomiczne i środowiskowe (Whitehead i Dahlgren, 2007). Badania prowadzone w wielu krajach potwierdziły, że wykształcenie jest jednym z najważniejszych czynników wpływających na stan zdrowia i długość życia ludności (m.in. Kulhanova, Hoffmann, Eikemo, Menvielle i Mackenbach, 2014; Leinsalu, Vågerö i Kunst, 2003; Leinsalu i in., 2009; Olshansky i in., 2012; Shkolnikov i in., 2006; Steingrimsdottir i in., 2012; Valkonen, Sihvonen i Lahelma, 1997). Wyniki tych analiz wskazały na znaczne różnice w umieralności i trwaniu życia osób z wyższym wykształceniem w porównaniu z tymi, które zakończyły edukację na wcześniejszych etapach.

Występowanie różnic w trwaniu życia w Europie w zależności od wykształcenia potwierdzają także informacje prezentowane w bazie Eurostatu *Life expectancy by age, sex and educational attainment level*². Dostępne są tam dane z lat 2007—2015 dotyczące przeciętnego trwania życia dla 19 krajów europejskich według płci, wieku oraz trzech grup wykształcenia zgodnie z klasyfikacją ISCED 2011³. W pierwszej grupie znalazł się poziom 0—2, określane jako niski poziom edukacji, w drugiej — 3 i 4, nazywane średnim poziomem, a w trzeciej — 5—8, czyli wysoki poziom edukacji.

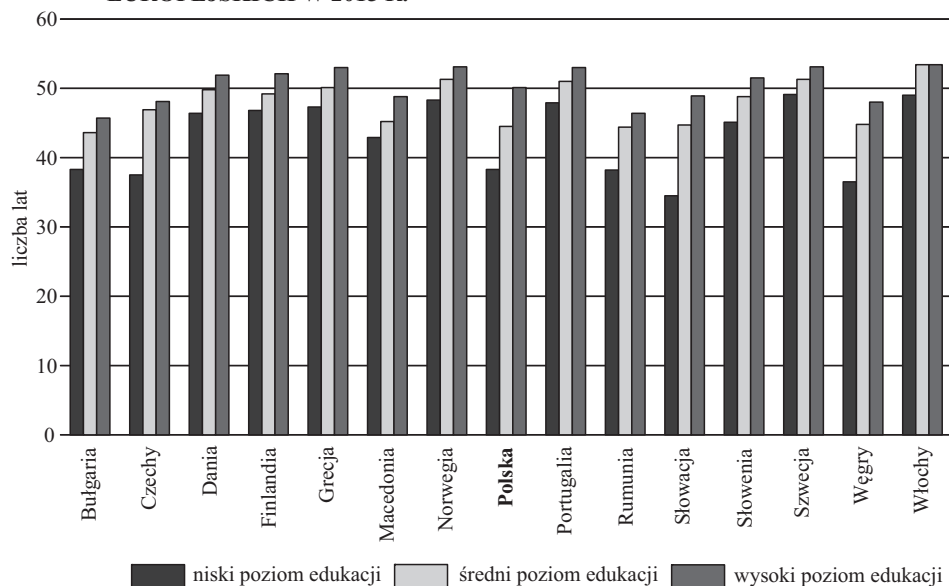
Polacy w wieku 30 lat z wykształceniem wyższym mają przed sobą przeciętnie prawie 12 lat życia więcej niż ci z wykształceniem na niskim poziomie. W przypadku kobiet ta różnica wynosi niewiele ponad 5 lat.

Największe różnice w trwaniu życia trzydziestolatków z wysokim i niskim wykształceniem odnotowano w przypadku mężczyzn dla Słowacji (14,4 roku), a w przypadku kobiet — dla Węgier (6,4) i Słowacji (6,3), zaś najmniejsze w przypadku mężczyzn dla Szwecji (4,0) i Włoch (4,4), a w przypadku kobiet — dla Grecji (2,4) i Szwecji (2,6). Ponadto zauważono, że im wyższe wykształcenie, tym zróźnicowanie dalszego trwania życia jest mniejsze zarówno dla kobiet, jak i mężczyzn. Najmniejsze wartości przeciętnej długości życia odnotowano dla osób z niskim wykształceniem. Trwanie życia ludności Polski, szczególnie z najniższym wykształceniem, znacznie odbiega od trwania życia mieszkańców krajów, dla których notowano największe wartości. W przypadku Polaków z niskim wykształceniem jest to ok. 11 lat mniej od Szwedów, natomiast Polki mają przed sobą przeciętnie o prawie 5 lat krótsze życie niż Portugalki. Przy najwyższym poziomie wykształcenia analogiczne różnice dla kobiet i mężczyzn wyniosły odpowiednio 2,3 i 3,3 roku.

² http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_mlexpecedu&lang=en (dostęp 29.03.2017 r.).

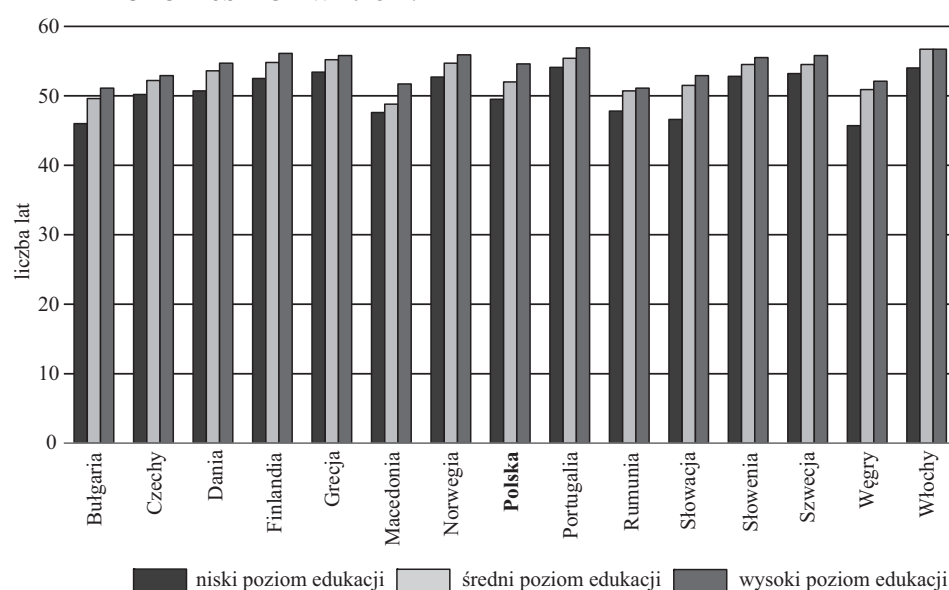
³ Międzynarodowa Standardowa Klasyfikacja Wykształcenia (ISCED) jest stosowana w celach statystycznych na potrzeby porównań międzynarodowych. Do 2013 r. dane o poziomie edukacji i wykształcenia były klasyfikowane według ISCED 1997 (7 poziomów edukacji), a od roku 2014 ujmuje się je zgodnie z ISCED 2011 (9 poziomów). [http://ec.europa.eu/eurostat/statistics-explained/index.php/International_Standard_Classification_of_Education_\(ISCED\)](http://ec.europa.eu/eurostat/statistics-explained/index.php/International_Standard_Classification_of_Education_(ISCED)).

Wykr. 1. PRZECIĘTNE DALSZE TRWANIE ŻYCIA MĘŻCZYŹN W WIEKU 30 LAT WEDŁUG WYKSZTAŁCENIA W POLSCE I WYBRANYCH KRAJACH EUROPEJSKICH W 2015 R.



Źródło: opracowanie własne na podstawie danych Eurostatu.

Wykr. 2. PRZECIĘTNE DALSZE TRWANIE ŻYCIA KOBIEC W WIEKU 30 LAT WEDŁUG WYKSZTAŁCENIA W POLSCE I WYBRANYCH KRAJACH EUROPEJSKICH W 2015 R.



Źródło: jak przy wykr. 1.

O tym, jak duża jest skala nierówności społecznych w zdrowiu według wykształcenia, świadczą wyniki badania, którym objęto wszystkie kraje UE. Stwierdzono, że powodują one dodatkowo ponad 700 tys. zgonów i ok. 33 mln przypadków zapadania na różne choroby w skali roku w porównaniu z hipotetyczną sytuacją, w której wskaźniki umieralności i zdrowia dla wszystkich byłyby takie jak dla osób z wyższym wykształceniem (Mackenbach, Menvielle, Jasilionis i de Gelder, 2015).

OPIS DANYCH I METODY ANALIZY

Do analizy wykorzystano dane GUS dotyczące:

- ludności według poziomu wykształcenia, płci oraz wieku z Narodowego Spisu Powszechnego Ludności i Mieszkań (NSP) 2002 oraz 2011;
- zgonów według wykształcenia, płci oraz wieku osób zmarłych w latach 2001—2003 oraz 2010—2012⁴.

Na potrzeby opracowania wykształcenie ludności oraz osób zmarłych przyporządkowano do jednej z czterech grup:

- I — wyższe (poziomy 5 i 6 według klasyfikacji ISCED 1997);
- II — policealne (poziom 4C) oraz średnie (poziom 3A);
- III — zasadnicze zawodowe (poziom 3C);
- IV — gimnazjalne⁵ (poziom 2A), podstawowe ukończone⁶ (poziom 1), podstawowe nieukończone oraz bez wykształcenia szkolnego (lub odpowiadająca kategoria dla danych o zgonach — niepełne podstawowe).

Na potrzeby analizy jako oddzielną wyodrębniono grupę osób z wykształceniem zasadniczym zawodowym. Według klasyfikacji ISCED 1997 jest ono klasyfikowane jako średni poziom edukacji, jednak analiza przeciętnego trwania życia przeprowadzona na podstawie danych dotyczących umieralności mężczyzn na Węgrzech w latach 1992—2002 pokazała, że bardziej adekwatne jest agregowanie osób z wykształceniem zasadniczym zawodowym do niższego obowiązkowego poziomu wykształcenia niż ujmowanie ich w średnim poziomie edukacji (Mackenbach i in., 2015). W przypadku Polski wyodrębnienie tego wykształcenia pozwoli ocenić, do którego poziomu edukacji najwłaściwiej jest włączać osoby z wykształceniem zasadniczym zawodowym.

Osoby, w przypadku których brakowało danych o wykształceniu zostały proporcjonalnie rozszacowane do poszczególnych grup. Odsetek kobiet w wieku 30—69 lat, dla których w NSP 2002 nie ustalono wykształcenia wyniósł 2,0%, a dla mężczyzn — 2,2%. Dane uzyskane na podstawie NSP 2011 charakteryzują

⁴ Aby zapewnić większe zbiory oraz wyeliminować wahania losowe, które mogły wystąpić w jednym roku, analizy prowadzono na podstawie danych o zgonach z trzech lat, w tym rok spisy sowy uwzględniono jako środkowy (Mackenbach i in., 2015; Marek i in., 2012).

⁵ Gimnazjalny poziom wykształcenia dostępny jest dla ludności według NSP 2011.

⁶ W polskim systemie edukacji przed reformą ukończona szkoła podstawowa (8-klasowa) odpowiadała poziomowi 2.

się większym udziałem osób o nieustalonym wykształceniu — 5,0% kobiet i 5,2% mężczyzn⁷.

Dane dotyczące osób zmarłych charakteryzowały się niższym odsetkiem informacji bez wzmianki o wykształceniu. W latach 2001—2003 w grupie osób w wieku 30—69 lat odsetek zmarłych kobiet, dla których nie notowano danych o wykształceniu wyniósł 0,7%, a dla mężczyzn — 1,0%. W przedziale czasowym 2010—2012 udziały te wyniosły odpowiednio 1,5% i 1,9%.

Badanie umieralności i trwania życia dla każdego analizowanego okresu zostało przeprowadzone w dwóch etapach. W pierwszym oszacowano przekrojowe tablice trwania życia według wykształcenia i płci. Do ich wyznaczenia zastosowano metodologię Chianga, opisaną m.in. przez Chianga (1979) oraz Prestona, Heuveline'a i Guillota (2001). Tablice trwania życia według grup wykształcenia i roczników szacowane były od wieku 30 lat. Taka granica wieku jest zazwyczaj przyjmowana w tego typu analizach i wynika z założenia, że większość osób w tym wieku zakończyła już edukację (Brønnum-Hansen i Baadsgaard, 2012; Nusselder i in., 2005; Shkolnikov i in., 2006, 2007).

W kolejnym etapie analiz obliczono parametr e_{30-69} , który w literaturze określany jest jako czasowe lub częściowe trwanie życia i oznacza średnią liczbę lat życia pomiędzy dwoma rocznikami. Uzyskane oszacowanie określa, ile lat z możliwych 40 przeżywają średnio osoby w danej grupie wykształcenia. Przyjmowanie ograniczenia dla górnej granicy wieku w analizach trwania życia według wykształcenia zalecane jest ze względu na znaczne różnice w strukturze edukacji osób w starszych grupach wieku, w których dominują jednostki z wykształceniem podstawowym lub niższym. Równie istotne jest pokazanie różnic w przedwczesnej umieralności osób według wykształcenia (Brønnum-Hansen i Baadsgaard, 2012; Mackenbach i in., 2015; Shkolnikov i in., 2007; Strību i in., 2010).

Wartość parametru e_{30-69} obliczono korzystając z oszacowanych parametrów tablic trwania życia:

$$e_{x, x+k} = \frac{T_x - T_{x+k}}{l_x}$$

gdzie:

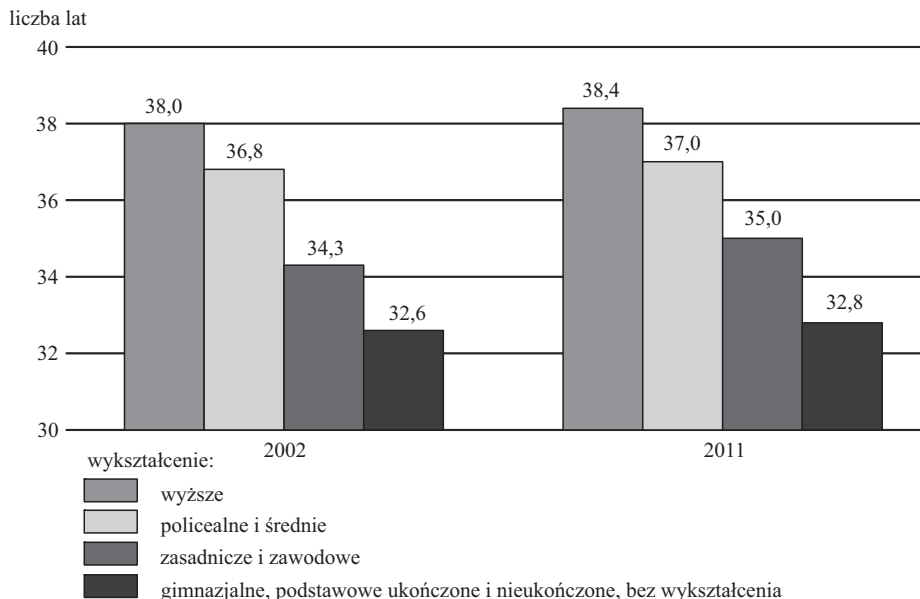
- T_x, T_{x+k} — łączna liczba lat, inaczej nazywana funduszem lat, którą mają do przeżycia wszystkie osoby w wieku x oraz $x+k$ ukończonych lat,
 l_x — liczba osób, które dożyły wieku x lat (Holzer, 2003).

⁷ Na tak znaczące różnice odsetka osób z nieustalonym wykształceniem pomiędzy danymi dla obydwu spisów mogła mieć wpływ metodologia. W 2002 r. NSP przeprowadzono metodą tradycyjną, a w 2011 — mieszaną, czyli z wykorzystaniem systemów informacyjnych administracji publicznej oraz danych zebranych w badaniu pełnym i reprezentacyjnym.

ANALIZA WYNIKÓW

Uzyskane wyniki oszacowania trwania życia od 30 do 69 lat potwierdzają występowanie zróźnicowania w poziomie tego parametru w Polsce zarówno według wykształcenia, jak i płci.

**Wykr. 3. TRWANIE ŻYCIA MĘŻCZYŹN W WIEKU 30—69 LAT (e_{30-69})
WEDŁUG WYKSZTAŁCENIA W POLSCE W LATACH 2002 I 2011**



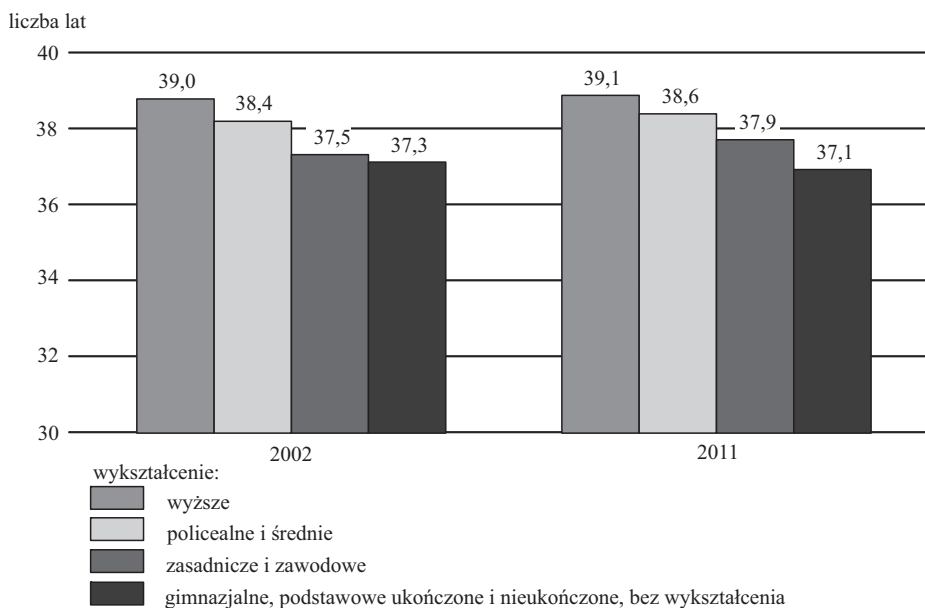
Ź r ó d ł o: opracowanie własne na podstawie danych GUS.

W 2002 r. wartość oszacowanego parametru e_{30-69} dla mężczyzn z wykształceniem wyższym wyniosła 38,0 lat, z policealnym i średnim — 36,8, a z zasadniczym zawodowym — 34,3. Najkrótsze trwanie życia w analizowanym przedziale wieku oszacowano dla mężczyzn, którzy zakończyli edukację na poziomie gimnazjalnym, wyniosło ono 32,6 roku. W 2011 r. zauważono niewielkie zmiany tego parametru dla mężczyzn. Najdłuższe trwanie życia (38,4) dotyczyło tych z wyższym wykształceniem. Niższą o 1,4 roku wartość parametru e_{30-69} obserwowano w przypadku mężczyzn z wykształceniem policealnym i średnim. Najmniejsze wartości zaobserwowano natomiast dla mężczyzn z dwóch najniższych grup wykształcenia.

Najwyższą wartość parametru (39 lat) w 2002 r. odnotowano dla kobiet z wyższym wykształceniem, a z wykształceniem policealnym i średnim — 38,4 roku. Dla kobiet z grup III i IV wartości parametru e_{30-69} były zbliżone i wyniosły odpowiednio 37,5 i 37,3. W 2011 r. najdłuższe trwanie życia odnotowano również dla kobiet z wyższym wykształceniem; dla tych z policealnym i średnim wartość e_{30-69} była mniejsza o pół roku. W przypadku kobiet z wykształceniem

zasadniczym zawodowym trwanie życia w wieku 30—69 lat wyniosło 37,9. Jedynie dla kobiet z wykształceniem co najwyżej gimnazjalnym odnotowano mniejszą wartość parametru (o 0,2) niż w pierwszym analizowanym okresie.

**Wykr. 4. TRWANIE ŻYCIA Kobiet W WIEKU 30—69 LAT (e_{30-69})
WEDŁUG WYKSZTAŁCENIA W POLSCE W LATACH 2002 I 2011**



Źródło: jak przy wykr. 3.

W każdej grupie wykształcenia oszacowane wartości parametru były wyższe dla kobiet niż dla mężczyzn i — co warto podkreślić — im niższa grupa wykształcenia, tym różnica była większa. W analizowanym czasie największą różnicę stwierdzono w grupie osób z najniższym wykształceniem (ponad 4 lata), najmniejszą zaś dla tych z wyższym (ok. 1 roku). Warto też zauważyć, że trwanie życia e_{30-69} kobiet z dwóch najniższych grup wykształcenia było dłuższe od trwania życia mężczyzn z wykształceniem policealnym i średnim.

Zróznicowanie pomiędzy grupami wykształcenia daje się wyraźnie zauważyć. Tak jak wskazywały inne badania, dla osób, które znalazły się w wyższych grupach wykształcenia odnotowano dłuższe czasowe trwanie życia niż dla tych, które zakończyły edukację na wcześniejszym etapie. W analizowanych latach różnica wartości oszacowanego parametru pomiędzy najwyższą a najniższą grupą wykształcenia była znacznie większa dla mężczyzn (powyżej 5 lat) niż dla kobiet (ok. 2 lat).

Analiza wartości oszacowanego parametru dla osób z wykształceniem zasadniczym zawodowym nie przyniosła jednoznacznej odpowiedzi na pytanie, do którego poziomu edukacji (średniego czy niskiego) należy włączać te osoby, gdy

rozpatrujemy umieralność według wykształcenia. W pierwszym badanym okresie wartości trwania życia osób z wykształceniem zasadniczym zawodowym, niezależnie od płci, były bardziej zbliżone do tych z najniższej grupy wykształcenia, a w drugim — do osób z policealnym i średnim wykształceniem.

Podsumowanie

Wyniki oszacowania parametru e_{30-69} pokazały, że wykształcenie jest czynnikiem wpływającym na występowanie znacznych różnic w trwaniu życia kobiet i mężczyzn w Polsce i potwierdziły wyniki badań w innych krajach.

W analizowanych okresach trwania życia osób z wyższym wykształceniem było dłuższe niż osób z niższym. W 2002 r. różnica ta wyniosła 5,4 roku w przypadku mężczyzn i 1,7 roku dla kobiet, natomiast w 2011 r. zwiększyła się i wyniosła odpowiednio 5,6 roku i 2,0 lata. W każdej grupie wykształcenia oszacowane wartości parametru e_{30-69} dla kobiet były większe niż dla mężczyzn, a różnica zwiększała się wraz ze zmniejszaniem się poziomu wykształcenia. Pomiędzy badanymi latami wartość oszacowanego parametru zmniejszyła się jedynie dla kobiet z grupy IV. W obydwu latach w przypadku kobiet wystąpiły mniejsze zróźnicowania wartości parametru e_{30-69} pomiędzy najwyższym a najniższym poziomem wykształcenia niż w przypadku mężczyzn.

Występowanie różnic w tym obszarze może mieć różne przyczyny, do których należy zaliczyć: czynniki społeczne, ekonomiczne, środowiskowe, a także styl życia, w szczególności palenie papierosów i nadużywanie alkoholu. Jak stwierdza Mielecka-Kubień (2012), *jednym z najważniejszych negatywnych skutków nadużywania alkoholu, palenia papierosów i używania narkotyków (...) jest przedwczesna umieralność*. Badanie w Polsce wykazało, że w latach 2006—2008 mężczyźni nadużywający alkoholu w porównaniu z całą populacją żyli przeciętnie o 17 lat krócej, a kobiety — o 8 lat. W przypadku palaczy było to odpowiednio 2,5 i 2,3 roku (Mielecka-Kubień, 2012). Nawyk palenia tytoniu jest typowy głównie dla osób z niższym wykształceniem (Brønnum-Hansen i Baadsgaard, 2012; Steingrimsdottir i in., 2012). Badania stanu zdrowia Polaków także potwierdzają występowanie zależności (bez względu na płeć) pomiędzy wykształceniem a paleniem tytoniu. Im wyższe wykształcenie, tym mniejszy odsetek osób palących (Marek i in., 2012). Analizy przeprowadzone w Wielkiej Brytanii pokazały, że w ok. 30-letniej perspektywie nawet 72% nierówności społecznych, jeśli chodzi o umieralność, może być wyjaśnionych przez związane ze zdrowiem elementy stylu życia, takie jak: palenie tytoniu, picie alkoholu, niezdrowa dieta i brak aktywności fizycznej. Stwierdzono również, że skłonności do niezdrowych zachowań częściej przejawiają osoby o niższej pozycji społeczno-ekonomicznej. Te znajdujące się wyżej w hierachii społecznej, najczęściej lepiej wykształcone, mają szersze możliwości dostępu do opieki zdrowotnej, a także większą otwartość na zmiany zachowania związane z żywieniem czy aktywnością fizyczną (Stringhini i in., 2010).

Ważnym aspektem badań prowadzonych nad umieralnością ze względu na wykształcenie jest analiza przyczyn zgonów. Analiza przyczyn zgonów możliwych do uniknięcia przeprowadzona na podstawie danych dotyczących osób zmarłych w wieku 30—64 lata według dwóch grup wykształcenia⁸ w 16 krajach europejskich (w tym w Polsce) wykazała, że dla wszystkich krajów i typów przyczyn zgonów możliwych do uniknięcia występują nierówności ze względu na wykształcenie. Osoby z niższym mają wyższe wskaźniki udziału przyczyn zgonów możliwych do uniknięcia. Największe nierówności odnotowano w Europie Środkowej i Wschodniej, najmniejsze zaś w krajach leżących na południu Europy (Stirbu i in., 2010). Z analizy danych o zgonach w Polsce w latach 2005—2007 wynika, że współczynniki zgonów dla wszystkich głównych przyczyn śmierci są najwyższe dla osób z niskim wykształceniem (Marek i in., 2012).

mgr Agnieszka Nocko — *Urząd Statystyczny w Lublinie, doktorantka w SGH*

Autorka bardzo dziękuje prof. SGH Wiktorii Wróblewskiej za cenne uwagi podczas przygotowywania artykułu oraz Recenzentom za sugestie pozwalające na lepsze opracowanie ostatecznej wersji tekstu.

LITERATURA

- Brønnum-Hansen, H., Baadsgaard, M. (2012). Widening social inequality in life expectancy in Denmark. A register-based study on social composition and mortality trends for the Danish population. *BioMed Central Public Health*, no. 12, s. 994—1000. DOI: 10.1186/1471-2458-12-994.
- Chiang, C.L. (1979). *Life tables and mortality analysis*. Genewa: WHO.
- Europejskie Regionalne Biuro WHO (2013). *Europejski raport zdrowia 2012: Droga do osiągnięcia dobrostanu*. Kopenhaga: WHO.
- Holzer, J.Z. (2003). *Demografia*. Warszawa: Polskie Wydawnictwo Ekonomiczne.
- Kulhanova, I., Hoffmann, R., Eikemo, T.A., Menvielle, G., Mackenbach, J.P. (2014). Educational inequalities in mortality by cause of death: first national data for the Netherlands. *International Journal of Public Health*, 59 (5), s. 687—696. DOI: 10.1007/s00038-014-0576-4.
- Leinsalu, M., Vågerö, D., Kunst, A.E. (2003). Estonia 1989—2000: enormous increase in mortality differences by education. *International Journal of Epidemiology*, 32 (6), s. 1081—1087. DOI: 10.1093/ije/dyg192.
- Leinsalu, M., Stirbu, I., Vagero, D., Kalediene, R., Kovács, K., Wojtyniak, B., Wróblewska, W., Mackenbach, J.P., Kunst, A.E. (2009). Educational inequalities in mortality in four Eastern European countries: divergence in trends during the post-communist transition from 1990 to 2000. *International Journal of Epidemiology*, 38 (2), s. 512—525. DOI: 10.1093/ije/dyn248.
- Mackenbach, J.P., Menvielle, G., Jasilionis, D., de Gelder, R. (2015). Measuring educational inequalities in mortality statistics. *OECD Statistics Working Papers*, 2015/08, Paryż: OECD Publishing. DOI: 10.1787/5jrppx182zs-en.

⁸ Według klasyfikacji ISCED 1997 niższy poziom edukacji to poziomy 0—2, natomiast wyższy — pozostałe poziomy.

- Marek, M., Chłoń-Domińczuk, A., Kaleta, D., Mazur, J., Miśkiewicz, P., Poznańska, D., Stokwieszewski, J., Wojtyniak, B., Zdrojewski, T. (2012). *Spoleczne nierówności w zdrowiu w Polsce*. Kopenhaga: WHO.
- Mielecka-Kubieñ, Z. (2012). *Nierówności trwania życia w Polsce. Analiza statystyczno-demograficzna*. Katowice: Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.
- Nusselder, W.J., Looman, C., Mackenbach, J.P., Huisman, M., van Oyen, H., Deboosere, P., Gadeyne, S., Kunst, A.E. (2005). The contribution of specific diseases to educational disparities in disability-free life expectancy. *American Journal of Public Health, 95* (11), s. 2035—2041. DOI: 10.2105/AJPH.2004.054700.
- Olshansky, J.S., Antonucci, T., Berkman, L., Binstock, R.H., Boersch-Supan, A., Cacioppo, J.T., Bruce, A., Carnes, B.A., Carstensen, L.L., Fried, L.P., Goldman, D.P., Jackson, J., Kohli, M., Rother, J., Zheng, Y., Rowe, J. (2012). Differences in life expectancy due to race and educational differences are widening, and many may not catch up. *Health Affairs, 31* (8), s. 1803—1813. DOI: 10.1377/hlthaff.2011.0746.
- Preston, S.H., Heuveline, P., Guillot, M. (2001). *Demography: measuring and modeling population processes*. Oxford: Blackwell Publishers.
- Rutkowska, L. (2016). *Trwanie życia w 2015 r.* Warszawa: GUS.
- Shkolnikov, V.M., Andreev, E.M., Jasilionis, D., Leinsalu, M., Antonova, O.I., McKee, M. (2006). The changing relation between education and life expectancy in central and eastern Europe in the 1990s. *Journal of Epidemiology and Community Health, 60* (10), s. 875—881. DOI: 10.1136/jech.2005.044719.
- Shkolnikov, V.M., Jasilionis, D., Andreev, E.M., Jdanov, D.A., Stankuniene, V., Ambrozaitiene, D. (2007). Linked versus unlinked estimates of mortality and length of life by education and marital status: evidence from the first record linkage study in Lithuania. *Social Science and Medicine, 64* (7), s. 1392—1406. DOI: 10.1016/j.socscimed.2006.11.014.
- Stirbu, I., Kunst, A.E., Bopp, M., Leinsalu, M., Regidor, E., Esnaola, S., Costa, G., Martikainen, P., Borrell, C., Kalediene, R., Rychtarikova, J., Artnik, B., Deboosere, P., Mackenbach, J.P. (2010). Educational inequalities in avoidable mortality in Europe. *Journal of Epidemiology and Community Health, 64* (10), s. 913—920. DOI: 10.1136/jech.2008.081737.
- Steingrimsdottir, O.A., Naess, Ø., Moe, J.O., Grøholt, E.K., Thelle, D.S., Strand, B.H., Bævre, K. (2012). Trends in life expectancy by education in Norway 1961—2009. *European Journal of Epidemiology, 27* (3), s. 163—171. DOI: 10.1007/s10654-012-9663-0.
- Stringhini, S., Sabia, S., Shipley, M., Brunner, E., Nabi, H., Kivimaki, M., Singh-Manoux, A. (2010). Association of socioeconomic position with health behaviors and mortality. *Journal of the American Medical Association, 303* (12), s. 1159—1166. DOI: 10.1001/jama.2010.297.
- Whitehead, M., Dahlgren, G. (2007). *Concepts and principles for tackling social inequities in health: levelling up part 1*. Kopenhaga: WHO.
- Valkonen, T., Sihvonen, A.-P., Lahelma, E. (1997). Health expectancy by level of education in Finland. *Social Science and Medicine, 44* (6), s. 801—808.

Summary. *The research on mortality and life expectancy by sex and education level was described in the article. To assess this phenomenon partial life expectancy rate of persons aged from 30 to 69 was used (e_{30-69}). The research was conducted for the years 2002 and 2011. At the first stage life expectancy tables for people at the age of 30 and more, by sex and four groups of education, were estimated by Chiang's method. Then on this basis the parameter e_{30-69} was computed. The obtained results confirmed that life expectancy is differentiated by sex and education. Longer life expectancy was noticed for people with higher*

education than for those with lower education. In the analysed period life expectancy of persons aged 30—69 was extended, except for women with at most lower secondary education. For women, the value of the assessed partial life expectancy parameter was less differentiated between the highest and the lowest level of education than for men.

Keywords: life expectancy, education, social inequalities, mortality, Poland.

Marlena PIEKUT

Klasyfikacja wiejskich gospodarstw domowych ze względu na kierunki i rodzaje rozchodów

Streszczenie. *Celem badania jest wyodrębnienie grup wiejskich gospodarstw domowych o podobnych rozchodach oraz ich opis za pomocą cech społeczno-demograficznych i ekonomicznych. Przeprowadzono je z wykorzystaniem wielowymiarowych metod statystycznych — analizy skupień metodą k-średnich oraz analizy dyskryminacyjnej. Materiał badawczy stanowiły dane z badania budżetów gospodarstw domowych GUS za lata 2004 i 2012. W wyniku badania dokonano podziału wiejskich gospodarstw domowych na cztery grupy ze względu na rozchody, przy czym jedna z nich objęła ponad 2/3 gospodarstw. Zmiennymi najbardziej dyskryminującymi przynależność gospodarstw domowych do poszczególnych grup okazały się liczba osób w gospodarstwie domowym oraz dochód rozporządzalny przypadający na osobę.*

Słowa kluczowe: konsumpcja, wydatki, budżety domowe, wielowymiarowe metody statystyczne.

JEL: D12, D14

W Polsce mieszkańcy wsi stanowią blisko 40% ludności ogółem. Cechą charakterystyczną wiejskich gospodarstw domowych są mniejsze dochody rozporządzalne na osobę w porównaniu z miejskimi gospodarstwami domowymi i tym samym niższe wydatki. W 2015 r. przeciętny miesięczny dochód rozporządzalny na osobę na wsi był o 41,6% niższy niż w mieście, a wydatki okazały się niższe o 38,8% (GUS, 2016). Wiejskie gospodarstwa domowe nie tworzą jednak jednolitej grupy. Są zróżnicowane pod względem: wykształcenia i wieku głowy domu, wysokości i źródeł uzyskiwanych dochodów, faktu posiadania lub nieposiadania gruntów rolnych i innych cech determinujących konsumpcję. Można się więc spodziewać, że charakteryzują je różne profile wydatków, w zależności od cech ekonomicznych i społeczno-demograficznych gospodarstwa domowego.

W Polsce wielu badaczy monitoruje sytuację społeczno-ekonomiczną w wiejskich gospodarstwach domowych, traktuje je jednak jako jednolite w swym składzie (Wrzochalska, 2005; Utzig, 2014). Część badaczy (Leszczyńska, 2007; Chmielewska, 2013a) wśród wiejskich gospodarstw domowych wyróżnia gospodarstwa rolników, wskazując na specyfikę ich zachowań konsumpcyjnych. Artykuł stanowi próbę wyodrębnienia typów wiejskich gospodarstw domowych ze względu na rozporządzanie budżetem domowym.

Celem opracowania jest wyłonienie grup wiejskich gospodarstw domowych o podobnej wysokości rozchodów i opis tych grup za pomocą cech społeczno-demograficznych i ekonomicznych.

MATERIAŁ I METODYKA BADANIA

Do przeprowadzenia analiz wykorzystano dane z badania budżetów gospodarstw domowych GUS za rok 2004 (10508 wiejskich gospodarstw domowych) i 2012 (15742 wiejskie gospodarstwa domowe). Wybór lat był podyktowany chęcią zaobserwowania zmian, jakie dokonały się w budżetach tych gospodarstw od wstąpienia Polski do Unii Europejskiej do 2012 r.

Analiza skupień

Do wyodrębnienia grup gospodarstw domowych o podobnych rozchodach zastosowano analizę skupień metodą *k*-średnich z wykorzystaniem techniki *data mining*. Celem algorytmu jest zmniejszenie całkowitej wariancji między skupiskami (Dulli, Furini i Peron, 2009). Grupowanie polega na określeniu liczby grup i ich centroidów (środków ciężkości), a następnie dokonaniu przyporządkowania każdego obiektu do grupy, której centroid jest mu najbliższy (np. w myśl odległości euklidesowej) (Everitt, Landau, Leese i Stahl, 2011). Po każdym takim działaniu centroidy są modyfikowane na podstawie aktualnej zawartości skupień, a następnie wykonywane jest kolejne przyporządkowanie obiektów według analogicznych zasad. Postępowanie to prowadzi się do momentu, w którym obiekty przestaną ulegać przemieszczeniom.

Algorytm *k*-średnich okazuje się stosunkowo prosty i szybki. Kluczową kwestią jest poddanie metodzie właściwej liczby skupień. Do oceny grupowania posłużono się wskaźnikami homogeniczności i heterogeniczności oraz poprawności skupień. W ocenie tej stosuje się koncepcję środka ciężkości grupy i odległości od niego. W badaniu wykorzystano podejście, w którym środek ciężkości danego skupienia zastąpiony został medianą Webera jej elementów. Mediana Webera stanowi wielowymiarowe uogólnienie klasycznego pojęcia mediany (Młodak, 2006).

Przed grupowaniem dokonano zabiegu normalizacji zmiennych. Wykorzystano metodę standaryzacji. Wartości przekształcone w wyniku standaryzacji mają rozkład o średniej 0 i odchyleniu standardowym 1. Dzięki standaryzacji danych

wejściowych wyniki analiz są całkowicie niezależne od jednostek pomiaru zmiennych.

Zmiennymi diagnostycznymi poddanymi analizie skupień było 18 kategorii rozchodów w budżetach domowych w przeliczeniu na osobę. Są to rozchody na: żywność i napoje bezalkoholowe, napoje alkoholowe i wyroby tytoniowe, odzież i obuwie, utrzymanie mieszkania i opłaty za energię, gaz i wodę, wyposażenie mieszkania i prowadzenie gospodarstwa domowego, zdrowie, transport, łączność, rekreację i kulturę, edukację, gastronomię i zakwaterowanie, pozostałe wydatki na dobra i usługi konsumpcyjne, pozostałe wydatki, podatki i inne opłaty oraz rozchody kapitałowe (rzeczowe), finansowe, bieżące, związane z gospodarstwem rolnym oraz inwestycyjne na gospodarstwo rolne.

Efektywny dobór zmiennych diagnostycznych do analizy skupień oraz zmiennych niezależnych w analizie dyskryminacyjnej wymagał weryfikacji zestawu wyjściowego pod kątem zmiennościowo-korelacyjnym, co też uczyniono.

Przedstawione w artykule wartości rozchodów poddano ważeniu. Dla pierwszej osoby dorosłej w gospodarstwie przydzielono wagę 1, dla każdego następnego członka gospodarstwa domowego w wieku 14 lat lub więcej — 0,5, a dla każdego dziecka poniżej 14 lat — 0,3.

Analiza dyskryminacyjna

Do odkrywania cech charakterystycznych poszczególnych skupień zastosowano analizę dyskryminacji. Metoda ta polega na zakwalifikowaniu obiektu do określonej, arbitralnie zdefiniowanej klasy na podstawie odpowiedniego zestawu zmiennych objaśniających i efektywnie skonstruowanej funkcji dyskryminacyjnej (Panek, 2009).

Analiza dyskryminacyjna służy do dwóch celów — dyskryminacji, czyli algebraicznego i graficznego opisu różnych właściwości obiektów należących do rozłącznych względem siebie zbiorów i klasyfikacji obiektów do dwóch lub więcej zbiorów. W pierwszym przypadku zadaniem badacza jest znalezienie „dyskryminatorów”, których wariancja separuje te zbiory od siebie tak bardzo jak to możliwe, a w drugim badacz poszukuje optymalnej reguły matematycznej, która zostanie wykorzystana w celu przyporządkowania nowych obiektów do właściwego zbioru (Johnson i Wichern, 1992). Z rachunkowego punktu widzenia wskazuje się, że analiza dyskryminacyjna podobna jest do analizy wariancji (Sokołowski, 2004).

Ważnymi miarami statystycznymi służącymi do rozstrzygnięcia istotności tego, czy dana zmienna istotnie dyskryminuje grupy są współczynnik lambda Wilksa oraz test F. Wartość współczynnika lambda Wilksa wskazuje na moc dyskryminacyjną całego modelu i określa, w jakim stopniu wartości funkcji dyskryminacji są zakłócone innymi czynnikami niż przynależność do wyróżnionych grup. Wartości statystyki lambda Wilksa zawierają się w przedziale od 0 do 1, przy czym wartość 0 oznacza doskonałą moc dyskryminacyjną, zaś wartość 1 — brak mocy dyskry-

minacyjnej. Oznacza to, że im niższa jest wartość omawianego współczynnika, tym jakość oszacowanego modelu okazuje się większa (Rószkiewicz, 2002).

Współczynnik lambda Wilksa obliczany jest jako:

$$\lambda = \frac{\det \mathbf{G}}{\det(\mathbf{G} + \mathbf{H})} \quad (1)$$

gdzie:

\mathbf{G} — macierz wariancji-kowariancji wewnątrzgrupowej,

\mathbf{H} — macierz wariancji-kowariancji międzygrupowej.

Wartość statystyki F dla określonej zmiennej wskazuje jej istotność statystyczną w dyskryminowaniu grup. Im indywidualny wkład zmiennej w przewidywanie przynależności do grupy jest większy, tym znaczenie tej zmiennej również jest większe. Statystyka F obliczana jest jako stosunek wariancji międzygrupowej do wariancji wewnątrzgrupowej. Jeśli wariancja międzygrupowa jest istotnie większa, to wynika to z tego, że występują istotne różnice między średnimi (Sokołowski, 2004). Test F wyznacza się jako:

$$F_k = \left(\frac{N - K - I}{K - 1} \right) \frac{1 - \lambda_k^{cz}}{\lambda_k^{cz}} \quad (2)$$

gdzie:

N — łączna liczba obiektów w próbie,

K — liczba zmiennych,

I — liczba rozważanych populacji,

λ_k^{cz} — cząstkowy współczynnik Wilksa, który obrazuje wzór $\lambda_k^{cz} = \frac{\lambda_1}{\lambda_0}$, gdzie:

λ_1 — wartość współczynnika lambda Wilksa dla modelu po wprowadzeniu do niego danej zmiennej,

λ_0 — wartość współczynnika lambda Wilksa dla modelu przed wprowadzeniem danej zmiennej.

Przy omawianiu wyników analizy dyskryminacyjnej pojawiają się pojęcia współczynników dyskryminacyjnego i strukturalnego. Współczynnik dyskryminacyjny (waga funkcji dyskryminacyjnej¹) wskazuje na ilościowy, statystyczny wkład zmiennej źródłowej w objaśnienie zmiennej dyskryminacyjnej. Na podstawie współczynników dyskryminacyjnych określa się siłę oraz kierunek

¹ Są to odpowiedniki estymowanych współczynników regresji w regresji wielorakiej.

wpływu poszczególnych zmiennych na zmienną zależną. Im większa jest wartość współczynnika dyskryminacyjnego, tym ów wpływ dyskryminacyjny jest silniejszy.

Współczynniki strukturalne (współczynniki korelacji, ładunki dyskryminacyjne) wskazują na ścisłość związków statystycznych zmiennych źródłowych z pierwiastkiem. Współczynniki te pokazują korelacje między zmiennymi i funkcjami dyskryminacyjnymi. Im większe wartości współczynników strukturalnych, tym za mocniejszą można uznać siłę dyskryminacyjną analizowanej zmiennej niezależnej. Wskazuje to na jej znaczną rolę w decydowaniu o przynależności badanych obiektów do wyróżnionych grup (Rószkiewicz, 2002). Kiedy współczynniki strukturalne mają wartość bliską zera, wtedy funkcję i zmienną mało łączy. Wartości współczynników strukturalnych bliskie -1 i 1 występują natomiast wtedy, gdy funkcja dostarcza tyle samo informacji, co zmienna. Miary te są częściej stosowane do rzeczywistej interpretacji funkcji, gdyż oznaczają proste korelacje między zmiennymi i funkcjami, zaś wagi dyskryminacyjne określają indywidualny wkład każdej zmiennej do funkcji dyskryminacyjnej.

Kolejnym pojęciem związanym z analizą dyskryminacyjną są pierwiastki. Największy wkład w wyjaśnianie badanej zmienności wnosi pierwszy pierwiastek, zaś kolejne — coraz mniejszy. Jednak udział pozostałych pierwiastków jest swoisty i dotyczy zmienności niewyjaśnionej przez pierwiastki wcześniejsze w hierarchii.

Potrzeba wyodrębnienia zmiennej grupującej (zależnej) oraz zmiennych niezależnych w modelu stanowi cel utworzenia funkcji dyskryminacyjnej. Zmienne powinny być słabo skorelowane ze sobą i jednocześnie cechować się wystarczającą zmiennością oraz w istotny sposób oddziaływać na zmienną grupującą.

W badaniu zmienną grupującą stanowiła zmienna określająca przynależność obiektów do skupień, natomiast zmiennymi niezależnymi były następujące cechy gospodarstwa domowego: dochód rozporządzalny na osobę, liczba osób w gospodarstwie domowym, wiek i wykształcenie głowy domu, faza cyklu życia rodziny², kategoria społeczno-ekonomiczna gospodarstwa domowego, powierzchnia użytkowa mieszkania i powierzchnia użytków rolnych.

Do konstrukcji modelu wykorzystano, obok zmiennych o charakterze mierzalnym, także zmienne jakościowe (np. wykształcenie głowy gospodarstwa domowego). Włączenie do modelu zmiennych niemierzalnych wymagało zastosowania procedury przekształcenia ich na zmienne zero-jedynkowe. Zmienna zero-jedynkowa to taka, która przyjmuje dwie wartości — 0 (brak danej cechy) lub 1 .

Warto pamiętać, że zmienne zero-jedynkowe mają dużą moc dyskryminacyjną, ponieważ dzielą zbiór obiektów na dwie części — posiadające daną cechę i jej nieposiadające. Liniowe funkcje dyskryminacyjne zawierające takie zmienne zachowują się dobrze, jeżeli klasy są liniowo separowalne, w przeciwnym

² Wyróżnia się następujące fazy cyklu życia rodziny: I — narzeczeństwo, II — wczesna faza małżeństwa, III — rodzina z małymi dziećmi, IV — dojrzewanie dzieci, V — odchodzenie dorosłych dzieci z domu, VI — tzw. puste gniazdo.

razie nie dają pożądaných efektów. Duża liczba zmienných zero-jedynkowych powoduje, że funkcje dyskryminacyjne nie są optymalne i bywają trudne do zinterpretowania (Gatnar, 1998).

Zabieg zmiany zmienných jakościowych na zero-jedynkowe przeprowadzono w odniesieniu do następujących predyktorów: faza cyklu życia rodziny, wykształcenie głowy domu oraz kategoria społeczno-ekonomiczna gospodarstwa domowego. Przy szacowaniu modeli z układami zmienných zero-jedynkowych pomija się w każdym układzie — w celu uniknięcia współliniowości — jedną ze zmienných zero-jedynkowych. Parametry dla pozostałych kategorii zmienných mają wtedy swoiste znaczenie, gdyż mierzą wpływ tych wariantów na zmienne objaśniane w odniesieniu do wariantu pominiętego (Podolec i Ulman, 2008).

Pominięte kategorie zmienných to w przypadku:

- fazy cyklu życia rodziny — gospodarstwa domowe VI fazy,
- wykształcenia głowy gospodarstwa domowego — gospodarstwa reprezentowane przez osoby ze średnim wykształceniem,
- grupy społeczno-zawodowej — w 2004 r. pracownicy użytkujący gospodarstwo rolne, a w 2012 r. gospodarstwa osób utrzymujących się ze świadczeń społecznych.

Konkludując, zmienną zależną wykorzystaną w analizie dyskryminacyjnej była zmienna określająca przynależność gospodarstwa domowego do jednego z czterech skupień. Zmienne niezależne to: wielkość gospodarstwa domowego, dochód rozporządzalny na osobę, powierzchnia gospodarstwa rolnego w hektarach, ogólna powierzchnia użytkowa mieszkania, wiek głowy domu, grupa społeczno-ekonomiczna (6 zmienných zero-jedynkowych), faza cyklu życia gospodarstwa domowego (5 zmienných zero-jedynkowych) i wykształcenie głowy domu (3 zmienne zero-jedynkowe). Fazy cyklu życia rodziny wskazują na liczbę lat, która upłynęła od założenia gospodarstwa domowego lub zawarcia związku małżeńskiego. Fazy I oraz V i VI to gospodarstwa domowe bez dzieci na utrzymaniu, natomiast fazy od II do IV — z dziećmi.

Wyniki analizy dyskryminacyjnej wykorzystuje się też do przewidywania przynależności danego obiektu (uczestniczącego w analizie lub nowego) do konkretnej grupy. Na użytek tej procedury tworzone są funkcje klasyfikacyjne, nie należy ich jednak mylić z dyskryminacyjnymi (Gatnar, 1995; Sokołowski, 2004), a różnica między nimi polega na występowaniu w pierwszej z nich komponentu losowego — błędu. Podsumowując, analiza dyskryminacyjna stanowi narzędzie bardzo przydatne w wykrywaniu tych zmienných, które pozwalają wyodrębnić różne grupy, a także w klasyfikowaniu przypadków do różnych grup (Sokołowski, 2004).

GRUPY GOSPODARSTW DOMOWYCH W 2004 R.

Zastosowanie analizy skupień metodą *k*-średnich z wykorzystaniem technik *data mining* pozwoliło na wyodrębnienie czterech skupień wiejskich gospodarstw domowych. Pierwsze trzy obejmowały od 5,8% gospodarstw domowych w I do 15,2%

w III skupieniu. W IV znalazło się 70,7% wiejskich gospodarstw domowych. Zakwalifikowanie do jednego skupienia tak dużej liczby gospodarstw domowych może wskazywać na to, że w znaczącej części wiejskich gospodarstw domowych realizowany jest podobny model rozdysponowywania rozchodów.

Przyjęto, że każde skupienie gospodarstw domowych reprezentuje odrębny model rozchodów, którym jest centroida skupienia. Jest to wektor ze średnich wartości poszczególnych rozchodów na osobę w gospodarstwie domowym.

W najmniejszym, I skupieniu wiejskich gospodarstw domowych zaobserwowano największe rozchody ogółem na osobę w gospodarstwie domowym (tabl. 1). Cechą charakterystyczną skupienia były niskie — w porównaniu z innymi skupieniami — udziały nakładów na żywność i napoje bezalkoholowe oraz użytkowanie mieszkania i nośniki energii w wydatkach ogółem. Dużą część budżetu przeznaczano natomiast na odzież i obuwie, transport, łączność i usługi. O ile niewielkie udziały wydatków na wymienione kategorie dóbr nie świadczyły o tym, że były one najniższe wśród omawianych skupień, o tyle najwyższe wpływały na parokrotnie większe wydatki na omawiane dobra i usługi w porównaniu z analogicznymi wydatkami w wiejskich gospodarstwach domowych innych skupień. W rozchodach ogółem w gospodarstwach domowych I skupienia zaobserwowano też wysoki udział rozchodów finansowych oraz bieżących i inwestycyjnych związanych z gospodarstwem rolnym. Fakt ten świadczy o inwestowaniu w gospodarstwa rolne oraz spłacie zadłużeń pojawiających się w ramach funkcjonowania tych gospodarstw.

W II skupieniu wiejskich gospodarstw domowych zaobserwowano największy poziom i udział wydatków na zdrowie oraz względnie wysokie udziały wydatków na użytkowanie mieszkania i nośniki energii, natomiast poziomy i udziały wydatków na: transport i łączność, usługi, napoje alkoholowe i wyroby tytoniowe oraz pozostałe wydatki były stosunkowo niskie, w szczególności w porównaniu z I i III skupieniem (tabl. 1).

TABL. 1. MIESIĘCZNE ROZCHODY NA OSOBĘ I ICH UDZIAŁY W ROZCHODACH OGÓŁEM W POSZCZEGÓLNYCH SKUPIENIACH WIEJSKICH GOSPODARSTW DOMOWYCH W 2004 R.

Wyszczególnienie	Skupienia							
	I		II		III		IV	
	w zł	w %	w zł	w %	w zł	w %	w zł	w %
Żywność i napoje bezalkoholowe	260	23,0	260	27,2	294	32,3	158	39,2
Napoje alkoholowe i wyroby tytoniowe	23	2,0	15	1,6	37	4,1	11	2,7
Odzież i obuwie	161	14,3	31	3,2	26	2,9	18	4,5
Użytkowanie mieszkania i nośniki energii	170	15,1	204	21,3	182	20,0	65	16,1
Wyposażenie mieszkania i prowadzenie gospodarstwa domowego	66	5,8	39	4,1	50	5,5	19	4,7
Zdrowie	35	3,1	176	18,4	34	3,7	15	3,7
Transport	119	10,5	62	6,5	82	9,0	41	10,2
Łączność	52	4,6	34	3,6	43	4,7	16	4,0
Rekreacja i kultura	72	6,4	36	3,8	48	5,3	19	4,7

TABL. 1. MIESIĘCZNE ROZCHODY NA OSOBĘ I ICH UDZIAŁY W ROZCHODACH OGÓŁEM W POSZCZEGÓLNYCH SKUPIENIACH WIEJSKICH GOSPODARSTW DOMOWYCH W 2004 R. (dok.)

Wyszczególnienie	Skupienia							
	I		II		III		IV	
	w zł	w %	w zł	w %	w zł	w %	w zł	w %
Edukacja	20	1,8	3	0,3	7	0,8	4	1,0
Gastronomia i zakwaterowanie	15	1,3	7	0,7	7	0,8	4	1,0
Pozostałe wydatki na towary i usługi ...	80	7,1	39	4,1	50	5,5	24	6,0
Pozostałe wydatki	56	5,0	50	5,2	49	5,4	9	2,2
Wydatki ogółem	1129	100,0	956	100,0	909	100,0	403	100,0
Podatki i inne opłaty	40	2,2	67	5,6	126	9,8	22	3,9
Rozchody kapitałowe (rzeczowe)	15	0,8	11	0,9	8	0,6	3	0,6
Rozchody finansowe	185	10,1	55	4,5	81	6,3	36	6,5
Rozchody bieżące związane z gospodarstwem rolnym	389	21,2	84	6,9	152	11,9	80	14,6
Rozchody inwestycyjne na gospodarstwo rolne	78	4,2	36	3,0	5	0,4	6	1,0
Rozchody ogółem	1836	100,0	1209	100,0	1281	100,0	550	100,0

Ź r ó ł o: opracowanie własne na podstawie indywidualnych danych z badania budżetów gospodarstw domowych.

W III skupieniu znalazły się wiejskie gospodarstwa domowe cechujące się największym udziałem wydatków na żywność i napoje bezalkoholowe w wydatkach ogółem (tudzież ich względnie wysokim poziomem), a także najwyższym poziomem i największym udziałem wydatków na napoje alkoholowe i wyroby tytoniowe (tabl. 1). W rozchodach ogółem zauważa się tutaj największy udział podatków i innych opłat.

W IV skupieniu znalazły się wiejskie gospodarstwa domowe z najmniejszymi rozchodami ogółem (tabl. 1). W skupieniu tym, z wyjątkiem wydatków przeznaczonych na edukację oraz rozchodów inwestycyjnych na gospodarstwo rolne, odnotowano najniższe rozchody na pozostałe kategorie. Cechą charakterystyczną skupienia był wysoki udział w wydatkach ogółem tych przeznaczonych na żywność i napoje bezalkoholowe, a także na odzież i obuwie oraz transport. To, że znaczna część wydatków została przeznaczona na żywność, w porównaniu z ich stosunkowo niskim poziomem w innych skupieniach, świadczy o niekorzystnej sytuacji materialnej tej części społeczeństwa. Jest to szczególnie niepokojące, gdyż skupienie IV obejmowało większość wiejskich gospodarstw domowych (ponad 69% ogółu).

W kolejnym etapie przeprowadzono analizę dyskryminacyjną celem powiązania cech wiejskich gospodarstw domowych z wyodrębnionymi grupami gospodarstw reprezentujących określoną wysokość i strukturę rozchodów. W analizie identyfikacyjnej uwzględniono 19 zmiennych określających cechy demograficzno-społeczne i ekonomiczne członków gospodarstw domowych oraz dotyczące mieszkania i użytków rolnych. Najbardziej dyskryminujące okazały się następujące zmienne (tabl. 2): liczba osób w gospodarstwie domowym, wiek głowy domu oraz dochód rozporządzalny na osobę. Duży wpływ w tym kontekście wykazały też zmienne zero-jedynkowe kodujące przynależność do grupy

emerytów oraz do II i V fazy cyklu życia rodziny, a także do gospodarstw prowadzonych przez osoby z wykształceniem podstawowym i zasadniczym zawodowym. Najmniejszą siłą wpływu wykazały zmienne zero-jedynkowe gospodarstwa z I fazy cyklu rozwoju rodziny oraz gospodarstwa utrzymujących się z niezarobkowych źródeł, a także zmienna określająca wielkość powierzchni gruntów rolnych.

Współczynnik lambda Wilksa dla utworzonego modelu wyniósł 0,625, przy $F(57;31)=93,63$ ($p<0,0000$). Nie jest to wartość zbyt wysoka, co oznacza, że istnieją inne zmienne, które mają związek z rozchodami w gospodarstwach domowych. W budżetach gospodarstw domowych dostępne zmienne informują o sytuacji materialnej i społeczno-zawodowej, o cechach demograficznych członków rodzin, a także o ich majątku. Na zachowania członków gospodarstw domowych wpływa również wiele zmiennych, zarówno o charakterze endogenicznym jak i egzogenicznym. Sposobem na podwyższenie wartości współczynnika lambda Wilksa jest zwiększenie liczby skupień, jednak w opracowaniu dążono do wyłonienia rozsądnej ich liczby, przy świadomości, że w rzeczywistości każde gospodarstwo domowe reprezentuje indywidualny profil konsumpcji.

TABL. 2. WYNIKI TESTU F DLA ZMIENNYCH RÓŻNICUJĄCYCH SKUPIENIA W 2004 R.

Wyszczególnienie	F	p
Liczba osób	1038	0,0000
Wiek głowy domu	562	0,0000
Dochód rozporządzalny na osobę	511	0,0000
GSE ^a — emeryci	390	0,0000
II faza cyklu rozwoju rodziny	302	0,0000
V faza cyklu rozwoju rodziny	177	0,0000
Wykształcenie głowy domu: podstawowe i poniżej	115	0,0000
zasadnicze zawodowe	114	0,0000
IV faza cyklu rozwoju rodziny	91	0,0000
Wykształcenie głowy domu — wyższe	86	0,0000
GSE — pracownicy na stanowiskach nierobotniczych	70	0,0000
III faza cyklu rozwoju rodziny	68	0,0000
GSE: pracownicy na stanowiskach robotniczych	64	0,0000
renciści	46	0,0000
Ogólna powierzchnia użytkowa mieszkania	41	0,0000
GSE — pracujący na własny rachunek	33	0,0000
Powierzchnia użytkowa rolnych	26	0,0000
GSE — niepracujący	24	0,0000
I faza cyklu rozwoju rodziny	15	0,0000

^a Grupa społeczno-ekonomiczna.

Źródło: jak przy tabl. 1.

Prawie wszystkie zmienne niezależne uwzględnione w wyjaśnianiu zmienności rozchodów w 2004 r., z wyjątkiem zmiennej zero-jedynkowej kodującej wiejskie gospodarstwa domowe rencistów, okazały się istotne statystycznie ($p<0,05$).

Wyodrębniono trzy funkcje dyskryminacyjne, wszystkie istotne dla $p < 0,005$. Modele powstające przez teoretyczny zabieg usuwania kolejnych pierwiastków wносиły coraz mniejsze wkłady w wyjaśnianie badanej zmienności. Wkłady kolejnych pierwiastków były swoiste i dotyczyły zmienności niewyjaśnionej przez poprzednie. Zdecydowanie największy wkład wniósł pierwszy pierwiastek. Współczynnik korelacji kanonicznej, który przyjmuje wartości z przedziału $<0, 1>$ i im jest wyższy, tym większa jest siła dyskryminacyjna danej funkcji, dla pierwszej funkcji wyniósł 0,568, dla drugiej — 0,257, a dla trzeciej — 0,108. Wszystkie trzy kanoniczne funkcje dyskryminacyjne okazały się istotne statystycznie, wartości p wyniosły poniżej 0,005. Skumulowany procent wariancji wskazał, że pierwsza funkcja dyskryminacyjna wyjaśniła 85,3% zmienności międzygrupowej, druga — 12,6%, a ostatnia — pozostałą część, czyli 2,1%.

TABL. 3. WSPÓLCZYNNIKI KORELACJI I DISKRYMINACYJNE FUNKCJI KANONICZNYCH DLA ZMIENNYCH PIERWSZEGO PIERWIASTKA MODELU W 2004 R.

Wyszczególnienie	Współczynniki	
	korelacji	dyskryminacyjne
Charakterystyka ogólna gospodarstw domowych		
Liczba osób	0,762	0,471
Dochód rozporządzalny na osobę	-0,536	-0,378
Powierzchnia użytków rolnych	0,056	-0,057
Ogólna powierzchnia użytkowa mieszkania	0,019	-0,100
Wiek głowy domu	-0,497	-0,138
Grupy społeczno-ekonomiczne (GSE)		
Pracownicy na stanowiskach robotniczych	-0,091	-0,059
Pracownicy na stanowiskach nierobotniczych	0,182	-0,042
Pracujący na własny rachunek	0,029	-0,034
Emeryci	-0,428	-0,143
Renciści	-0,097	-0,043
Niepracujący	0,105	0,050
Faza cyklu rozwoju rodziny		
I	-0,020	0,169
II	0,402	0,417
III	0,192	0,363
IV	0,222	0,394
V	-0,301	0,101
Wykształcenie głowy domu		
Podstawowe	-0,126	0,184
Zasadnicze zawodowe	0,243	0,082
Wyższe	-0,174	-0,115

U w a g a. Pogrubiono nazwy zmiennych, przy których bezwzględne wartości obu współczynników przekraczają 0,1.
Ź r ó d ł o: jak przy tabl. 1.

Pierwiastek pierwszy wartościami dodatnimi odróżnił skupienie IV od pozostałych, dla których średnie kanoniczne okazały się ujemne. Najważniejszymi zmiennymi w procesie dyskryminacyjnym dla pierwszego pierwiastka kanonicznego były wielkość gospodarstwa domowego, dochód rozporządzalny na osobę oraz wiek głowy domu (tabl. 3). Nieco mniejszą siłę wykazały zmienne zero-jedynkowe kodujące przynależność wiejskich gospodarstw domowych do faz życia rodziny z dziećmi na utrzymaniu oraz do gospodarstw, w których głową domu były osoby z wykształceniem podstawowym lub wyższym.

Wartości współczynników korelacji oraz dyskryminacyjnych wskazały, że wiejskie gospodarstwa domowe tworzące największe, tj. IV skupienie składały się z większej liczby domowników, przy czym głową domu były na ogół osoby w młodszym wieku oraz z mniejszymi dochodami rozporządzalnymi na osobę. W gospodarstwach tych — częściej niż w gospodarstwach domowych z pozostałych skupień — były dzieci na utrzymaniu, a osoba uzyskująca największe dochody na rzecz gospodarstwa częściej miała niższe wykształcenie, co najwyżej zasadnicze zawodowe (tabl. 4).

Z analizy struktury wydatków w gospodarstwach domowych wynika, że IV skupienie obejmowało ponad 2/3 ogółu wiejskich gospodarstw domowych i charakteryzowało się ubogim modelem konsumpcji. Wysokie udziały wydatków na żywność — przy ich ogólnym niskim poziomie — wskazują na niekorzystną sytuację materialną gospodarstw domowych z tego skupienia. Często były to rodziny z dziećmi na utrzymaniu, co można wiązać z większym obciążeniem budżetów domowych kosztami edukacji. Mimo wysokiego udziału wydatków związanych z kształceniem, poziom tej kategorii jest jednym z niższych wśród wyłonionych grup gospodarstw domowych.

Niskie wydatki na potrzeby wyższego rzędu (rekreacja i kultura oraz usługi gastronomiczne i zakwaterowanie) potwierdzają niekorzystną sytuację materialną gospodarstw domowych z IV skupienia. Należy jednak zwrócić uwagę na hierarchię potrzeb w tych gospodarstwach i rytm ich pracy zarobkowej. W rolniczych gospodarstwach domowych, przede wszystkim powiązanych z produkcją roślinną, praca zarobkowa przypada na czas letni. Okres wakacyjny nie sprzyja więc wyjazdom wypoczynkowym, gdyż wtedy priorytetem staje się uzyskanie dochodów na cały rok.

W omawianym skupieniu odnotowano także relatywnie wysokie udziały wydatków na rozchody bieżące związane z gospodarstwem rolnym przy najniższym poziomie tych wydatków. Może to wskazywać na niesfinansowanie wszystkich potrzeb pojawiających się w gospodarstwie rolnym.

Drugi pierwiastek oddzielił skupienia I i III wartościami dodatnimi od II i IV, które miały wartości ujemne. W tym pierwiastku największą siłę związku z dochodami wykazały zmienne wiek i wykształcenie głowy domu oraz dochód rozporządzalny na osobę. Wysoką siłą charakteryzowały się także zmienne związane z wielkością mieszkania, niektórymi fazami cyklu rozwoju rodziny oraz przynależnością do określonych grup społeczno-ekonomicznych.

TABL. 4. PODSUMOWANIE KLASYFIKACJI GOSPODARSTW DOMOWYCH W 2004 R.

Skupienie	Wysokie		Niskie		Wysokość dochodów/wielkość gospodarstwa domowego	Średni wiek głowy domu	Przeciętna ogólna powierzchnia użytkowa mieszkań/użytków rolnych	Faza cyklu życia rodziny	GSE
	wyposażenie, transport, usługi, rozchody finansowe, rozchody związane z gospodarstwem rolnym	udziały wydatków/rozchodów na	transport, rekreację i kulturę, edukację, napoje alkoholowe i wyroby tytoniowe	rozchody inwestycyjne na gospodarstwo rolne					
I — bogate	wyposażenie, transport, usługi, rozchody finansowe, rozchody związane z gospodarstwem rolnym	udziały wydatków/rozchodów na	użytkowanie mieszkania		wysokie — 1047 zł/ /3 osoby	niski — 47,9	93 m ² /5,0 ha	młodsze — od I do IV	emeryci, rolnicy
II — starsze	zdrowie		transport, rekreację i kulturę, edukację, napoje alkoholowe i wyroby tytoniowe		średnie — 886 zł/ /2 osoby	wysoki — 65,9	69 m ² /1,2 ha	starsze — VI	emeryci, renciści
III — ryzykowne	napoje alkoholowe i wyroby tytoniowe, podatki i inne opłaty		rozchody inwestycyjne na gospodarstwo rolne		średnie — 897 zł/ /2,4 osoby	średni — 55,2	81,7 m ² /2,3 ha	starsze — V i VI	emeryci, renciści
IV — ubogie	żywność i napoje bezalkoholowe		użytkowanie mieszkań i nośniki energii		niskie — 483 zł/ /4,1 osoby	niski — 47,0	80,1 m ² /3,0 ha	z dziećmi — od II do IV	różne

Źródło: jak przy tabl. 1.

W drugim pierwiastku uwypuklone zostały przede wszystkim różnice między gospodarstwami domowymi ze skupień I i II. W tych z I skupienia wiek głowy domu był niższy niż w tych z II — odpowiednio 48 i 66 lat. W I skupieniu odnotowano też najniższy odsetek gospodarstw, w których głowa domu miała podstawowe wykształcenie (21% gospodarstw z I skupienia), a w II było to 61%. Gospodarstwa domowe z I skupienia cechowały się największą powierzchnią użytkową mieszkania (ponad 93 m²), podczas gdy w II skupieniu — tylko 69 m². W I skupieniu obserwowano też największy udział gospodarstw domowych reprezentujących I fazę cyklu życia rodziny (12% ogółu gospodarstw z I skupienia), podczas gdy II skupienie obejmowało najmniejszy udział tych gospodarstw (ok. 4% ogółu gospodarstw z II skupienia). Charakterystyczną cechą II skupienia okazał się niski udział gospodarstw domowych pracowników na stanowiskach robotniczych i nierobotniczych oraz pracujących na własny rachunek (ok. 8%). Gospodarstwa domowe emerytów stanowiły zaś w tym skupieniu 59% ogółu badanych, będąc tym samym dominującą grupą społeczno-zawodową.

II skupienie cechowało się starszym wiekiem głowy domu i związaną z tym częstszą przynależnością do grupy gospodarstw domowych, w których głównym źródłem utrzymania była emerytura. W II skupieniu obserwowano znaczący udział gospodarstw, w których głowa domu miała wykształcenie podstawowe. Ze względu na wiek członków gospodarstw domowych zdrowie stanowiło ważną pozycję w strukturze wydatków, pochłaniającą znaczną część budżetu domowego. Koszty związane z ochroną zdrowia w II skupieniu były parokrotnie większe niż w innych. Konkludując, II skupienie utworzyły gospodarstwa domowe osób starszych, co miało wpływ na hierarchię wydatków charakterystyczną dla tej grupy.

W trzecim pierwiastku skupienia I, II i IV oddzieliły się wartościami dodatnimi od skupienia III, które miało wartości ujemne. Największą siłą dyskryminacyjną w tym pierwiastku odznaczyły się zmienne zero-jedynkowe kodujące przynależność gospodarstw domowych do pracujących na własny rachunek oraz do V fazy cyklu życia rodziny. Wysokimi współczynnikami korelacji i dyskryminacji charakteryzowały się też zmienne: dochód rozporządzalny na osobę, powierzchnia użytków rolnych, przynależność do grupy emerytów i zasadnicze zawodowe wykształcenie głowy domu.

W III skupieniu — w porównaniu z innymi skupieniami — zauważono największy odsetek gospodarstw domowych z V fazy cyklu życia rodziny. Takie gospodarstwa stanowiły ok. 32% ogółu gospodarstw w III skupieniu, podczas gdy w pozostałych skupieniach ich udziały kształtowały się od 11% w IV skupieniu do 24% w I skupieniu. W III skupieniu, które obejmowało ok. 1/5 wiejskich gospodarstw domowych, znalazły się gospodarstwa domowe z relatywnie korzystną sytuacją materialną. W dużej części były to gospodarstwa osób starszych, aktywnych zawodowo, z usamodzielnionymi dziećmi.

IV skupienie miało największy odsetek poprawnie sklasyfikowanych obiektów, to jest ok. 92% ogółu zakwalifikowanych gospodarstw. Na drugim końcu skali znalazły się skupienia I i III, w których odnotowano odpowiednio 17% i 18% poprawnie sklasyfikowanych obiektów.

Trafność predykcji dla modeli rozchodów w wiejskich gospodarstwach domowych w roku 2004 wyniosła ok. 73%. Wynik ten, przy tak dużej liczbie gospodarstw domowych, można uznać za korzystny.

GRUPY GOSPODARSTW DOMOWYCH W 2012 R.

W modelu gospodarstw domowych w 2012 r. znalazły się dwa małe skupienia obejmujące 0,6% i 1,3% ogółu gospodarstw domowych oraz dwa większe — 25,8% i 72,3%. Po raz drugi wyłaniają się więc trzy relatywnie mniejsze skupienia i jedno duże.

Charakterystyczna cecha gospodarstw domowych zakwalifikowanych do I skupienia to największe rozchody ogółem, wynoszące miesięcznie 10614 zł na osobę (tabl. 5). Były one związane z wydatkowaniem znaczących kwot na potrzeby gospodarstwa rolnego. Rozchody bieżące i inwestycyjne na rolnictwo pochłaniały 81% rozchodów ogółem, podczas gdy dla ogółu gospodarstw zlokalizowanych na wsi wartość ta wynosiła 15%. W gospodarstwach domowych należących do I skupienia notowano nieco wyższy odsetek wydatków na odzież i obuwiu (6,9%), użytkowanie mieszkania i nośniki energii (12,8%) oraz transport (18,6%) w wydatkach ogółem. Niski odsetek w strukturze wydatków gospodarstw z I skupienia dotyczył dóbr i usług rekreacyjno-kulturalnych, edukacyjnych i restauracyjno-hotelarskich (6,1%).

W II skupieniu znalazły się gospodarstwa domowe z największymi rozchodami ogółem — 3719 zł na osobę miesięcznie (tabl. 5). W strukturze wydatków gospodarstw z II skupienia odnotowano niski udział żywności i napojów bezalkoholowych (12,6% wydatków ogółem) oraz użytkowania mieszkania i nośników energii (11,6%), a wysoki — na dobra i usługi rekreacyjno-kulturalne, edukacyjne oraz gastronomiczne i zakwaterowanie (łącznie 31,7% wydatków ogółem). Gospodarstwa domowe z II skupienia cechowały niskie rozchody na podatki i inne opłaty (1,5% w strukturze rozchodów ogółem) oraz bieżące związane z rolnictwem (2,6%), nie odnotowano natomiast rozchodów inwestycyjnych na gospodarstwo rolne.

TABL. 5. MIESIĘCZNE ROZCHODY NA OSOBĘ I ICH UDZIAŁ W ROZCHODACH OGÓŁEM W POSZCZEGÓLNYCH SKUPIENIACH W 2012 R.

Wyszczególnienie	Skupienia							
	I		II		III		IV	
	w zł	w %	w zł	w %	w zł	w %	w zł	w %
Żywność i napoje bezalkoholowe	358	23,9	385	12,6	320	27,4	230	30,3
Napoje alkoholowe i wyroby tytoniowe	25	1,7	45	1,5	19	1,6	23	3,0
Odzież i obuwiu	103	6,9	170	5,6	43	3,7	39	5,1
Użytkowanie mieszkania i nośniki energii	192	12,8	353	11,6	232	19,9	153	20,1
Wyposażenie mieszkania i prowadzenie gospodarstwa domowego	115	7,7	189	6,2	62	5,3	33	4,3
Zdrowie	66	4,4	138	4,5	68	5,8	33	4,3
Transport	279	18,6	315	10,3	131	11,2	79	10,4
Łączność	55	3,7	77	2,5	63	5,4	27	3,6
Rekreacja i kultura	75	5,0	706	23,2	78	6,7	44	5,8
Edukacja	8	0,5	46	1,5	10	0,9	6	0,8
Gastronomia i zakwaterowanie	9	0,6	212	7,0	21	1,8	12	1,6
Pozostałe wydatki na towary i usługi ...	95	6,3	239	7,8	65	5,6	58	7,6
Pozostałe wydatki	121	8,1	174	5,7	54	4,6	23	3,0
Wydatki ogółem	1501	100,0	3049	100,0	1166	100,0	760	100,0

TABL. 5. MIESIĘCZNE ROZCHODY NA OSOBĘ I ICH UDZIAŁ W ROZCHODACH OGÓŁEM W POSZCZEGÓLNYCH SKUPIENIACH W 2012 R. (dok.)

Wyszczególnienie	Skupienia							
	I		II		III		IV	
	w zł	w %	w zł	w %	w zł	w %	w zł	w %
Podatki i inne opłaty	12	0,1	57	1,5	94	5,9	19	1,9
Rozchody kapitałowe (rzeczowe)	356	3,4	52	1,4	23	1,5	9	0,9
Rozchody finansowe	144	1,4	465	12,5	149	9,5	77	7,7
Rozchody bieżące związane z gospodarstwem rolnym	6487	61,1	96	2,6	120	7,6	111	11,2
Rozchody inwestycyjne na gospodarstwo rolne	2114	19,9	—	—	26	1,7	18	1,8
Rozchody ogółem	10614	100,0	3719	100,0	1578	100,0	994	100,0

Źródło: jak przy tabl. 1.

W III skupieniu znalazły się gospodarstwa domowe, w których zaobserwowano względnie wysokie udziały wydatków na żywność i napoje bezalkoholowe (ponad 27,4% wydatków ogółem), użytkowanie mieszkania i nośniki energii (ok. 19,9%), a także na podatki i inne opłaty (5,9%) (tabl. 5).

W IV skupieniu gospodarstw domowych rozchody na większość dóbr i usług były najniższe (wydatki ogółem stanowiły 760 zł na osobę) (tabl. 5). Odnotowano wysoki udział wydatków na żywność i napoje bezalkoholowe (ponad 30% wydatków ogółem) oraz na użytkowanie mieszkania i nośniki energii (ponad 20%). Wspomniane potrzeby pochłaniały zatem ponad połowę wydatków ogółem gospodarstw domowych, co wskazuje na niekorzystną sytuację materialną znacznej części wiejskiej społeczności. Większy niż w pozostałych skupieniach był w tej grupie udział wydatków na napoje alkoholowe i wyroby tytoniowe (3,0% wydatków ogółem).

Analiza dyskryminacyjna wykorzystana do identyfikacji gospodarstw domowych reprezentujących modele rozchodów w 2012 r. wykazała, że wszystkie zmienne niezależne zastosowane w wyjaśnianiu zmienności rozchodów w wiejskich gospodarstwach domowych okazały się istotne statystycznie (tabl. 6). Największą wartość statystyki F miała zmienna określająca powierzchnię użytków rolnych, a następnie dochody rozporządzalne na osobę oraz wielkość gospodarstwa domowego. Dużym udziałem cechowały się też zmienne zero-jedynkowe kodujące przynależność gospodarstw domowych do grupy rolników oraz do grupy gospodarstw prowadzonych przez osoby z wyższym wykształceniem. Najmniejszy wkład w wyjaśnianie zmienności rozchodów wykazały zmienne zero-jedynkowe kodujące gospodarstwa domowe rencistów i utrzymujących się z pozostałych źródeł niezarobkowych.

TABL. 6. WYNIKI TESTU F DLA ZMIENNYCH RÓŻNICUJĄCYCH SKUPIENIA W 2012 R.

Wyszczególnienie	F	p
Wielkość powierzchni użytków rolnych w ha	419	0,0000
Dochód rozporządzalny na osobę	352	0,0000
Liczba osób	280	0,0000
GSE — rolnicy	276	0,0000
Wyższe wykształcenie głowy domu	205	0,0000

TABL. 6. WYNIKI TESTU *F* DLA ZMIENNYCH RÓŻNICUJĄCYCH SKUPIENIA W 2012 R. (dok.)

Wyszczególnienie	<i>F</i>	<i>p</i>
Ogólna powierzchnia mieszkania	107	0,0000
GSE: pracownicy na stanowiskach nierobotniczych	83	0,0000
pracownicy na stanowiskach nierobotniczych	80	0,0000
Podstawowe wykształcenie głowy domu	59	0,0000
GSE — pracujący na własny rachunek	53	0,0000
Zasadnicze wykształcenie głowy domu	49	0,0000
Wiek głowy domu	47	0,0000
II faza cyklu życia rodziny	38	0,0000
GSE — emeryci	36	0,0000
I faza cyklu życia rodziny	24	0,0000
III faza cyklu życia rodziny	17	0,0000
IV faza cyklu życia rodziny	15	0,0000
V faza cyklu życia rodziny	12	0,0000
GSE: osoby utrzymujące się z pozostałych niezarobko- wych źródeł utrzymania	5	0,0032
renciści	4	0,0060

Źródło: jak przy tabl. 1.

Współczynnik lambda Wilksa dla utworzonego modelu wyniósł 0,767, co oznacza, że zmienne niezależne wprowadzone do modelu wyjaśniły niewielką część zmienności rozchodów w wiejskich gospodarstwach domowych. Wyodrębniono trzy funkcje dyskryminacyjne, wszystkie były istotne ($p < 0,005$). Największy wkład w wyjaśnianie zmienności wniósł pierwszy pierwiastek.

Skumulowany procent wariancji wskazał, że pierwsza funkcja dyskryminacyjna wyjaśniła 56,5% zmienności międzygrupowej, druga — 33,9%, a ostatnia — pozostała część, czyli 9,6% niewyjaśnionej wcześniej zmienności międzygrupowej.

Pierwszy pierwiastek kanoniczny oddzielił skupienie IV wartościami dodatnimi od pozostałych skupień. Najważniejszymi zmiennymi w procesie dyskryminacyjnym dla pierwszego pierwiastka kanonicznego były wielkość gospodarstwa domowego, dochód rozporządzalny na osobę oraz przynależność do III fazy cyklu życia rodziny z małymi dziećmi (tabl. 7). Nieco mniejszą siłę wykazały takie zmienne, jak podstawowe wykształcenie oraz wiek głowy domu.

W IV skupieniu znalazły się gospodarstwa domowe o (tabl. 8):

- stosunkowo dużej liczbie osób — średnio 3,5 osoby (w pozostałych skupieniach wartość ta kształtowała się od 2,5 do 3,4);
- najmniejszych dochodach rozporządzalnych na osobę, średnio miesięcznie 1034 zł (w pozostałych skupieniach było to od 1037 zł do 3028 zł).

W drugim pierwiastku skupienie I oddzieliło się od II i IV. Największą siłę wykazały zmienne związane z powierzchnią mieszkania i użytków rolnych oraz z gospodarstwami domowymi rolników.

W I skupieniu zaobserwowano największą powierzchnię zajmowanego mieszkania (148 m²; w pozostałych skupieniach było to od 96 do 132 m²) oraz gruntów rolnych (średnio ok. 34 ha; w pozostałych — od 1,0 ha do 2,4 ha).

Trzeci pierwiastek oddzielił skupienia II i I od III. W trzecim pierwiastku kanonicznym największą siłą cechowały się następujące zmienne: wielkość gospodarstwa domowego, dochód rozporządzalny na osobę, wiek głowy domu oraz wyższe wykształcenie głowy domu.

W II skupieniu znalazły się gospodarstwa domowe o stosunkowo małej liczbie osób (średnio 2,5 osoby), największym dochodzie rozporządzalnym na osobę (3028 zł) oraz z młodszymi głowami domu (47,1 roku).

Charakterystyka III skupienia jest natomiast następująca: gospodarstwa domowe z nieco większymi dochodami rozporządzalnymi na osobę (1481 zł) niż w skupieniu I i IV oraz z najstarszymi głowami domu (53,4 roku). Z tego powodu stosunkowo częstym źródłem utrzymania — obok najemnej pracy zarobkowej — były tam emerytury.

TABL. 7. WSPÓLCZYNNIKI KORELACJI I DYSKRYMINACYJNYCH FUNKCJI KANONICZNYCH DLA ZMIENNYCH PIERWSZEGO PIERWIĄSTKA MODELU W 2012 R.

Wyszczególnienie	Współczynniki	
	korelacji	dyskryminacyjne
Charakterystyka ogólna gospodarstw domowych		
Liczba osób	0,552	0,606
Liczba niepełnosprawnych	0,155	0,026
Dochód rozporządzalny na osobę	-0,624	-0,383
Wiek głowy domu	-0,148	-0,062
Liczba użytków rolnych w ha	0,058	-0,009
Ogólna powierzchnia mieszkania	-0,300	-0,279
Wykształcenie głowy domu		
Podstawowe	0,262	0,359
Zasadnicze	0,241	0,214
Wyższe	-0,462	-0,209
GSE		
Pracownicy na stanowiskach robotniczych	0,283	-0,242
Pracownicy na stanowiskach nierobotniczych	-0,302	-0,269
Rolnicy	0,150	-0,096
Pracujący na własny rachunek	-0,193	-0,282
Emeryci	-0,117	-0,206
Renciści	0,023	-0,112
Utrzymujący się z niezarobkowych źródeł	0,039	-0,014
Fazy cyklu życia rodziny		
I	-0,138	0,009
II	0,156	0,142
III	0,126	0,020
IV	0,128	-0,030
V	-0,111	0,048

U w a g a. Pogrubiono nazwy zmiennych, przy których bezwzględne wartości obu współczynników przekraczają 0,1.
Ź r ó d ł o: jak przy tabl. 1.

Na podstawie macierzy błędnych klasyfikacji można stwierdzić, że IV skupienie miało największy odsetek poprawnie sklasyfikowanych obiektów — 94%. W skupieniach II i III odnotowano natomiast najmniej poprawnie sklasyfikowanych obiektów — odpowiednio ok. 30% i 14%. Trafność predykcji dla modelu rozchodów w wiejskich gospodarstwach domowych za 2012 r. wyniosła ponad 72%.

TABL. 8. PODSUMOWANIE KLASYFIKACJI GOSPODARSTW DOMOWYCH W 2012 R.

Skupienie	Wysokie		Niskie		Wysokość dochodów/wielkość gospodarstwa domowego	Średni wiek głowy domu	Przeciętna ogólna powierzchnia użytkowa mieszkań/użytków rolnych	Faza cyklu życia rodziny	GSE
	Wysokie	Niskie	Wysokie	Niskie					
I — inwestujące	odzież i obuwie, transport, rozchody kapitałowe, rozchody związane z gospodarstwem rolnym	rozkłady finansowe	niskie — 1037 zł/ /3,4 osoby	niski — 46,6	148 m ² /33,8 ha	IV i V	rolnicy		
II — bogate	usługi, rozchody finansowe	żywność i napoje bezalkoholowe, użytkowanie mieszkań, rozchody na gospodarstwo rolne	wysokie — 3028 zł/ /2,5 osoby	niski — 47,1	132 m ² /1,0 ha	V	pracownicy na stanowiskach nirobotniczych, pracujący na własny rachunek		
III — starsze	żywność i napoje bezalkoholowe, podatki i inne opłaty	odzież i obuwie	średnie — 1481 zł/ /2,7 osoby	wysoki — 53,4	107 m ² /2,0 ha	V i VI	emeryci, pracownicy na stanowiskach robotniczych i nirobotniczych		
IV — ubogie	żywność i napoje bezalkoholowe, użytkowanie mieszkania	rozkłady kapitałowe	niskie — 1034 zł/ /3,5 osoby	średni — 50,3	96 m ² /2,4 ha	IV i V	pracownicy na stanowiskach robotniczych, emeryci		

Źródło: opracowanie własne na podstawie wyników badania.

DYSKUSJA WYNIKÓW

Konsumpcję większości gospodarstw domowych zlokalizowanych na wsi cechował wysoki udział wydatków na potrzeby podstawowe (wyżywienie i utrzymanie mieszkania). Inni badacze (Chmielewska, 2013) również podkreślają, że mimo wielu działań kohezyjnych poziom życia i wyposażenia infrastrukturalnego na wsi nadal odbiega od warunków miejskich. Wiejskie gospodarstwa domowe są narażone na większe ryzyko ubóstwa i mają mniejsze szanse wyjścia z niego (Sączewska-Piotrowska, 2016). Wartości wskaźników ubóstwa na wsi są dużo większe niż w miastach (Tarkowska, 2010).

Porównanie w latach 2004 i 2012 sytuacji w ubogich gospodarstwach domowych, które stanowiły największą grupę gospodarstw wiejskich, wykazuje spadek udziału wydatków na wyżywienie w wydatkach ogółem, a wzrost na towary i usługi wyższego rzędu, co świadczy o poprawie sytuacji materialnej. W innych badaniach wskazuje się (Chmielewska, 2013a; Kozera, Stanisławska i Wysocki, 2014), że po wstąpieniu Polski do Unii Europejskiej — choć nie zaszły rewolucyjne zmiany w poziomie życia mieszkańców wsi — społeczność wiejska zaczęła w szerszym zakresie korzystać z dóbr kulturalnych i rekreacyjnych, wzrósł także odsetek gospodarstw domowych, w których własny samochód osobowy stał się podstawowym środkiem transportu. Jak zauważyła Utzig (2015), wzrost dochodów wiejskich gospodarstwach domowych w mniejszym stopniu wpływał jednak na zwiększenie wydatków służących zaspokajaniu wyższych potrzeb. Mogło to wynikać z mniejszej dostępności kultury i rekreacji, a także z różnic w preferencjach i obieranych modelach życia.

Podsumowanie

Na podstawie uzyskanych wyników można stwierdzić, że wiejskie gospodarstwa domowe w różny sposób — pod względem poziomu i struktury — rozdysponowują swoje przychody. Ponad połowa z nich charakteryzuje się ubogim modelem konsumpcji, gdzie priorytet stanowią wydatki o znaczeniu podstawowym (żywność i napoje bezalkoholowe). Funkcjonują też mniejsze grupy gospodarstw domowych z charakterystyczną dla siebie konsumpcją. Można wskazać gospodarstwa domowe osób starszych, w których głównym źródłem utrzymania są emerytury, gdzie relatywnie duże udziały mają wydatki na ochronę zdrowia. Niewielka liczba wiejskich gospodarstw domowych cechuje się też wysokimi rozchodami — względnie wysokie ich odsetki przypadają tam na usługi oraz rozchody finansowe. Gospodarstwa te są tworzone przez stosunkowo młode osoby. W 2012 r. wyłoniono także grupę gospodarstw domowych (zwaną inwestującą) — zakładanych przez relatywnie młodych ludzi — w których odnotowano ponadprzeciętną wielkość gospodarstwa rolnego. W budżetach domowych tej grupy wysokie udziały wydatków dotyczyły potrzeb odzieżowo-obuwniczych, transportu, a także rozchodów kapitałowych i na gospodarstwo rolne.

Duża część wiejskich gospodarstw domowych charakteryzowała się niskimi dochodami przypadającymi na osobę i niskimi wydatkami na dobra i usługi. Jest to tym bardziej niepokojące, że są to rodziny z dziećmi na utrzymaniu. Ważną kwestią pozostaje więc monitorowanie wydatków tej części gospodarstw domowych pod kątem zaspokojenia potrzeb.

dr inż. Marlena Piekut — *Politechnika Warszawska*

LITERATURA

- Chmielewska, B. (2013a). Zmiany poziomu i struktury wydatków gospodarstw domowych jako wyraz przemian społecznych na wsi. *Journal of Agribusiness and Rural Development*, nr 2 (28), s. 19—31.
- Chmielewska, B. (2013b). *Ekonomiczno-społeczna sytuacja gospodarstw domowych rolników po akcesji Polski do Unii Europejskiej*. Warszawa: Instytut Ekonomiki Rolnictwa i Gospodarki Żywnościowej — Państwowy Instytut Badawczy.
- Dulli, S., Furini, S., Peron, E. (2009). *Data mining: metodi e strategie*. Mediolan: Springer Science & Business Media.
- Everitt, R.S., Landau, S., Leese, M., Stahl, D. (2011). *Cluster Analysis*. 5th Edition. John Wiley & Sons.
- Gatnar, E. (1995). *Klasyfikacja danych za pomocą pakietu statystycznego SPSS for Windows*. Warszawa: Wydawnictwo PLJ.
- GUS (2016). *Sytuacja gospodarstw domowych w 2015 r. w świetle wyników badania budżetów gospodarstw domowych*. Warszawa.
- Johnson, R.A., Wichern, D.W. (1992). *Applied multivariate statistical analysis*. New Jersey: Prentice-Hall.
- Kozera, A., Stanisławska, J., Wysocki, F. (2014). Sytuacja finansowa gospodarstw domowych zamieszkujących obszary wiejskie w Polsce po wstąpieniu do Unii Europejskiej. *Roczniki Naukowe Ekonomii Rolnictwa i Rozwoju Obszarów Wiejskich*, nr 101(2), s. 91—101.
- Leszczyńska, M. (2007). Zróżnicowanie dochodów ludności rolniczej i ich uwarunkowania. *Prace Naukowe Wydziału Ekonomii Uniwersytetu Rzeszowskiego. Monografie i Opracowania*, nr 2.
- Młodak, A. (2006). *Analiza taksonomiczna w statystyce regionalnej*. Warszawa: Difin.
- Panek, T. (2009). *Statystyczne metody wielowymiarowej analizy porównawczej*. Warszawa: Oficyna Wydawnicza SGH.
- Podolec, B., Ulman, P. (2008). Modele regresyjne w analizie wydatków gospodarstw domowych. *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie*, nr 790, s. 5—25.
- Rószkiewicz, M. (2002). *Narzędzia statystyczne w analizach marketingowych*. Warszawa: Wydawnictwo C.H. Beck.
- Sączewska-Piotrowska, A. (2016). Dynamika ubóstwa w miejskich i wiejskich gospodarstwach domowych. *Wiadomości Statystyczne*, nr 7, s. 39—59.
- Sokołowski, A. (2004). *Analizy wielowymiarowe*. Kraków: StatSoft.
- Tarkowska, E. (2010). Ubóstwo dzieci w Polsce. *Biuletyn Informacyjny PKESPU*, nr 6.
- Utzig, M. (2015). Wydatki konsumpcyjne wiejskich i miejskich gospodarstw domowych a zmiany ich dochodów. *Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu*, nr 17(4), s. 330—335.
- Utzig, M. (2014). Konwergencja dochodowa ludności wiejskiej i miejskiej w Polsce. *Roczniki Naukowe Ekonomii Rolnictwa i Rozwoju Obszarów Wiejskich*, nr 101(4), s. 144—152.

Wrzochalska, A. (2005). *Wybrane cechy społeczno-ekonomiczne ludności wiejskiej a rozwój wsi i rolnictwa*. Warszawa: Instytut Ekonomiki Rolnictwa i Gospodarki Żywnościowej — Państwowy Instytut Badawczy.

Summary. *The aim of the study is to isolate groups of rural households with similar outgoings and to describe them by socio-demographic and economic characteristics. It was carried out using multivariate statistical methods such as k-means cluster and discriminant function analysis. Data from the CSO survey of household budgets for the years 2004 and 2012 were used for the research purpose. The research resulted in the division of rural households into four groups considering the outgoings, where one group covered more than 2/3 of the households. Variables which discriminated the membership of rural households to certain groups to the largest extent were the number of people in the household and disposable income per capita.*

Keywords: consumption, expenditure, household budget, multivariate statistical methods.

Romana GŁOWICKA-WOŁOSZYN
Agnieszka KOZERA
Feliks WYSOCKI

Identyfikacja wewnętrznych uwarunkowań samodzielności i atrakcyjności finansowej gmin wiejskich województwa wielkopolskiego z wykorzystaniem biplotu

Streszczenie. *Badanie ma na celu wyodrębnienie i scharakteryzowanie grup gmin wiejskich województwa wielkopolskiego o zbliżonej samodzielności finansowej i podobnych uwarunkowaniach wewnętrznych w 2013 r., a ponadto pozwala na określenie relacji pomiędzy wskaźnikami finansowymi i cechami społeczno-ekonomicznymi. W badaniu posłużono się narzędziem podwójnego wykresu — biplotem PCA (the Principal Component Analysis). Podstawę informacyjną badania stanowiły dane z Banku Danych Lokalnych GUS oraz z Ministerstwa Finansów. Stwierdzono, że największa samodzielność finansowa cechowała grupę gmin wiejskich położonych w bezpośrednim sąsiedztwie stolicy województwa. Gminy te charakteryzowały się najwyższą oceną takich cech, jak: saldo migracji, odsetek mieszkańców korzystających z kanalizacji i instalacji gazowych, wskaźniki zagęszczenia oraz wskaźnik wykorzystania miejsc noclegowych.*

Słowa kluczowe: gminy wiejskie, samodzielność finansowa, atrakcyjność finansowa, biplot.

JEL: C38, H72

Istotą funkcjonowania gmin jest m.in. zaspokajanie podstawowych potrzeb społeczności lokalnych, jak również podejmowanie inicjatyw służących zwiększaniu zakresu świadczonych usług i podwyższaniu ich standardu, zapewnienie odpowiednich warunków do rozwoju działalności gospodarczej oraz podejmowanie działań na rzecz podnoszenia konkurencyjności regionu (Jastrzębska,

2004, s. 105; Heller i Farelnek, 2013, s. 84; Głowicka-Wołoszyn i Wysocki, 2014, s. 35). Podstawą działania samorządu terytorialnego jest: jego osobowość prawna, określony zakres uprawnień, własne mienie i źródła dochodów. Upodmiotowienie samorządu gminnego wiąże się z przekazaniem gminom samodzielności¹ w podejmowaniu decyzji o wykonywanych zadaniach (dotyczy to zwłaszcza zadań własnych gmin). Możliwość swobodnego decydowania — zgodnie z obowiązującymi przepisami prawa — przez organy samorządu gminnego o uzyskiwaniu i wydatkowaniu środków finansowych oraz wykonywaniu budżetu określana jest mianem samodzielności finansowej² (Patrzalek, 2010, s. 68; Jastrzębska, 2004, s. 106). Jest ona uzależniona od wielkości własnych środków finansowych i majątku oraz stanowi podstawę samorządności lokalnej (Surówka, 2013, s. 21; Zawora, 2010, s. 138). Samodzielność finansowa gmin związana jest więc m.in. z zapewnieniem stałych i wydajnych źródeł dochodów własnych oraz odpowiednim inwestowaniem środków, aby zapewnić wzrost atrakcyjności finansowej i harmonijny rozwój gminy.

Do podstawowych mierników samodzielności finansowej gmin zalicza się najczęściej (Heller, 2006, s. 138; Jaworska i Kożuch, 2011, s. 223; Duś, 2011, s. 74; Filipiak, 2009, s. 196):

- dochody własne w przeliczeniu na mieszkańca gminy — określające możliwości gminy w podejmowaniu decyzji o wydatkowaniu tych środków;
- udział dochodów własnych w dochodach budżetowych ogółem (wskaźnik samodzielności finansowej wydatkowej I stopnia) — informujący o stopniu swobody rozdysponowywania dochodów przez JST; wzrost wartości tego wskaźnika zwiększa możliwości decyzyjne gmin w zakresie przeznaczenia dochodów własnych.

Analizując samodzielność dochodową gmin, należy również zwrócić uwagę na konstytucyjne zagwarantowanie im dochodów w formie subwencji, które charakteryzują się niezależnością od źródeł znajdujących się w gminie. Są one bowiem formą ogólnego zasilenia budżetu gminy przez budżet państwa (Feret, 2013, s. 136 i 137). Subwencje ogólne mają charakter dofinansowania nieoznaczonego, a więc gminy decydują o sposobie ich wykorzystania. Stanowią zatem źródło dochodów, które wspomaga samodzielność wydatkowania samorządów lokalnych. W literaturze przedmiotu proponowany jest również wskaźnik samo-

¹ Samodzielność jednostki samorządu terytorialnego (JST) może być rozpatrywana na różnych płaszczyznach (Patrzalek, 2010, s. 67 i 68; Zawora, 2010, s. 137; Jaworska i Kożuch, 2011, s. 220; Jańczuk, 2013, s. 79—92), tj. w ujęciu: prawnym, politycznym, organizacyjnym i gospodarczym.

² W literaturze wskazywane są następujące aspekty samodzielności finansowej gmin (Jastrzębska, 2004, s. 108—110; Patrzalek, 2010, s. 69—81; Kosek-Wojnar, 2006, s. 77—86; Feret, 2013, s. 126—147): dochodowa (gminy posiadają uprawnienia i kompetencje w zakresie zwiększania dochodów własnych, pozyskiwania ich z różnych źródeł i różnymi metodami), wydatkowa (swoboda decyzji władz gmin co do poziomu i kierunku wydatków, a także sposobu i formy finansowania usług publicznych) i budżetowa (uprawnienia do zasad i trybu procedury budżetowej, czyli: planowania, uchwalania, wykonywania i kontroli wykonania budżetu gmin).

dzielności finansowej gmin wyrażający udział w dochodach budżetowych gmin sumy dochodów własnych i subwencji (Zawora, 2010, s. 138; Dylewski, Filipiak i Gorzałczyńska-Koczkodaj, 2004).

Władze gmin poprzez kształtowanie polityki finansowej mają możliwość zwiększania zarówno ich atrakcyjności inwestycyjnej, jak i osadniczej. Wiąże się to ze zdolnością kreowania rozwoju lokalnego, co może skutkować zwiększeniem możliwości przyciągania inwestorów i nowych mieszkańców (Filipiak, 2009, s. 196).

Wśród miar atrakcyjności finansowej gmin najczęściej w literaturze wymieniane są: wskaźnik określający wielkość wydatków inwestycyjnych przypadających na mieszkańca gminy (Filipiak, 2009, s. 198) i wskaźnik względny, podający udział wydatków inwestycyjnych w wydatkach ogółem (Zawora, 2010, s. 138; Jaworska i Kożuch, 2011, s. 227; Jastrzębska, 2012, s. 56).

Samodzielność finansowa gmin uwarunkowana jest czynnikami zewnętrznymi, takimi jak: koniunktura gospodarcza w kraju, stopy procentowe czy regulacje prawne, które m.in. określają podział zadań między państwo i samorząd, regulują system dochodów i wydatków oraz ustalają deficyt budżetowy (Patrzalek, 2010, s. 75—77; Zawora, 2010, s. 140). Ważnym czynnikiem oddziałującym na potencjał dochodowy gmin, niezależnym od władz lokalnych, jest położenie, zwłaszcza w odniesieniu do dużych miast (Zawora, 2010, s. 140), szczególnie w przypadku gmin wiejskich (Duś, 2011, s. 86; Głowicka-Wołoszyn i Wysocki, 2014, s. 38).

Samodzielność finansowa gmin uzależniona jest także od uwarunkowań wewnętrznych (Zawora, 2010, s. 140 i 141; Głowicka-Wołoszyn i Wysocki, 2014, s. 36 i 37): sytuacji demograficznej (np. liczby ludności, migracji) i społecznej (np. liczba bezrobotnych wśród ludności w wieku produkcyjnym), a także cech związanych z zagospodarowaniem infrastrukturalnym (np. sieć kanalizacyjna i gazowa) oraz sytuacji gospodarczej (np. liczba podmiotów gospodarczych). Cechy te wpływają na możliwość uzyskania dochodów własnych gminy oraz na wielkość wydatków, m.in. inwestycyjnych w zakresie infrastruktury technicznej.

W ocenie samodzielności finansowej gmin występują duże dysproporcje związane z ich typem administracyjnym. Gminy wiejskie w porównaniu z miejsko-wiejskimi i miejskimi cechuje mniejsza samodzielność i dużo większe jej zróżnicowanie (Heller, 2006, s. 145 i 146).

Wśród metod analizy wielowymiarowej, które umożliwiałyby pełniejsze zrozumienie powiązań pomiędzy samodzielnością i atrakcyjnością finansową gmin a czynnikami demograficzno-społecznymi, ekonomicznymi i geograficznymi, na uwagę zasługuje narzędzie graficzne nazywane biplotem, czyli podwójnym wykresem, wprowadzone przez Gabriela (1971). Na takim wykresie punkty ilustrują obserwacje, natomiast zmienne przedstawione są jako wektory o wspólnym początku. Biplot pozwala nie tylko na identyfikację czynników wpływających na badane zjawisko, ale również na określenie powiązań pomiędzy analizowanymi wskaźnikami finansowymi i uwarunkowaniami endogenicznymi gmin a obiektami, czyli gminami.

Celem artykułu jest opisanie relacji między gminami wiejskimi województwa wielkopolskiego a wskaźnikami charakteryzującymi ich sytuację finansową (samodzielność i atrakcyjność finansowa) oraz przestrzenno-przyrodniczą, demograficzno-społeczną, gospodarczą, jak również w zakresie wyposażenia infrastrukturalnego.

Na tej podstawie podjęto próbę wyodrębnienia i scharakteryzowania grup gmin wiejskich o zbliżonych uwarunkowaniach samodzielności i atrakcyjności finansowej.

Celem autorów było również podkreślenie zalet zastosowanego narzędzia diagnostycznego (biplotu), które stwarza możliwość pogłębionej analizy zależności między gminami a opisującymi je cechami. W konsekwencji prowadzi to do bardziej przejrzystej interpretacji wyników.

MATERIAŁ I METODY BADAWCZE

Badania przeprowadzono na podstawie danych GUS dotyczących gmin wiejskich województwa wielkopolskiego za 2013 r., pobranych z Banku Danych Lokalnych. Do obliczeń wykorzystano program Statistica oraz udostępniany do niego skrypt BiPlot (<http://www.statsoft.pl/Zasoby/Do-pobrania/Skrypty-VB>).

Jako narzędzie badawcze zastosowano biplot PCA (*the Principal Component Analysis*) (Gower i Hand, 1996; Gower, Lubbe i le Roux, 2010; Greenacre, 2010; Sagan, 2004; Wysocki, 2010), który umożliwia (na jego diagramie) analizę położenia względem siebie obiektów, czyli gmin oraz cech, czyli zarówno wskaźników finansowych, jak i cech reprezentujących uwarunkowania wewnętrzne. Pozwalało to na:

- wyodrębnienie grup gmin najbardziej podobnych do siebie ze względu na cechy aktywne, o czym świadczą małe odległości pomiędzy punktami reprezentującymi gminy, ponieważ długości odcinków aproksymują odległość Mahalanobisa pomiędzy gminami w przekroju wszystkich cech;
- poznanie typów funkcjonalnych gmin, określonych relacjami między gminami i cechami, poprzez rzuty prostopadłe punktów reprezentujących gminy na wektory odpowiadające cechom; na podstawie odległości tych rzutów od środka ciężkości możliwe jest określenie wartości ocen gmin w przekroju każdej cechy;
- analizę powiązań pomiędzy wskaźnikami samodzielności i atrakcyjności finansowej gmin a ich determinantami społeczno-ekonomicznymi; siła powiązań pomiędzy cechami jest aproksymowana przez kosinusy kątów pomiędzy wektorami reprezentującymi wskaźniki finansowe i ich uwarunkowania w gminach.

Na biplotcie wektory wychodzące ze środka ciężkości reprezentują wskaźniki finansowe i uwarunkowania przestrzenno-przyrodnicze, demograficzne, społeczne i ekonomiczne.

Wyboru cech dokonano na podstawie przesłanek merytorycznych, literatury przedmiotu oraz po analizie uzyskanych wstępnie wyników biplotów, poprzez

usunięcie ze zbioru danych tych cech, które miały niewielkie znaczenie z punktu widzenia ich mocy dyskryminacyjnej.

Do określenia samodzielności finansowej przyjęto następujące wskaźniki (cechy aktywne):

- udział dochodów własnych gmin w dochodach ogółem (DW/D),
- dochody własne gminy w zł *per capita* (DW/os.),
- udział subwencji w dochodach ogółem (Subw./D).

Atrakcyjność finansową określono udziałem wydatków majątkowych inwestycyjnych w wydatkach ogółem w latach 2011—2013 (WI/W). Ze względu na dużą zmienność tejsze cechy wyznaczono ją jako średnią z tych lat.

Ponadto uwzględniono cechy aktywne reprezentujące uwarunkowania społeczno-gospodarcze samodzielności i atrakcyjności finansowej gmin:

- saldo migracji w osobach w przeliczeniu na 1 tys. mieszkańców;
- odsetek mieszkańców korzystających z instalacji gazowej w %;
- liczbę pracujących w indywidualnych gospodarstwach rolnych w przeliczeniu na 100 mieszkańców gminy w wieku produkcyjnym³;
- udział gospodarstw indywidualnych o powierzchni powyżej 15 ha w ogólnej liczbie gospodarstw indywidualnych w %;
- liczbę noclegów udzielonych od stycznia do grudnia 2013 r. w przeliczeniu na 1 tys. mieszkańców gminy.

Liniowe kombinacje wymienionych cech aktywnych, maksymalizujące ilość wyjaśnianej całkowitej ich wariancji, utworzyły osie główne (główne składowe) biplotu.

Na diagramie dodatkowo przedstawiono cechy pasywne, które nie brały udziału w tworzeniu głównych składowych, ale pozwoliły pogłębić analizowane relacje:

- odległość od Poznania⁴ w km;
- liczba zarejestrowanych bezrobotnych w przeliczeniu na 100 mieszkańców gminy w wieku produkcyjnym;
- odsetek mieszkańców gminy korzystających z kanalizacji w %;
- powierzchnia użytków rolnych (UR) znajdujących się w granicach administracyjnych gminy w przeliczeniu na mieszkańca gminy w ha;
- udział gospodarstw indywidualnych o powierzchni 1—10 ha w ogólnej liczbie gospodarstw indywidualnych w %;
- wskaźnik Charvata — zagęszczenie bazy noclegowej na km²;
- udział dotacji w dochodach ogółem (Dot./D) w %.

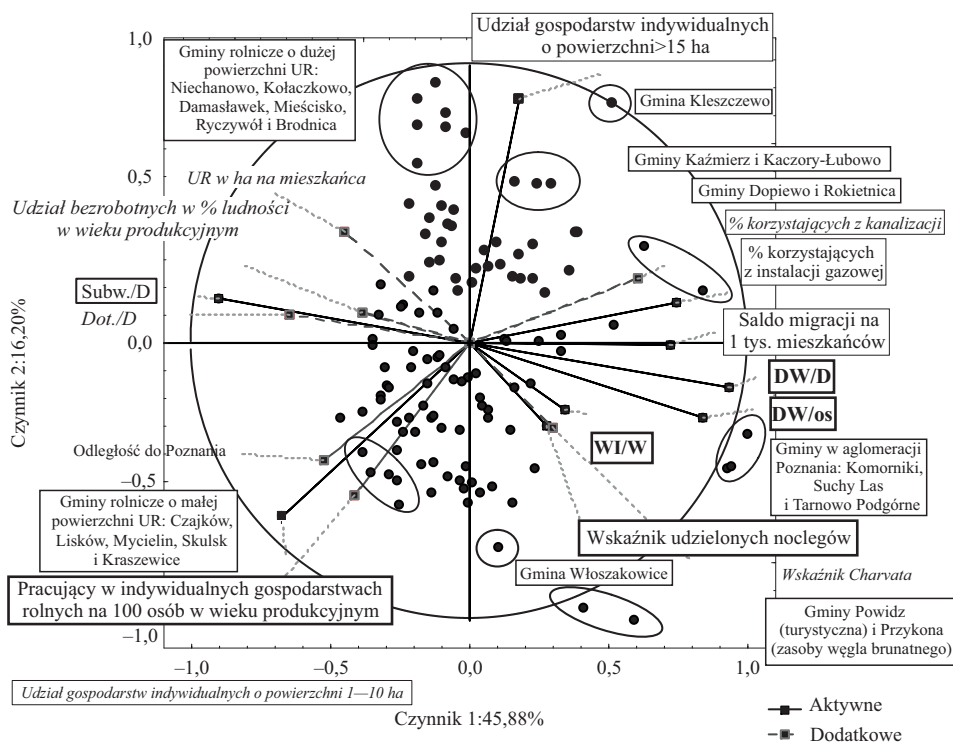
³ Liczba pracujących w indywidualnych gospodarstwach rolnych została oszacowana na podstawie liczby gospodarstw indywidualnych w gminie i przeciętnego wskaźnika pracujących w indywidualnych gospodarstwach rolnych w przeliczeniu na 1 gospodarstwo.

⁴ Odległości gmin od Poznania zostały obliczone jako odległości pomiędzy geograficznymi centrami gmin i Poznaniem z wykorzystaniem Google Maps.

WYNIKI BADAŃ

Z biplotu skonstruowanego dla gmin wiejskich w województwie wielkopolskim (wykres) wynika, że dwie osie główne, czyli nowe zmienne, które zostały utworzone z liniowych kombinacji cech aktywnych, wyjaśniły łącznie 62,08% całkowitej zmienności cech. Pierwsza składowa główna to 45,88% tej zmienności, a druga — tylko 16,20%.

**BIPILOT SPORZĄDZONY PO ANALIZIE SKŁADOWYCH GŁÓWNYCH
DLA SAMODZIELNOŚCI I ATRAKCYJNOŚCI FINANSOWEJ I JEJ UWARUNKOWAŃ
SPOŁECZNO-EKONOMICZNYCH W GMINACH WIEJSKICH WOJEWÓDZTWA
WIELKOPOLSKIEGO W 2013 R.**



U w a g a. DW/D — udział dochodów własnych gmin w dochodach ogółem; DW/os — poziom dochodów własnych gminy w złotych *per capita*; Subw./D — udział subwencji w dochodach ogółem; Dot./D — udział dotacji w dochodach ogółem.

Ź r ó d ł o: opracowanie własne na podstawie danych GUS pobranych z Banku Danych Lokalnych.

Na diagramie wektory wychodzące ze środka ciężkości zbioru punktów reprezentują wskaźniki samodzielności i atrakcyjności finansowej gmin i ich uwarunkowania wewnętrzne. Wskaźniki samodzielności finansowej gmin — DW/D,

a także DW/os. i Subw./D — okazały się cechami poprawnie reprezentowanymi przez punkty w dwuwymiarowym układzie głównych składowych. Długości tych wektorów były bowiem zbliżone do wartości 1. Również jedna z cech opisujących uwarunkowania ekonomiczne (pracujący w indywidualnych gospodarstwach rolnych w przeliczeniu na 100 mieszkańców gminy w wieku produkcyjnym) była poprawnie reprezentowana na diagramie.

Rozpatrywane wskaźnik atrakcyjności finansowej gmin (WI/W) oraz udział gospodarstw indywidualnych o powierzchni powyżej 15 ha UR w ogólnej liczbie gospodarstw indywidualnych okazały się natomiast cechami mniej dokładnie odwzorowanymi na wykresie.

Stwierdzono również, że najbardziej oddalona od środka ciężkości i jednocześnie położona najbliżej wektorów reprezentujących wskaźniki samodzielności i atrakcyjności finansowej była grupa składająca się z trzech gmin: Tarnowo Podgórne, Suchy Las i Komorniki, znajdujących się w bezpośrednim sąsiedztwie miasta Poznania. Rzuty prostopadłe punktów reprezentujących te gminy na przedłużeniu wektorów cech pozwoliły stwierdzić, że gminy te charakteryzowały się największą samodzielnością finansową, mierzoną największym udziałem dochodów własnych w dochodach ogółem i najwyższym poziomem dochodów własnych gmin *per capita*. Ponadto cechowała je pod tym względem najwyższa ocena atrakcyjności finansowej. Analizując położenie tej grupy względem wektorów reprezentujących uwarunkowania wewnętrzne, stwierdzono, że gminy te cechowały również najwyższe, dodatnie saldo migracji w przeliczeniu na 1 tys. mieszkańców oraz najlepsza dostępność sieci kanalizacyjnej i gazowej. Wiąże się to z sąsiedztwem aglomeracji miasta Poznania i postępującymi procesami zmian kategorii funkcjonalnej tych gmin wiejskich z rolniczej na mieszkalną. Jednocześnie stwierdzono tam najniższe oceny takich wskaźników, jak: powierzchnia UR znajdujących się w granicach administracyjnych gmin w przeliczeniu na mieszkańca, udział subwencji oraz dotacji w dochodach gmin ogółem i liczba zarejestrowanych bezrobotnych na 100 mieszkańców gminy w wieku produkcyjnym. W przypadku tej ostatniej cechy należy jednak zwrócić uwagę na jej mniej dokładne odwzorowanie przez punkty na wykresie.

Na biplotcie wyróżniono jeszcze dwie gminy — Dopiewo i Rokietnica, oraz oddzielnie zaznaczono gminę Kleszczewo. Są to gminy, które również znajdują się w obszarze bezpośredniego oddziaływania aglomeracji poznańskiej. Charakteryzowały się one relatywnie wysoką, w porównaniu z pozostałymi gminami, samodzielnością finansową, a ponadto cechowało je wysokie saldo migracji, jak również wysoki odsetek mieszkańców korzystających z infrastruktury kanalizacyjnej i gazowej, co wiąże się ze zjawiskiem suburbanizacji. Jednakże rzut punktu reprezentującego Kleszczewo na wektor cechy określającej udział gospodarstw indywidualnych o powierzchni powyżej 15 ha UR w ogólnej liczbie gospodarstw indywidualnych i jednocześnie duże oddalenie od środka ciężkości potwierdziły również jej podwójny charakter — rolniczy i mieszkalny.

W opozycji do pierwszych omówionych gmin znajdowała się wyróżniona na diagramie grupa gmin rolniczych, w skład której wchodziły: Niechanowo, Kołaczkowo, Damasławek, Mieścisko, Ryczywół i Brodnica. Rzuty prostopadłe reprezentujących je punktów na wektory cech opisujących udział dochodów własnych w dochodach ogółem oraz dochodów własnych przypadających na mieszkańca gminy były ujemne. Jednocześnie rzuty te nie były najbardziej oddalone od środka ciężkości, co świadczyło o relatywnie niskiej samodzielności finansowej. Tę grupę gmin charakteryzowały równocześnie wysokie wartości udziału subwencji w dochodach ogółem gmin.

Subwencja ogólna jako źródło dochodów uzupełniających dochody własne budżetów gmin, przekazywana jest z budżetu państwa w celu przeciwdziałania utrwalaniu się dysproporcji regionalnych w zaspokajaniu potrzeb społeczności, a jednocześnie służy stymulowaniu władz lokalnych do podnoszenia konkurencyjności regionu. Nie ogranicza ona samodzielności wydatkowej gmin w wykonywaniu i finansowaniu zadań publicznych, natomiast uzależnia gminy od finansów państwa — w przypadku tego źródła zasilania budżetów można zatem mówić o ograniczonej samodzielności dochodowej jednostek samorządowych. Tę grupę gmin, podobnie jak Kleszczewo, cechował najwyższy odsetek gospodarstw indywidualnych o powierzchni powyżej 15 ha UR. O rolniczym charakterze tej grupy gmin świadczyły również relatywnie wysokie zasoby UR w przeliczeniu na mieszkańca gminy, choć ta cecha dodatkowa była relatywnie mało dokładnie odwzorowana przez punkty na wykresie.

W III ćwiartce diagramu zaznaczono pięć punktów położonych najdalej od środka ciężkości, a stosunkowo blisko siebie, które reprezentują gminy o charakterze rolniczym: Czajków, Lisków, Mycielin, Skulsk i Kraszewice. Grupę tę wyróżniały najwyższe oceny następujących cech: udział pracujących w indywidualnych gospodarstwach rolnych na 100 mieszkańców gminy w wieku produkcyjnym, odsetek gospodarstw indywidualnych o powierzchni od 1 do 10 ha UR oraz odległość od miasta Poznania, przy relatywnie niskich ocenach wskaźników samodzielności finansowej. Były to gminy wiejskie położone we wschodniej lub południowo-wschodniej części województwa.

Na wykresie zaznaczono również gminy Przykona i Powidz, które nie znajdują się w zasięgu bezpośredniego oddziaływania Poznania, ale również wykazywały wysoką samodzielność finansową. Położenie gminy Włoszakowice także świadczy o relatywnie wysokiej ocenie jej samodzielności finansowej. W przypadku tych trzech gmin duże znaczenie mają uwarunkowania przyrodnicze, które przyczyniają się do ukierunkowania podejmowanej przez nie działalności gospodarczej na turystykę i wydobywanie węgla brunatnego, a w konsekwencji wpływają pozytywnie na ich sytuację finansową. Zasoby naturalne węgla brunatnego i związane z jego wydobywaniem opłaty eksploatacyjne przyczyniają się bezpośrednio do wzrostu dochodów własnych gminy Przykona (Głowicka-Wołoszyn i Wysocki, 2014, s. 40). W przypadku gmin Powidz i Włoszakowice

uwarunkowania przyrodnicze (jeziora, lasy) sprzyjają rozwojowi turystyki. Charakteryzowały się one najwyższymi wartościami wskaźnika wykorzystania bazy turystycznej w województwie wielkopolskim — ok. 6 tys. udzielonych noclegów w roku w przeliczeniu na 1 tys. mieszkańców⁵ — w porównaniu z przeciętną wartością w gminach wiejskich województwa wielkopolskiego wynoszącą 70 noclegów.

Wnioski

1. Największą samodzielnością i atrakcyjnością finansową charakteryzowała się grupa gmin wiejskich, które położone są w bliskim sąsiedztwie stolicy województwa. Cechowały się one również najwyższą oceną uwarunkowań demograficznych i gospodarczych obejmujących saldo migracji i odsetek mieszkańców korzystających z kanalizacji i instalacji gazowych. Ponadto gminy te wyróżniały wysokie wartości wskaźników zagęszczenia i wykorzystania miejsc noclegowych. Świadczy to nie tyle o turystycznej roli tych gmin, ile o pełnieniu przez nie funkcji bazy noclegowej dla turystyki i biznesu, rozwijających się w Poznaniu.

2. Specyficzne uwarunkowania przyrodnicze trzech gmin wiejskich — Przykona (zasoby węgla brunatnego) oraz Powidz i Włoszakowice (rozwój turystyki) wpłynęły na kierunki ich rozwoju gospodarczego oraz pozwoliły im osiągnąć wysoką samodzielność finansową.

3. W najmniej korzystnej sytuacji pod względem samodzielności i atrakcyjności finansowej znajdowały się zarówno gminy posiadające stosunkowo duże zasoby UR przypadających na mieszkańca (które ponadto skupione są w większych gospodarstwach indywidualnych — powyżej 15 ha UR), jak i te najbardziej oddalone od Poznania, położone we wschodniej części województwa, które charakteryzowały się wysokim udziałem pracujących w indywidualnych gospodarstwach rolnych i największym odsetkiem małych gospodarstw indywidualnych (poniżej 1 ha UR).

4. Biplot PCA umożliwił przeprowadzenie pogłębionej analizy zależności między obiektami (gminami) a cechami (oceną samodzielności i atrakcyjności finansowej oraz ich wewnętrznymi uwarunkowaniami). Okazał się przydatnym narzędziem do identyfikacji uwarunkowań samodzielności i atrakcyjności finansowej gmin wiejskich w województwie wielkopolskim.

dr Romana Głowicka-Wołoszyn, dr Agnieszka Kozera, prof. dr hab. Feliks Wysocki —
Uniwersytet Przyrodniczy w Poznaniu

⁵ Obliczenia własne na podstawie informacji z Banku Danych Lokalnych z 2013 r.

LITERATURA

- Duś, E. (2011). Wysokość i struktura dochodów gmin wiejskich jako miernik stopnia ich peryferyjności. W: M. Wesołowska (red.), *Wiejskie obszary peryferyjne — uwarunkowania i czynniki aktywizacji. Studia Obszarów Wiejskich 26*, s. 73—88. Warszawa: Wydawnictwo Bernardinum.
- Dylewski, M., Filipiak, B., Gorzałczyńska-Koczkodaj, M. (2004). *Analiza finansowa w jednostkach samorządu terytorialnego*. Warszawa: Municipium SA.
- Feret, E. (2013). Samodzielność finansowa jednostek samorządu terytorialnego gwarancją samorządowego działania społeczności lokalnych. W: W. Skrzydło, W. Szapował, K. Eckhardt, P. Steciuk (red.), *Konstytucyjne podstawy budowania i rozwoju społeczeństwa obywatelskiego w Polsce i na Ukrainie — dobre praktyki, t. 1*, s. 126—148. Seria: Prawo naszych sąsiadów, red. cyklu publikacji J. Połuszny. Artykuł powstał w ramach konkursu *Współpraca w dziedzinie dyplomacji publicznej*, realizowanego za pośrednictwem MSZ RP w 2013 r.
- Filiipiak, B. (2009). *Metodyka kompleksowej oceny gospodarki finansowej jednostki samorządu terytorialnego*. Warszawa: Difin.
- Gabriel, K.R. (1971). The Biplot Graphic Display of Matrices with Application to Principal Component Analysis. *Biometrics, Vol. 58, No. 3*, s. 453—467.
- Głowicka-Wołoszyn, R., Wysocki, F. (2014). *Uwarunkowania społeczno-ekonomiczne samodzielności finansowej gmin województwa wielkopolskiego. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, nr 346*, s. 34—44.
- Gower, J.C., Hand, D.J. (1996). *Biplots*. Londyn: Chapman & Hall.
- Gower, J.C., Lubbe, S., le Roux, N. (2010). *Understanding Biplots*. Nowy Jork: Wiley.
- Greenacre, M. (2010). *Biplots in Practice*. Bilbao: Fundacion BBVA.
- GUS. *Bank Danych Lokalnych*. Pobrane z: www.stat.gov.pl.
- Heller, J. (2006). Samodzielność finansowa samorządów terytorialnych w Polsce. *Studia Regionalne i Lokalne, nr 2 (24)*, s. 138.
- Heller, J., Farelnek, E. (2013). Finanse i samodzielność ekonomiczna a ustrój samorządów terytorialnych w Polsce. *Studia Regionalne i Lokalne, nr 2 (52)*, s. 81—94.
- Jańczuk, L. (2013). *Samodzielność jednostek samorządu terytorialnego jako podmiotów administracji publicznej w Polsce*. Wydawnictwo KUL.
- Jastrzębska, M. (2004). Samodzielność ekonomiczna i finansowa jednostek samorządu terytorialnego. *Ekonomia, nr 13*. Uniwersytet Warszawski, s. 100—112.
- Jastrzębska, M. (2012). *Finanse jednostek samorządu terytorialnego*. Warszawa: Wolters Kluwer Polska.
- Jaworska, M., Kożuch, A. (2011). Możliwości rozwoju lokalnego w kontekście zróżnicowania samodzielności finansowej podstawowych jednostek samorządu terytorialnego (na przykładzie gmin miejsko-wiejskich Polski Wschodniej). *Optimum. Studia Ekonomiczne, nr 6 (54)*, s. 218—233.
- Kosek-Wojnar, M. (2006). Samodzielność jednostek samorządu terytorialnego w sferze wydatków. *Zeszyty Naukowe Wyższej Szkoły Ekonomicznej w Bochni, nr 4*, s. 75—88.
- Patrzalek, L. (2010). *Finanse samorządu terytorialnego*. Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu.
- Sagan, A. (2004). *Jeden obraz ukazuje więcej niż 10 liczb, czyli jak budować mapy zadowolenia klienta z wykorzystaniem programu STATISTICA*. StatSoft, <http://www.statsoft.pl>.
- Surówka, K. (2013). *Samodzielność finansowa samorządu terytorialnego w Polsce*. Warszawa: Polskie Wydawnictwo Ekonomiczne.
- Wysocki, F. (2010). *Metody taksonomiczne w rozpoznawaniu typów ekonomicznych rolnictwa i obszarów wiejskich*. Wydawnictwo Uniwersytetu Przyrodniczego w Poznaniu.
- Zawora, J. (2010). Samodzielność finansowa samorządów gminnych Podkarpacia. *Zeszyty Naukowe SGGW. Ekonomika i organizacja gospodarki żywnościowej, nr 81*, s. 137.

Summary. *The research aims at isolating and describing groups of rural gminas in Wielkopolskie voivodship with comparable financial autonomy and similar internal determinants in 2013. Furthermore, it allows to examine relations between the financial indicators and the socio-economic determinants. The PCA (Principal Component Analysis) biplot was used in the research. Data published by the Local Data Bank of the CSO and the Ministry of Finance was applied as a study informational basis. It was found that the highest level of financial autonomy was observed in the group of rural gminas close to the capital city of the voivodship. These gminas were distinguished by such determinants as net migration, percentage of inhabitants using sewage and gas supply system, as well as density and occupancy rate of accommodation establishments.*

Keywords: rural gminas, financial autonomy, financial attractiveness, biplot.

INFORMACJE. PRZEGLĄDY. RECENZJE

XXXV Międzynarodowa Konferencja Naukowa *Wielowymiarowa Analiza Statystyczna 2016*

Od 7 do 9 listopada 2016 r. w Centrum Szkoleniowo-Konferencyjnym Uniwersytetu Łódzkiego odbyła się XXXV Międzynarodowa Konferencja Naukowa *Wielowymiarowa Analiza Statystyczna 2016* (WAS 2016) (*The 35th Annual Conference on Multivariate Statistical Analysis*, MSA 2016). Jej organizatorami przy udziale Polskiego Towarzystwa Statystycznego (PTS) byli: Instytut Statystyki i Demografii Uniwersytetu Łódzkiego (UŁ), Katedra Metod Statystycznych UŁ oraz Komitet Statystyki i Ekonometrii PAN. Honorowy patronat nad konferencją objęła prezydent miasta Łodzi — Hanna Zdanowska. Konferencja została wsparta finansowo przez NBP oraz PAN ze środków przeznaczonych na Działalność Upowszechniającą Naukę. Partnerem Konferencji była firma StatSoft Polska Sp. z o.o. Funkcję przewodniczącego Komitetu Organizacyjnego pełnił prof. dr hab. Czesław Domański, a sekretarzami naukowymi byli mgr Piotr Szczepocki oraz dr hab. inż. Jacek Białek.

Głównym celem XXXV Konferencji WAS 2016 była prezentacja najnowszych osiągnięć z zakresu wielowymiarowej analizy statystycznej oraz ich zastosowań w naukach ekonomicznych.

W Konferencji wzięło udział 78 osób z ośrodków akademickich w Polsce (Katowice, Krakowa, Gdańska, Poznania, Rzeszowa, Słupska, Szczecina, Warszawy i Wrocławia), przedstawiciele GUS i Urzędu Statystycznego w Łodzi, a także goście z Włoch oraz studenci UŁ. Podczas siedemnastu sesji (plenarnych i równoległych) uczestnicy wysłuchali 63 referatów oraz 2 wykładów zaproszonych.

Obrady otworzył prof. Czesław Domański, a następnie wystąpili prezes GUS dr Dominik Rozkrut oraz prodziekan Wydziału Ekonomiczno-Socjologicznego UŁ dr hab. Michał Przybyliński, prof. UŁ. Pierwszy wykład zaproszony wygłosił prof. Józef Pociecha (Uniwersytet Ekonomiczny w Krakowie, Komitet Statystyki i Ekonometrii PAN) na temat *Statystyka Polski — pierwszy rocznik statystyczny ziem polskich*.

Sesja I (plenarna) — historyczna, poświęcona była wybitnym przedstawicielom polskiej myśli statystycznej. Przewodniczył jej Włodzimierz Okrasa (GUS, Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie). Wystąpili:

- Jan Berger (GUS) wygłosił *Wspomnienie o Witoldzie Załęskim*;
- Mirosław Krzyśko (Uniwersytet im. Adama Mickiewicza Wydział Matematyki i Informatyki) przedstawił referat *Wspomnienie o Mikołaju Olekiewiczu*;
- Jan Kordos (Wyższa Szkoła Menedżerska w Warszawie) przygotował referat *Janusz Buga — wspomnienie*;
- referat *Jan Śniadecki — pierwszy polski probabilista* wygłosił Czesław Domański.

Sesji II (plenarnej) przewodniczyła Danuta Strahl (Uniwersytet Ekonomiczny we Wrocławiu, Komitet Statystyki i Ekonometrii PAN).

Włodzimierz Okrasa wygłosił wykład zaproszony na temat *Jakość życia a jakość otoczenia: wieloźródłowa — wielopoziomowa analiza zależności z użyciem danych z badań budżetu czasu*.

W pierwszym dniu Konferencji wysłuchano też innych referatów:

- Dominik Rozkrut (GUS, Uniwersytet Szczeciński) — *Big Data w statystyce publicznej*;
- Mirosław Krzyśko, Tomasz Górecki, Waldemar Wołyński (Uniwersytet im. Adama Mickiewicza w Poznaniu) — *Funkcjonalna analiza kanoniczna — nowe podejście*;
- Andrzej Sokołowski (Uniwersytet Ekonomiczny w Krakowie, Komitet Statystyki i Ekonometrii Polskiej Akademii Nauk), Marek Sobolewski (Politechnika Rzeszowska) — *O niewłaściwym wykorzystaniu współczynnika zmienności*;
- Grażyna Trzpiot (Uniwersytet Ekonomiczny w Katowicach, Komitet Statystyki i Ekonometrii PAN), Justyna Majewska (Uniwersytet Ekonomiczny w Katowicach) — *Modelowanie i prognozowanie oczekiwanej długości życia z zastosowaniem regresji Poissona logarytmiczno-dwuliniowej*.

Obradom sesji IIIA (równoległej) przewodniczyła Małgorzata Markowska (Uniwersytet Ekonomiczny we Wrocławiu). A oto autorzy i tematyka referatów:

- Artur Zaborski (Uniwersytet Ekonomiczny we Wrocławiu) — *Wpływ wyboru triad na wyniki skalowania preferencji*;
- Daniel Kosiorowski, Małgorzata Snarska (Uniwersytet Ekonomiczny w Krakowie), Dominik Mielczarek, Jerzy Rydlewski (Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie) — *Uogólniona metoda wygładzania wykładniczego w zagadnieniu predykcji hierarchicznego funkcjonalnego szeregu czasowego*;
- Daniel Kosiorowski (Uniwersytet Ekonomiczny w Krakowie), Ewa Szlachetowska, Zygmunt Zawadzki (Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie) — *Wybrane algorytmy analizy skupisk we wstępnej analizie oraz specyfikacji funkcjonalnego szeregu czasowego*;
- Piotr Szczepocki (Uniwersytet Łódzki) — *Zastosowanie filtru Kalmana do modeli stochastycznej zmienności typu Ornsteina-Uhlenbecka*.

Sesji IIIB (równoległej) przewodniczył Andrzej Dudek (Uniwersytet Ekonomiczny we Wrocławiu); wysłuchano 4 referatów:

- Romana Głowicka-Wołoszyn (Uniwersytet Przyrodniczy w Poznaniu) — *Zastosowanie przestrzennego taksonomicznego miernika rozwoju do oceny samodzielności finansowej samorządów gminnych*;
- Jacek Białek (Uniwersytet Łódzki), Radosław Pietrzyk (Uniwersytet Ekonomiczny we Wrocławiu) — *Zastosowanie indeksu Divisia z powiązаныmi czynnikami do analizy fluktuacji indeksu WIG*;
- Dorota Rozmus (Uniwersytet Ekonomiczny w Katowicach) — *Porównanie wyników badania stabilności grupowania z zastosowaniem programu R*;
- Tomasz Bartłomowicz (Uniwersytet Ekonomiczny we Wrocławiu) — *Estymacja wektora wag w metodzie Analytic Hierarchy Process z wykorzystaniem pakietu AHP programu R*.

Obradom sesji IVA (równoległej) przewodniczył Grzegorz Kończak (Uniwersytet Ekonomiczny w Katowicach). Tematyka referatów była następująca:

- Iwona Markowicz (Uniwersytet Szczeciński) — *Zróżnicowanie modelu trwania firm w powiatach województwa zachodniopomorskiego*;
- Beata Bieszk-Stolorz, Iwona Markowicz (Uniwersytet Szczeciński) — *Modele analizy trwania w ocenie wpływu bessy na sytuację subsektorów giełdowych*;
- Małgorzata Markowska, Marek Sobolewski (Politechnika Rzeszowska) — *Metody grupowania hierarchicznego z warunkiem spójności w analizach regionalnych*;
- Artur Mikulec (Uniwersytet Łódzki) — *Przekrojowe tablice trwania przedsiębiorstw w województwie łódzkim*.

Przewodniczącą sesji IVB (równoległej) była Grażyna Trzpiot. Wygłoszono 4 referaty:

- Grażyna Dehnel (Uniwersytet Ekonomiczny w Poznaniu), Tomasz Klimanek (Uniwersytet Ekonomiczny w Poznaniu, Urząd Statystyczny w Poznaniu) — *Osoby niepełnosprawne w Polsce w latach 2002 i 2011 w świetle estymacji pośredniej*;
- Michał Bernardelli, Barbara Kowalczyk (SGH) — *Optymalna alokacja próby w badaniu cechy drażliwej*;
- Wioletta Grzenda (SGH) — *Modelowanie czasu trwania pierwszej pracy z wykorzystaniem bayesowskich modeli przyspieszonej porażki AFT*;
- Elżbieta Gołata (Uniwersytet Ekonomiczny w Poznaniu) — *Aktywność ekonomiczna osób niepełnosprawnych w ujęciu regionalnych rynków pracy*.

Sesja V rozpoczęła drugi dzień spotkania naukowego. Przewodniczył jej Bronisław Ceranka (Uniwersytet Przyrodniczy w Poznaniu), a autorami referatów byli:

- Marek Walesiak (Komitet Statystyki i Ekonometrii PAN), Andrzej Dudek (Uniwersytet Ekonomiczny we Wrocławiu) — *Wybór optymalnej procedury skalowania wielowymiarowego dla danych metrycznych z wykorzystaniem programu R*;

- Józef Dziechciarz (Komitet Statystyki i Ekonometrii PAN), Marta Dziechciarz-Duda (Uniwersytet Ekonomiczny we Wrocławiu) — *Analiza danych niemetrycznych w badaniu wyposażenia gospodarstw domowych w dobra trwałe. Wybrane aspekty*;
- Tadeusz Bednarski, Magdalena Skolimowska-Kulig (Uniwersytet Wrocławski) — *Zgodna z dokładnością do skali estymacja parametrów w modelach regresji ze zmienną „frailty”*;
- Grzegorz Kończak — *O testowaniu istotności wielowymiarowego współczynnika korelacji rang.*

Obradom sesji VIA (równoległej) przewodniczyła Grażyna Dehnel. Wysłuchano 4 referatów:

- Mirosław Krzyśko, Wojciech Łukaszonek, Waldemar Wołyński (Uniwersytet im. Adama Mickiewicza w Poznaniu, Państwowa Wyższa Szkoła Zawodowa w Kaliszu) — *Współrzędne dyskryminacyjne dla danych podwójnie wielowymiarowych*;
- Elżbieta Roszko, Jacek Białek (Uniwersytet Łódzki) — *Zastosowanie metod oceny grupowej do oceny innowacyjności krajów Unii Europejskiej*;
- Joanna Landmesser (SGGW) — *Pomiar różnic między rozkładami dochodów w Polsce w latach 2000—2012*;
- Aneta Ptak-Chmielewska (Bank Gospodarstwa Krajowego) — *Modele oceny ryzyka kredytowego przedsiębiorstw — teoria i praktyka.*

Sesji VIB (równoległej) przewodniczył Tadeusz Bednarski (Uniwersytet Wrocławski). Wygłoszono wówczas 4 referaty:

- Adrianna Mastalerz-Kodzis, Ewa Pośpiech (Uniwersytet Ekonomiczny w Katowicach) — *Zastosowanie funkcji Höldera do badania intensywności rozprzestrzeniania się zjawisk przestrzennych*;
- Ewa Pośpiech, Adrianna Mastalerz-Kodzis (Uniwersytet Ekonomiczny w Katowicach) — *Zastosowanie regresji przestrzennej do modelowania charakterystyk zatrudnienia*;
- Agnieszka Kozera, Romana Głowicka-Wołoszyn (Uniwersytet Przyrodniczy w Poznaniu) — *Autokorelacja przestrzenna potencjału dochodowego gmin w wybranych obszarach metropolitalnych*;
- Monika Miśkiewicz-Nawrocka, Katarzyna Zeug-Żebro (Uniwersytet Ekonomiczny w Katowicach) — *Sytuacja demograficzna a poziom bezrobocia w Polsce w latach 2005—2014.*

Prowadzącym sesję VIIA (równoległą, w języku polskim) był Wojciech Zieliński (SGGW). Przedstawiono 4 referaty:

- Katarzyna Słupor (Politechnika Śląska) — *Testy statystyczne dla porównania klasyfikatorów: przegląd, pułapki i zalecenia*;
- Grzegorz Sitek, Janusz Wywiiał (Komitet Statystyki i Ekonometrii PAN, Uniwersytet Ekonomiczny w Katowicach) — *O dokładności estymacji parametrów biprostych regresji*;
- Piotr Sulewski (Akademia Pomorska) — *Nowy test niezależności dla tablic dwudzielnych 2x2*;

— Michał Trzęsiok (Uniwersytet Ekonomiczny w Katowicach) — *Wybrane metody pomiaru jakości modeli statystycznych.*

Obrady sesji VIIB (równoległej) poprowadził Jerzy Korzeniewski (Uniwersytet Łódzki). Wygłoszono następujące referaty:

— Dominik Krężolek (Uniwersytet Ekonomiczny w Katowicach) — *Wybrane modele klasy GARCH na rynku metali — testowanie wsteczne Value-At-Risk;*

— Dominik Krężolek, Grażyna Trzpiot — *Efektywność miary GlueVaR w ocenie ryzyka na rynku metali — zastosowanie wskaźnika Omega;*

— Alicja Ganczarek-Gamrot (Uniwersytet Ekonomiczny w Katowicach) — *Wielowymiarowe modelowanie zmienności i analiza ryzyka zmian cen energii elektrycznej w Europie;*

— Anna Chmielewska (SGH) — *Strukturalne zmiany w kształtowaniu się zmienności i współzależności obserwowane w kontekście kryzysu finansowego 2007—2009 i późniejszych zmian regulacyjnych.*

Sesji VIIIA (równoległej, w języku angielskim) przewodniczyła Alina Jędrzejczak (Uniwersytet Łódzki). Wysłuchano wtedy:

— Francescii Greselin (Universita degli Studi di Milano Bicocca) — *Parsimonious and robust clustering;*

— Dariusza Parysa (Uniwersytet Łódzki) — *Some modifications of weighted multiple procedures;*

— Małgorzaty Krzciuk (Uniwersytet Ekonomiczny w Katowicach) — *On the simulation study of jackknife and bootstrap MSE estimators of some domain mean predictor for Fay-Herriot model.*

Sesji VIIIB (równoległej) przewodniczyła Iwona Markowicz (Uniwersytet Szczeciński). Wygłoszono 2 referaty:

— Joanna Trzęsiok (Uniwersytet Ekonomiczny w Katowicach) — *Obserwacje odstające a problem odporności nieparametrycznych metod regresji;*

— Andrzej Wołoszyn, Romana Głowicka-Wołoszyn, Agnieszka Kozera (Uniwersytet Przyrodniczy w Poznaniu) — *Zastosowanie uogólnionej miary odległości GDM do oceny subiektywnego poziomu bezpieczeństwa finansowego gospodarstw domowych rolników.*

Sesja IXA (równoległa) rozpoczęła trzeci dzień konferencji, przewodniczył jej Janusz Wywiół. Autorzy zgłosili następujące tematy:

— Anna Szymańska (Uniwersytet Łódzki) — *Zastosowanie zmodyfikowanego modelu Bülmanna-Strauba do estymacji stawek składki netto w systemach bonus-malus ubezpieczeń komunikacyjnych OC;*

— Aleksandra Baszczyńska (Uniwersytet Łódzki) — *Jedna wartość parametru wygładzania vs. przedział wartości parametru wygładzania w estymacji jądrowej funkcji gęstości;*

— Kamil Sapała, Marcin Piołun-Noyszewski, Marcin Weiss (Free Construction) — *Wykrywanie zagrożeń występujących w systemach teleinformatycznych przy pomocy wybranych narzędzi i metod Data Mining;*

— Mariola Chrzanowska, Iwona Pomianek (SGGW) — *Wielowymiarowa analiza porównawcza warunków życia w powiatach województwa mazowieckiego w latach 2005—2015.*

Obradom sesji IXB (równoległej) przewodniczył Marek Walesiak. Wygłoszono 4 referaty:

- Monika Zielińska-Sitkiewicz (SGGW) — *Wpływ wyboru formuły normalizacyjnej na klasyfikację spółek przemysłu materiałów budowlanych;*
- Angelina Rajda-Tasior (Uniwersytet Ekonomiczny w Katowicach) — *Ocena zdolności systemu pomiarowego dla pomiarów niepowtarzalnych w procesach produkcyjnych;*
- Michał Szymczak (Uniwersytet Łódzki) — *Wybrane metody opisu niezawodności urzędzeń;*
- Małgorzata Misztal (Uniwersytet Łódzki) — *O zastosowaniu analizy redundancji do badania przestępczości przeciwko mieniu w Polsce.*

Sesji X (plenarnej, w języku angielskim) przewodniczył Andrzej Sokołowski, a wysłuchano 3 referatów:

- Jerzy Korzeniewski (Uniwersytet Łódzki) — *Efficient stock portfolio construction by means of clustering;*
- Tomasz Żądło, Tomasz Stachurski (Uniwersytet Ekonomiczny w Katowicach) — *On calibration estimation supported by auxiliary variables from past periods;*
- Jan Kubacki (Urząd Statystyczny w Łodzi), Alina Jędrzejczak (Uniwersytet Łódzki) — *Estimation of small area characteristics using multivariate Rao-Yu model.*

Ostatniej sesji XI (plenarnej) przewodniczył Józef Dziechciarz (Uniwersytet Ekonomiczny we Wrocławiu, Komitet Statystyki i Ekonometrii PAN). Podczas niej referaty wygłoszili:

- Bronisław Ceranka, Małgorzata Graczyk (Uniwersytet Przyrodniczy w Poznaniu) — *Problem efektywności w układach wagowych;*
- Małgorzata Graczyk, Bronisław Ceranka (Uniwersytet Przyrodniczy w Poznaniu) — *D-optymalne układy wagowe o ujemnie skorelowanych błędach: nowe klasy;*
- Wojciech Zieliński, Dominik Sieradzki (SGGW) — *Estymacja frakcji w dwuwarstwowej populacji skończonej.*

Podsumowania obrad dokonał prof. dr hab. Czesław Domański, który podziękował wszystkim uczestnikom za udział w Konferencji i zaprosił na kolejną WAS 2017, która odbędzie się od 6 do 8 listopada 2017 r. w Łodzi.

Wydawnictwa GUS — lipiec 2017 r.

Z lipcowej oferty wydawniczej GUS warto zwrócić uwagę na publikację cykliczną **„Zielone Płuca Polski w 2015 r.”** oraz wydanie jednorazowe **„Kompendium wiedzy o rachunkach regionalnych”**.



Pierwsza z nich, ukazująca się co 3 lata, to już ósme wydanie publikacji opisującej statystycznie „Zielone Płuca” Polski. Idea „Zielonych Płuc”, sformułowana w 1983 r. przez Krzysztofa Wolframa, zakłada integrację ochrony środowiska z rozwojem gospodarczym i postępem cywilizacyjnym w północno-wschodnim regionie kraju. Zaliczany on jest do obszarów o nieskażonej przyrodzie oraz niezwykłych walorach krajoznawczych i kulturowych w Europie. Opracowanie przygotowano w Urzędzie Statystycznym w Białymstoku.

Publikacja jest opracowaniem tabelarycznym wzbogaconym grafiką. Przedstawiono w niej m.in.: charakterystykę warunków naturalnych, zagrożeń i ochrony środowiska, sytuacji demograficznej i społecznej, infrastruktury komunalnej i mieszkaniowej oraz kultury i turystyki według województw oraz na tle kraju w 2015 r. Niektóre informacje ujęto także według powiatów i gmin. Ponadto w opracowaniu zamieszczono tablicę przeglądową zawierającą ważniejsze dane o tym ekoregionie z lat: 1993, 1996, 1999, 2003, 2006, 2009, 2012 i 2015.

Publikacja uwzględnia również uwagi metodyczne, w których przywołano podstawowe definicje pojęć niezbędne do właściwej interpretacji danych. W porównaniu z poprzednią edycją opracowanie uzupełniono informacjami o ścieżkach rowerowych, zbiornikach bezodpływowych oraz indywidualnych przydomowych oczyszczalniach ścieków.

Publikacja ukazała się w wersji polsko-angielskiej i jest dostępna również na stronie internetowej GUS. Wszystkie tablice udostępniono w formacie MS Excel, co ułatwia prowadzenie własnych analiz.



„Kompedium wiedzy o rachunkach regionalnych” zawiera zbiór najważniejszych informacji na temat rachunków regionalnych przeznaczonych nie tylko dla statystyków. Ma ono na celu ułatwienie korzystania z informacji statystycznych o tym bardzo zawężonym dziale statystyki.

W publikacji przedstawiono informacje merytoryczne: pojęcie regionu, jednostki i ich grupowanie, transakcje, wycena, zakres zestawianych rachunków, niektóre kategorie makroekonomiczne, zasady ich zestawiania, metody regionalizacji oraz źródła danych. Ponadto Czytelnicy znajdą tam inne informacje związane m.in. z systemem rachunków oraz obligatoryjnym zakresem rachunków regionalnych w Unii Europejskiej. W kompendium opisano również podstawy prawne, dokumenty, podręczniki i klasyfikacje, bazy danych i publikacje oraz znaczenie i wykorzystanie danych z rachunków regionalnych.

Publikacja ukazała się po polsku i jest dostępna również na stronie internetowej Urzędu.

W lipcu br. ukazały się także: „Mały Rocznik Statystyczny Polski 2017”, „Aktywność ekonomiczna ludności Polski — I kwartał 2017 r.”, „Biuletyn Statystyczny Nr 6/2017”, „Budownictwo — wyniki działalności w 2016 r.”, „Ceny robót budowlano-montażowych i obiektów budowlanych — maj 2017 r.”, „Informacja o sytuacji społeczno-gospodarczej kraju w pierwszym półroczu 2017 r.”, „Polska w Unii Europejskiej 2017” (folder), „Powierzchnia i ludność w przekroju terytorialnym w 2017 r.”, „Produkcja ważniejszych wyrobów przemysłowych w czerwcu 2017 r.”, „Produkcja wyrobów przemysłowych w 2016 r.”, „Statistics in Transition — new series Vol. 18 No. 2”, „Trwanie życia w 2016 r.”, „Uczestnictwo w sporcie i rekreacji ruchowej w 2016 r.”, „Wskaźniki zielonej gospodarki w Polsce 2017”, „Zwierzęta gospodarskie w 2016 r.” oraz „Wiadomości Statystyczne nr 7/2017 (674)”.

Oprac. Justyna Gustyn

Do Autorów

Szanowni Państwo!

- W „Wiadomościach Statystycznych” publikowane są artykuły o charakterze naukowym poświęcone teorii i praktyce statystycznej, prezentujące wyniki oryginalnych badań teoretycznych lub analitycznych wykorzystujących metody statystyki matematycznej, opisowej lub ekonometrii. W miesięczniku zamieszczane są również artykuły przeglądowe, popularnonaukowe, recenzje publikacji naukowych oraz inne opracowania informacyjne. Prezentowany w artykule naukowym problem badawczy powinien być jednoznacznie zdefiniowany oraz istotny dla oceny zjawisk społecznych lub gospodarczych. Wyniki studiów przeprowadzanych w artykułach winny oddziaływać na rozwój myśli statystycznej oraz edukacji, wnosząc oryginalny wkład do tej dziedziny.

Zasopismo publikuje także artykuły i opracowania prezentujące informacje o teorii i praktyce statystycznej, jak również o problemach edukacji statystycznej. Dotyczą one: programów badań statystycznych statystyki publicznej, systemu zbierania i udostępniania informacji statystycznych, zastosowań informatyki w statystyce, informacji o konferencjach naukowych, działalności organów doradczych prezesa GUS oraz edukacji statystycznej.

- Artykuły kierowane do opublikowania w „Wiadomościach Statystycznych” powinny zawierać precyzyjny opis badanych zjawisk i stosowanych metod oraz autorskie wnioski i sugestie dotyczące rozwoju badań i analiz statystycznych. Autorzy winni wyraźnie określić cel artykułu oraz jasno przedstawić uzyskane wyniki przeprowadzonej analizy. W przypadku prezentacji badań prowadzonych przez Autorów należy opisać zastosowaną w nich metodę. Przy prezentacji nowatorskich metod analizy pożądane jest podanie przykładu pokazującego ich zastosowanie w praktyce statystycznej.
- Artykuły zamieszczane w „Wiadomościach Statystycznych” powinny wyrażać opinie własne Autorów. Autorzy ponoszą odpowiedzialność za treści prezentowane w artykułach. W razie zgłaszania przez czytelników zastrzeżeń odnoszących się do tych treści, Autorzy są zobligowani do udzielenia odpowiedzi na łamach miesięcznika.
- Po wstępnej ocenie przez Redakcję „Wiadomości Statystycznych” tematyki artykułu pod względem zgodności z profilem czasopisma, artykuły mające charakter naukowy przekazywane są do recenzji osobom specjalizującym się w poszczególnych dziedzinach, które w ocenie kierują się kryterium oryginalności i jakości opracowania, w tym treści i formy, a także potencjalnego zainteresowania czytelników.
- Recenzowanie artykułów odbywa się zgodnie z jednym z dwóch systemów rekomendowanych przez MNiSW — stosowana jest zasada single-blind. Szczegółowe informacje dotyczące procedury recenzowania, kryteria oceny oraz wzór karty recenzji artykułu znajdują się na stronie internetowej „Wiadomości Statystycznych” w zakładce *Recenzowanie artykułów*.

- Autorzy artykułów, które otrzymały pozytywne recenzje, wprowadzają zasugerowane przez recenzentów poprawki i dostarczają Redakcji zaktualizowaną wersję opracowania. Autorzy poświadczają w przysłanym piśmie uwzględnienie wszystkich poprawek. Jeśli pojawi się różnica zdań co do zasadności proponowanych zmian, należy wyjaśnić, które poprawki zostały uwzględnione, a w przypadku ich nieuwzględnienia przedstawić motywy swojego stanowiska.
- Kontroli poprawności stosowanych przez Autorów metod statystycznych dokonują redaktorzy statystyczni.
- Decyzję o publikacji artykułu podejmuje Kolegium Redakcyjne „Wiadomości Statystycznych”. Podstawą tej decyzji jest wynik dyskusji dotyczącej zgłoszonego artykułu, w której uwzględniane są opinie przedstawione w recenzjach wraz z rekomendacją ich opublikowania.
- Redakcja „Wiadomości Statystycznych” przestrzega zasady nietolerowania przejawów nierzetelności naukowej autorów artykułów polegającej na:
 - nieujawnianiu współautorów, mimo że wnieśli oni istotny wkład w powstanie artykułu, określanemu w języku angielskim terminem „ghostwriting”;
 - podawaniu jako współautorów osób o znikomym udziale lub niebiorących udziału w opracowaniu artykułu, określanemu w języku angielskim terminem „guest authorship”.

Stwierdzone przypadki nierzetelności naukowej w tym zakresie mogą być ujawniane. W celu przeciwdziałania zjawiskom „ghostwriting” i „guest authorship” należy dołączyć do przesłanego artykułu oświadczenie, którego wzór zamieszczono na stronie internetowej czasopisma (link do załącznika znajduje się w zakładce „Do Autorów”).

Główną odpowiedzialność za rzetelność przekazanych informacji, łącznie z informacją na temat wkładu poszczególnych współautorów w powstanie artykułu, ponosi zgłaszający artykuł.

Redakcja „Wiadomości Statystycznych” zastrzega sobie prawo dokonywania w artykułach zmian tytułów, skrótów i przeredagowania tekstu i tablic bez naruszenia zasadniczej myśli Autora.

Redakcja „Wiadomości Statystycznych” oświadcza, że nie wypłaca Autorom honorariów za opracowanie artykułów zamieszczanych na łamach naszego czasopisma oraz nie pobiera opłat za ich publikację.

Uprzejmie informujemy, że od 2007 r. „Wiadomości Statystyczne” znajdują się na liście polskich punktowanych czasopism naukowych Ministerstwa Nauki i Szkolnictwa Wyższego.

W komunikacie MNiSW z dnia 18 grudnia 2015 r. w sprawie wykazu czasopism naukowych wraz z liczbą punktów przyznawanych za publikację w tych czasopismach, miesięcznikowi „Wiadomości Statystyczne” przyznano **12 punktów**.

Miesięcznik posiada konto w Polskiej Bibliografii Naukowej. „Wiadomości Statystyczne” są indeksowane w następujących bazach: Index Copernicus, CEJSH (*Central European Journal of Sciences and Humanities*) oraz w BazEkon.

- Redakcja „Wiadomości Statystycznych” oświadcza:
- wersja elektroniczna czasopisma jest jego wersją pierwotną;
 - dostęp do pełnej zawartości czasopisma jest otwarty;
 - datą publikacji numeru „Wiadomości Statystycznych” w otwartym dostępie w Internecie jest dzienna data jego zamieszczenia na stronie internetowej czasopisma;
 - wersja artykułów zamieszczonych w otwartym dostępie jest ich wersją ostateczną;
 - **materiały zamieszczone w „Wiadomościach Statystycznych” są chronione prawem autorskim. Przedruk tekstu może nastąpić wyłącznie za zgodą Redakcji. Treści cytowane z „Wiadomości Statystycznych” powinny być opatrzone dokładną informacją o źródle ich pochodzenia.**

Informacje dotyczące wymaganej formy oraz kompletności przesyłanych artykułów do „Wiadomości Statystycznych”

- Artykuły należy dostarczać pocztą elektroniczną na adres:

a.swiderska@stat.gov.pl lub e.grabowska@stat.gov.pl
Redakcja „Wiadomości Statystyczne”
Główny Urząd Statystyczny
al. Niepodległości 208, 00-925 Warszawa

- Konieczne jest dołączenie do artykułu skróconej informacji (streszczenia) o jego treści (ok. 10 wierszy) w językach polskim i angielskim. **Streszczenie powinno być utrzymane w formie bezosobowej i zawierać: zwięźle sprecyzowany cel badania, przybliżony jego zakres i przyjętą metodologię badania oraz ważniejsze wnioski.**
- Prosimy o podanie słów kluczowych, w językach polskim i angielskim, przybliżających zagadnienia w artykule.
- Prosimy również o podanie kodów klasyfikacji JEL (*Journal of Economic Literature*).
- **Redakcja rozpoczyna postępowanie kwalifikujące artykuł do opublikowania po spełnieniu warunku przesłania przez Autora oświadczenia.**
- Pytania dotyczące przesłanego artykułu, co do jego aktualnego statusu itp., należy kierować do redakcji na adres: a.swiderska@stat.gov.pl lub e.grabowska@stat.gov.pl lub tel. 22 608-32-25.

Wymogi czasopisma dotyczące przygotowania artykułu

Artykuł powinien mieć optymalną objętość (łącznie z wykresami, tablicami i literaturą) 10—20 stron przygotowanych zgodnie z poniższymi wytycznymi:

1. Edytor tekstu — Microsoft Word, format *.doc lub *.docx.
2. Czcionka:
 - autor — Arial, wersalik, wyrównanie do lewej, 12 pkt.,
 - tytuł opracowania — Arial, wyśrodkowany, 16 pkt.,
 - tytuły rozdziałów i podrozdziałów — Times New Roman, wyśrodkowany, kursywa, 14 pkt.,
 - tekst główny — Times New Roman, normalny, wyjustowany, 12 pkt.,
 - przypisy — Times New Roman, 10 pkt.
3. Marginesy przy formacie strony A4 — 2,5 cm z każdej strony.
4. Odstęp między wierszami półtorej linii oraz interlinia przed tytułami rozdziałów.
5. Pierwszy wiersz akapitu wcięty o 0,4 cm, enter na końcu akapitu.
6. Wyszczególnianie rozmaitych kategorii należy zacząć od kropek, a numerowanie od cyfr arabskich.

7. Strony powinny być ponumerowane automatycznie.
8. Wykresy powinny być zamieszczone w artykule oraz konieczne przesłane w oddzielnym oryginalnym pliku, np. Excelu lub Corelu.
9. Tablice należy zamieszczać w tekście, zgodnie z treścią artykułu. W tablicach nie należy stosować rastrów, cieniowania, pogrubiania czy też podwójnych linii itp.
10. Pod wykresami i tablicami należy podać informacje dotyczące źródła opracowania.
11. Oznaczenia literowe należy wyróżniać następująco: macierze — wersalik, proste, pogrubione (np. **P**, **N_{ij}**); wektory — małe litery, kursywa, pogrubione (np. **w**, **x_i**); pozostałe zmienne — małe lub duże litery, kursywa, bez pogrubienia (np. *w*, *x_i*, *Z*).
12. Stosowane są skróty: tablica — tabl., wykres — wykr.
13. Przypisy do tekstu należy umieszczać na dole strony.
14. Wykaz literatury załącznikowej i przytoczenia konkretnych prac w treści artykułu należy przygotować według stylu APA (*American Psychological Association*).

Zasady przywoływania pracy w tekście:

- a.** Jeden autor: bez względu na to ile razy przywoływana jest praca, zawsze należy podać nazwisko autora i datę publikacji pracy, w przypadku więcej niż jednej pracy danego autora opublikowanej w tym samym roku należy dodać kolejne litery alfabetu przy dacie (np., 2001a), zasada ta obowiązuje także w przypadku większej liczby autorów danej pracy.

Przykład zapisu:

Jak stwierdza Iksiński (2001)...

Badania wskazują, iż... (Iksiński, 2001).

- b.** Dwóch autorów: bez względu na to ile razy przywoływana jest praca, zawsze należy podać nazwiska obu autorów i datę publikacji pracy, w przypadku więcej niż jednej pracy tych autorów opublikowanej w tym samym roku należy dodać kolejne litery alfabetu przy dacie. Nazwiska autorów zawsze należy łączyć spójnikiem „i”, nawet w przypadku przywoływania publikacji obcojęzycznej.

Przykład zapisu:

Jak sugerują Iksiński i Nowak (1999)...

Badania wskazują, iż... (Iksiński i Nowak, 1999).

- c.** Od trzech do pięciu autorów: przywołanie po raz pierwszy — należy wymienić nazwiska wszystkich autorów, rozdzielając je przecinkami i stawiając spójnik „i” pomiędzy dwoma ostatnimi nazwiskami. Przy kolejnych wskazaniach tej samej pracy można zastosować określenie „i współpracownicy” (w przypadku umieszczenia przywołania nazwisk w strukturze zdania) lub „i in.” (w przypadku, gdy nazwiska autorów nie stanowią części struktury zdania).

Przykład zapisu:

Przywołanie po raz pierwszy:

Jak sugerują Nowak, Iksiński i Jankiewicz (2003)...

Badania (Nowak, Iksiński i Jankiewicz, 2003) wskazują, iż...

Kolejne przywołania:

Badania Nowaka i współpracowników (2003)...

Badania te wskazują, iż... (Nowak i in., 2003).

- d.** Sześciu i więcej autorów: wymienić należy tylko nazwisko pierwszego autora, zarówno gdy praca przywoływana jest po raz pierwszy, jak i w późniejszych przywołaniach, natomiast pozostałych autorów należy zastąpić skrótem „i in.” (gdy nazwiska nie stanowią części struktury zdania). W literaturze cytowanej należy umieścić nazwiska wszystkich autorów pracy.

Przykład zapisu:

Nowakowski i współpracownicy twierdzą, iż ... (1997).

Pierwsze badania na ten temat (Nowakowski i in., 1997) sugerują...

- e. Przywoływanie jednocześnie kilku prac: należy wymienić je alfabetycznie, według nazwiska pierwszego autora. Przywołania kolejnych prac muszą być oddzielone średnikiem i umieszczone w nawiasie. Lata wydania prac tego samego autora/autorów muszą być oddzielone przecinkiem.

Przykład zapisu:

(Iksiński, 2001; Nowak i Iksiński, 1999)

(Iksiński, 1997, 1999, 2004a, 2004b; Nowak i Iksiński, 1999).

- f. Przywoływanie pracy za innym autorem: stosujemy w tekście, natomiast w literaturze cytowanej umieszczamy jedynie pracę czytaną.

Przykład zapisu:

Jak wykazał Nowakowski (1990; za: Zieniecka, 2007)...

Badania sugerują, iż ... (Nowakowski, 1990; za: Zieniecka, 2007).

15. Wykaz literatury należy zamieszczać na końcu opracowania.

Prace zapisujemy przy zachowaniu kolejności alfabetycznej cytowanych dzieł, przy czym decyduje pierwsza litera nazwiska autora.

Każdą nową pracę zaczynamy bez wcięcia, wyrównanie do lewego marginesu, a kolejne wiersze danego adresu bibliograficznego powinny być zapisane z wcięciem 0,4 cm.

Zasady zapisu literatury załącznikowej:

Poniżej znajdują się schematy zapisów bibliograficznych podstawowych źródeł (artykułów i książek). Sposoby zapisu innych, rzadziej powoływanych źródeł są szczegółowo opisane w szóstym wydaniu „Publication Manual of the American Psychological Association”.

- a. artykuł w czasopiśmie, w którym każdy kolejny numer/zeszyt (*issue*) w ramach jednego rocznika ma osobną numerację stron (w każdym zeszycie pierwsza strona opatrzona jest numerem 1):

Nazwisko, X., Nazwisko2, X. Y., Nazwisko3, Z. (rok). Tytuł artykułu. *Tytuł Czasopisma, nr rocznika* (nr zeszytu), strona początku—strona końca.

- b. artykuł w czasopiśmie, w którym kolejne numery/zeszyty (*issues*) w ramach jednego rocznika nie mają osobnej numeracji stron (pierwsza strona w kolejnym zeszycie opatrzona jest numerem kolejnym, po ostatniej stronie w zeszycie poprzednim):

Nazwisko, X., Nazwisko2, X. Y., Nazwisko3, Z. (rok). Tytuł artykułu. *Tytuł Czasopisma, nr rocznika*, strona początku—strona końca.

- c. jeśli artykuł ma numer DOI (*Digital Object Identifier*), należy podać go na końcu zapisu bibliograficznego: Nazwisko, X., Nazwisko2, X. Y. (rok). Tytuł artykułu. *Tytuł Czasopisma, nr rocznika*, strona początku—strona końca. DOI: xxxxx.

- d. książka:

Nazwisko, X., Nazwisko, X. Y. (rok). *Tytuł książki*. Miejsce wydania: Wydawnictwo.

- e. książka napisana pod redakcją:

Nazwisko, X. (red.). (rok). *Tytuł książki*. Miejsce wydania: Wydawnictwo.

- f. rozdział w pracy zbiorowej:

Nazwisko, X. (rok). Tytuł rozdziału. W: Y. Nazwisko, B. Nazwisko (red.), *Tytuł książki* (s. strona początku—strona końca). Miejsce wydania: Wydawnictwo.

W stylu APA proponuje się zapis bibliograficzny bez użycia dwukropka po przyimku *W* (*In*), pisany wielką literą. W polskim zapisie jednak przyjmujemy zasadę pisania dwukropka po *W*.

- g.** jeśli dany tekst znajduje się na stronie internetowej i nie jest artykułem w czasopiśmie, książką ani rozdziałem w książce, należy podać autora, datę publikacji (jeśli jest znana), tytuł, a następnie zamieścić informacje o stronie, skąd został pobrany tekst:

Nazwisko, X. (rok). *Tytuł tekstu*. Pobrane z: adres strony internetowej.

16. W wykazie literatury należy zamieścić wyłącznie pozycje przytoczone w artykule.
17. Opracowanie przygotowane w sposób niezgodny z powyższymi wskazówkami będzie odesłane z prośbą o dostosowanie jego formy do wymagań redakcji.

Charakterystyka zakresu tematycznego poszczególnych działów „Wiadomości Statystycznych”

STUDIA METODOLOGICZNE

W dziale tym zamieszczane są artykuły naukowe zawierające prezentacje teoretycznych rozwiązań metodologicznych, ze wskazaniem ich praktycznej użyteczności, w tym prace o charakterze przeglądowym i porównawczym oraz dotyczące zagadnień etyki statystycznej. Poruszane tu zagadnienia mogą obejmować różnorodne dziedziny statystyki, ekonomii matematycznej i ekonometrii, a prezentowane rezultaty badawcze stwarzają możliwość efektywnego zastosowania w empirycznych badaniach i analizach statystycznych, umożliwiając doskonalenie ich jakości i zasobu informacyjnego.

STATYSTYKA W PRAKTYCE

Dział ten dotyczy prac naukowych poświęconych nowatorskim zastosowaniom znanych narzędzi i modeli statystycznych w praktyce, analizie i statystycznej oceny określonych zjawisk społeczno-ekonomicznych i innych, a w szczególności artykułów wykorzystujących dane pochodzące z zasobów statystyki publicznej. Publikowane są tutaj także teksty sygnalizujące praktyczne problemy związane z: projektowaniem badań statystycznych, uzyskiwaniem, integracją i przetwarzaniem danych oraz generowaniem wyników informacji statystycznych i kontrolą ich ujawniania wraz z propozycjami efektywnych metod rozwiązywania owych problemów.

EDUKACJA STATYSTYCZNA

Artykuły publikowane w tym dziale dotyczą metod i efektów nauczania statystyki oraz popularyzacji myślenia statystycznego. W szczególności odnosi się to do problemów związanych z kształceniem w zakresie stosowania statystyki na wszystkich poziomach edukacji, a także wykorzystywania nowoczesnych idei i metod dydaktycznych (w tym eksperymentów i pokazów) oraz pomocy naukowych (np. komputerów, Internetu i innych urządzeń) w nauczaniu statystyki. Szczególną uwagę koncentruje się tutaj na rozumieniu prawdopodobieństwa i statystyki, badaniach z zakresu nauczania statystyki, postaw i zachowań społecznych w odniesieniu do statystyki, jak również rozumieniu informacji statystycznych. Ponadto ukazywane są problemy związane z prezentacją danych statystycznych oraz ich interpretacją w powszechnym obiegu informacyjnym (np. w środkach społecznego przekazu).

STATYSTYKA W SPOŁECZEŃSTWIE INFORMACYJNYM

Jest to blok tematyczny zawierający artykuły z zakresu wykorzystania narzędzi informatycznych do użytkowania i przetwarzania informacji statystycznych, naliczania danych wyników, ich prezentacji i rozpowszechniania oraz dotyczące nowoczesnych technik programistycznych, interaktywnych i komunikacyjnych umożliwiających potencjalnym użytkownikom danych statystycznych ich wykorzystanie w oczekiwanym przez siebie zakresie i formie. W dziale tym przedstawiane mogą być również artykuły dotyczące: wykorzystania technologii informacyjnych i komunikacyjnych (ICT), gospodarki opartej na wiedzy, problematyki innowacyjności, zagadnień dotyczących przepływu informacji we współczesnym społeczeństwie (w tym z użyciem Internetu) oraz przetwarzania i analizy zagadnień związanych z Big Data.

Z DZIEJÓW STATYSTYKI

Prace należące do tego działu tematycznego poświęcone są historii prowadzenia obserwacji statystycznych, rozwoju i doskonalenia ich metodologii oraz narzędzi. Ponadto zamieszczane są opisy wartościowych faktów dotyczących życia i osiągnięć zawodowych wybitnych statystyków, jak również wiodących instytucji i organizacji statystycznych w Polsce i za granicą.

INFORMACJE. PRZEGLĄDY. RECENZJE

Dział ten obejmuje informacje o najważniejszych wydarzeniach w życiu statystyki polskiej i międzynarodowej, działalności Rady Statystyki oraz z życia Polskiego Towarzystwa Statystycznego, a także sprawozdania z prestiżowych konferencji naukowych, recenzje książek naukowych i popularnonaukowych z zakresu statystyki i ekonometrii, jak również rekomendacje nowych, istotnych i ciekawych pozycji wydawniczych dotyczących tego obszaru wiedzy. Jest to jedyna część czasopisma zawierająca teksty niemające charakteru artykułów naukowych.