

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 POSZCZĘGÓLNYCH 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 d 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 użytkownika/mieszkania/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 OGÓLEM 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 niezarobkowych 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	rozkłady finansowe	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	żywność i napoje bezalkoholowe, użytkowanie mieszkań, rozchody na gospodarstwo rolne	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	odzież i obuwie	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	rozkłady kapitałowe	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.