

WIADOMOŚCI STATYSTYCZNE

CZASOPISMO GŁÓWNEGO URZĘDU STATYSTYCZNEGO
I POLSKIEGO TOWARZYSTWA STATYSTYCZNEGO

STUDIA METODOLOGICZNE

Beata KASPRZYK, Barbara FURA

Wykorzystanie modeli logitowych do identyfikacji gospodarstw domowych zagrożonych ubóstwem

Problem identyfikacji i badania ubóstwa w społeczeństwie polskim jest wciąż aktualny. W artykule przedstawiono wyniki badań dotyczących identyfikacji gospodarstw domowych ubogich i nieubogich (w szczególności czynników społeczno-ekonomicznych określających te poziomy) za pomocą modeli logitowych. Badanie przeprowadzono na próbie losowej, w której jednostkami badanymi były gospodarstwa domowe z woj. podkarpackiego. Do analizy danych wykorzystano model regresji logistycznej. Pozwoliło to na ilościową analizę wpływu determinant sprzyjających ubóstwu ekonomicznemu.

IDENTYFIKACJA UBÓSTWA

Problem ubóstwa w sensie ekonomicznym związany jest z zasadniczą kwestią jego identyfikacji i pomiaru. S. Rowntree zdefiniował ubóstwo jako stan, w którym całkowite dochody są wystarczające zaledwie do zaspokojenia minimum potrzeb fizycznych (Watson, 2000). Badacz ten spopularyzował pojęcie linii ubóstwa, które stało się z jednej strony podstawową kategorią metodologiczną

w badaniach nad ubóstwem, a z drugiej — kategorią praktyczną, na podstawie której ustalane są działania w dziedzinie zabezpieczenia społecznego¹.

Linia ubóstwa dzieli populację na ubogich i nieubogich. W literaturze dostępne są trzy odrębne metody jej wyznaczania: absolutna, względna oraz oparta na subiektywnych miarach. Za pomocą metody absolutnej linia ubóstwa jest ustalana na poziomie minimalnej ilości zasobów w danym momencie, a jej położenie jest uaktualniane jedynie w wyniku zmian cen. Metoda względna ustala linię ubóstwa jako miejsce w rozkładzie dochodów i wydatków, stąd jej położenie jest dostosowywane automatycznie do zmian w standardzie życia. Natomiast subiektywne metody ustalania linii ubóstwa oparte są na opiniach dotyczących minimalnego dochodu czy minimalnego poziomu wydatków, które są konieczne do zaspokojenia podstawowych potrzeb. Subiektywne ustalanie linii ubóstwa jest rzadziej stosowane (Zheng, 2001).

Główne pytania naukowców ciągle dotyczą określenia pojęcia oraz możliwości obiektywnego wyznaczania, mierzenia i zapobiegania ubóstwu. Problem stanowi zmienność w rozumieniu zjawiska, które różni się w zależności od stopnia rozwoju społeczeństw, poglądów badaczy czy celów prowadzonych badań (Baczewski, 2008). Badania w zakresie jakości i poziomu życia społeczeństwa koncentrują się na diagnozie rozmiarów i czynników ubóstwa (biedy) czy wykluczenia społecznego (ekskluzji).

Jedna z propozycji badania poziomu życia (zgodna z ustaleniami Rady Europejskiej oraz Komisji Europejskiej) proponuje i mierzy zjawiska związane z wykluczeniem i ubóstwem ekonomicznym w krajach Unii Europejskiej (UE) poprzez analizę w zakresie dochodów (Atkinson i in., 2002). Również w Polsce GUS opracowuje listę wskaźników charakteryzujących wykluczenie ekonomiczne, a także wykonuje szacunki w zakresie wskaźników zagrożenia ubóstwem czy w zakresie wskaźników kwintylowego zróżnicowania dochodów. W latach 80. XX w. nastąpiła reorientacja badań społecznych, polegająca na przejściu od obszernych zestawów wskaźników do badań szerzej wykorzystujących miary subiektywne. Wobec tego powstała konieczność interpretacji wyników badań w zakresie pomiaru ubóstwa na tle innych wskaźników (demograficznych, ekonomicznych, przestrzennych i politycznych). Natomiast do analizy wpływu cech typologicznych (determinant) na poziom życia można w szczególności stosować modele logitowe i probitowe (Anselin-Dorantes, Ser-rando-Padial, 2010; Maaden, Simpson, 1996).

Celem artykułu jest omówienie ustalania linii ubóstwa, która jest wyznaczana na podstawie mierników: absolutnych, względnych czy też subiektywnych miar ubóstwa. W opracowaniu przedstawiono wyniki badań dotyczących identyfika-

¹ Definiowanie ubóstwa jest wielopłaszczyznowe i obszerne. Definicje wskazują na ubóstwo jako na problem o charakterze ekonomicznym, ale także fizycznym, psychologicznym, społecznym itd. (Sen, 2002). Nasze rozważania i analiza danych empirycznych skupione są na wymiarze ekonomicznym ubóstwa (czyli braku i niedostatku środków materialnych w gospodarstwie domowym).

cji ubóstwa ekonomicznego, w zależności od różnych kryteriów demograficzno-społecznych, na przykładzie gospodarstw domowych z woj. podkarpackiego.

USTALANIE LINII (GRANIC) UBÓSTWA

Główne dyskusje dotyczące ubóstwa toczą się wokół tego, czy ubóstwo należy postrzegać jako zjawisko o charakterze absolutnym czy relatywnym. Podejście absolutne oznacza, że ubogie są te gospodarstwa, które nie mają możliwości zaspokojenia potrzeb uznanych w danych warunkach za podstawowe.

W podejściu względnym ubóstwo uważane jest za formę nierówności, dystansu między poziomem życia poszczególnych grup społecznych. Ubogimi będą te rodziny, dla których poziom życia jest znacznie niższy w porównaniu do pozostałych grup społeczeństwa. Ubóstwo można zatem rozpatrywać i oceniać także w ujęciu obiektywnym i subiektywnym.

Poszukiwanie uzasadnionej naukowo granicy ubóstwa materialnego, w nurcie definiowania ubóstwa bezwzględnego i obiektywnego, oznacza określenie linii (granic dochodu) wyznaczających minimum egzystencji oraz minimum socjalnego, które stanowią oficjalne linie ubóstwa. Minimum to stanowi dolne kryterium ubóstwa skrajnego. Z kolei minimum socjalne stanowi taki poziom dochodów, który pozwala normalnie uczestniczyć w życiu społeczeństwa i nie pozwala popaść w ubóstwo. Zaspokojenie potrzeb na niższym poziomie od wyznaczonego może prowadzić do zerwania więzi społecznych. Minimum socjalne należy uznać za graniczny wskaźnik integracji i zarazem wykluczenia społecznego. Poziom minimum socjalnego oblicza się na podstawie normatywnie zdefiniowanego koszyka potrzeb społecznie niezbędnych. Minimum socjalne nie jest linią ubóstwa, lecz linią ostrzegającą, że poniżej przyjętej granicy powstaje zagrożenie ubóstwem. J. Veit-Wilson (1994) nazywa taką linię standardem dochodu minimalnego.

Dysponując poziomem linii ubóstwa można podać frakcję osób poniżej minimum egzystencji, czyli liczbę osób skrajnie ubogich. Z kolei sferę niedostatku stanowią osoby, których dochód jest poniżej linii minimum socjalnego (*Warunki życia...*, 1999—2007).

Ustalanie poziomu ubóstwa odbywa się na podstawie granicy dochodu ustalonej arbitralnie i oddzielnie dla każdego kraju, w sposób odmienny. Istnieje spór o wielkość tego świadczenia oraz kto ma decydować o jego wielkości: eksperci, społeczeństwo, politycy czy wreszcie sami zainteresowani. Nadal przeważają opinie ekspertów, którzy określają składniki i poziom progu ubóstwa (Kot, 2000). Natomiast według zasady P. Desai'a (1995): *poziom progu ubóstwa, tj. określonych składowych poziomu życia powinien być wyznaczony przez społeczeństwo*. Jego zdaniem eksperci są potrzebni jedynie w celu pobrania reprezentatywnej próby statystycznej, na podstawie której szacuje się rzeczywiste zachowania konsumpcyjne. Stosowane linie ubóstwa są ustalane w odmienny sposób, a źródłem sprzeczności jest różnorodność metod opracowanych przez

ekonomistów i ekspertów. Na tle tych metodologicznych sporów przedstawiono analizę czynników determinujących problem ubóstwa w gospodarstwach domowych woj. podkarpackiego.

ZAŁOŻENIA TEORETYCZNE MODELU LOGITOWEGO

Przedmiotem wyjaśnienia w dwumianowych modelach logitowych jest prawdopodobieństwo p_i przyjmowania przez zmienną y_i jednej z dwu możliwości. Jeżeli y_i przyjmuje wartość 0 lub 1 (dla odpowiednich wariantów zmiennej jakościowej), gdzie i oznacza jednostkę ekonomiczną, to: $P(y_i = 1) = p_i$, $P(y_i = 0) = 1 - p_i$. Przyjmuje się wówczas, że to prawdopodobieństwo jest funkcją wektora zmiennych objaśniających x_i oraz wektora parametrów β , czyli ich iloczynu skalarnego $x_i^T \beta$. W zależności od typu funkcji F wyróżnia się kilka rodzajów modeli dwumianowych: liniowy model prawdopodobieństwa, model probitowy i model logitowy (Aldrich, Nelson, 1984). W modelu logitowym przyjmuje się założenie, że prawdopodobieństwo p_i odpowiada dystrybucie rozkładu logistycznego (Cramer, 2004):

$$p_i = F(x_i^T \beta) = \frac{1}{1 + \exp(-x_i^T \beta)} = \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)}$$

Parametry modelu są szacowane przy użyciu metody największej wiarygodności (MNW).

Jeżeli obserwacje są niezależne, to funkcja wiarygodności n -elementowej próby ma postać:

$$L = \prod_{y=1} p_i \prod_{y=0} (1 - p_i) = \prod_{i=1}^n [F(x_i^T \beta)]^{y_i} [1 - F(x_i^T \beta)]^{1-y_i}$$

Po zlogarytmowaniu funkcji wiarygodności otrzymujemy (Chow, 1995):

$$\ln L = \sum_{i=1}^n \{y_i \ln F(x_i^T \beta) + (1 - y_i) \ln [1 - F(x_i^T \beta)]\}$$

Przyrównując gradient $\frac{\partial \ln L}{\partial \beta}$ do zera, uzyskujemy układ równań normalnych, po którego rozwiązaniu przy użyciu metod numerycznych wyliczane są wartości wektora β . Po uzyskaniu ich wartości wyznacza się wartość $x_i^T \beta$, a następnie wartość dystrybuanty rozkładu logistycznego.

Poprawność oszacowanego modelu można sprawdzić wykorzystując test ilorazu wiarygodności, służący do weryfikacji hipotezy zerowej dotyczącej modelu tylko z wyrazem wolnym (wszystkie parametry modelu poza wyrazem wolnym są równe zero). Statystyka testu wyrażona jest wzorem:

$$\chi^2 = 2(\ln L_{UR} - \ln L_R)$$

gdzie:

L_{UR} — wartość funkcji wiarygodności dla modelu pełnego,

L_R — wartość funkcji wiarygodności dla modelu zawierającego jedynie wyraz wolny.

Statystyka ta ma rozkład χ^2 z liczbą stopni swobody równą liczbie zmiennych objaśniających model. Wśród licznych miar dopasowania modelu do danych empirycznych znajduje się wykorzystany w artykule współczynnik Aldricha-Nelsona: $R_{A-N}^2 = \frac{2(\ln L_U - \ln L_{UR})}{2(\ln L_U - \ln L_{UR}) + n}$, będący odpowiednikiem współczynnika determinacji R^2 (Gruszczyński, 2001).

WYNIKI BADAŃ EMPIRYCZNYCH — MODELOWANIE LOGITOWE WYBRANYCH CZYNNIKÓW UBÓSTWA

W celu identyfikacji czynników ubóstwa wykorzystano materiał empiryczny uzyskany z ankiety badania na losowej próbie gospodarstw domowych z woj. podkarpackiego². Porównanie ekwiwalentnych miesięcznych dochodów rozporządzalnych gospodarstw domowych z 2008 r. z bezwzględną granicą ubóstwa w Polsce³ pozwoliło podzielić badane gospodarstwa na ubogie i nieubogie. Próba badawcza liczyła 373 gospodarstwa domowe i objęła 1270 osób (968 dorosłych osób i 302 dzieci). Do kwalifikacji danego gospodarstwa domowego *i* jako ubogie/nieubogie wykorzystano zmienną binarną *Y*, która jako zmienna zależna charakteryzowała kategorię gospodarstwa domowego — ubogie/nieubogie we-

² Reprezentatywność badanej próby losowej w stosunku do populacji generalnej uzyskano ze względu na skład osobowy gospodarstw domowych, otrzymując w miarę zadowalające wskaźniki struktury badanej próby (w nawiasach podano wskaźniki struktury dla populacji generalnej gospodarstw domowych w woj. podkarpackim): dla gospodarstw 1-osobowych — 11,8% (22,45%); 2-osobowych — 18,2% (22,01%); 3-osobowych — 20,9% (19,83%); 4-osobowych — 30,0% (17,18%); 5-osobowych i więcej — 19,0% (17,93%) (*Rocznik...*, 2008).

³ Respondenci byli proszeni o wskazanie dochodu rozporządzalnego gospodarstwa domowego w kwietniu 2008 r. Dochód rozporządzalny stanowił dochód brutto (obejmujący wszystkie przychody pieniężne i niepieniężne) po odliczeniu podatku i powiększeniu o wszystkie dodatki i transfery socjalne i prywatne.

dług dochodu przypadającego na osobę w badanym gospodarstwie. Zmienną Y zdefiniowano następująco:

$$y_i = \begin{cases} 0 & \text{jeżeli } i\text{-te gospodarstwo domowe jest ubogie} \\ 1 & \text{jeżeli } i\text{-te gospodarstwo domowe jest nieubogie} \end{cases}$$

gdzie $i = 1, \dots, n$ — liczba badanych gospodarstw domowych.

Wartości zmiennej zależnej Y uzyskano dokonując porównania dochodu przypadającego na osobę z ustaloną arbitralnie linią ubóstwa z . Przyjęto, że zmienna losowa $y_i = 0$, gdy dochód na osobę w gospodarstwie domowym jest mniejszy lub równy linii ubóstwa z (gospodarstwo jest ubogie), natomiast $y_i = 1$, gdy dochód na osobę w gospodarstwie domowym jest większy od linii ubóstwa z (gospodarstwo jest nieubogie). Jako linię ubóstwa z przyjęto wielkość minimum socjalnego, wyznaczoną dla określonych typów gospodarstw domowych przez Instytut Pracy i Spraw Socjalnych (IPiSS). Przyjęto wartość średnioroczną z 2008 r.⁴ Należy mieć na uwadze, że zaspokojenie potrzeb na poziomie minimum socjalnego nie oznacza ubóstwa, ale dostateczne, choć raczej skromne warunki bytowania, adekwatne do przeciętnego poziomu życia w kraju.

W modelowaniu określono podział na gospodarstwa domowe, które funkcjonowały w biedzie bądź nie czy też były raczej w sferze niedostatku niż w skrajnej biedzie. Na podstawie posiadanej wiedzy merytorycznej o czynnikach wpływających na ryzyko „popadnięcia” w ubóstwo skonstruowano listę potencjalnych zmiennych objaśniających o charakterze demograficzno-społecznym. Wykluczono podstawową zmienną — dochód rozporządzalny przypadający na osobę, gdyż stanowił on podstawę kwalifikacyjną zmiennej dychotomicznej Y . Wejściowy zbiór zmiennych niezależnych stanowiły zmienne charakteryzujące społeczno-demograficzne cechy danego gospodarstwa domowego: przynależność do grupy społeczno-ekonomicznej, wykształcenie głowy gospodarstwa, miejsce zamieszkania (wieś/miasto), liczba osób w gospodarstwie domowym, liczba dzieci w rodzinie, rodzina niepełna (tak/nie), wiek członków rodziny, posiadanie oszczędności (tak/nie) i struktura wydatków. Selekcji zmiennych objaśniających dokonano na podstawie macierzy korelacji, wybierając zmienne znacząco skojarzone ze zmienną zależną Y . Ze zbioru wstępnych zmiennych objaśniających wzięto pod uwagę wyłącznie te, które uznano za w miarę słabo skorelowane ze sobą (tabl. 1).

⁴ W analizie przyjęto następujące wartości zmodyfikowanego minimum socjalnego na 1 osobę dla gospodarstw pracowniczych: 1-osobowe — 865,1 zł, 2-osobowe — 714,6 zł, 3-osobowe — 724,75 zł, 4-osobowe — 685,4 zł, 5-osobowe — 668,5 zł (Kurowski, 2009).

TABL. 1. MACIERZ WSPÓŁCZYNNIKÓW KORELACJI DLA WYBRANYCH ZMIENNYCH OBJAŚNIAJĄCYCH

Wyszczególnienie	Współczynniki korelacji			
	oszczędności	liczba dorosłych	liczba dzieci	wykształcenie
Oszczędności	1,000000	-0,165300	-0,085412	-0,027001
Liczba dorosłych	-0,165300	1,000000	0,282214	-0,113524
Liczba dzieci	-0,085412	0,282214	1,000000	-0,274372
Wykształcenie	-0,027001	-0,113524	-0,274372	1,000000

Źródło: obliczenia własne.

Zbiór zmiennych objaśniających modelu ograniczył się zatem do kilku podstawowych zmiennych: oszczędności (zmienna dychotomiczna: 0 — gospodarstwo nie posiada oszczędności, 1 — gospodarstwo posiada oszczędności), liczba dorosłych, liczba dzieci, wykształcenie (1 — podstawowe, 2 — zasadnicze, 3 — średnie, 4 — wyższe).

Wyniki estymacji modelu logitowego dla danych ogółem przedstawiono w tabl. 2. Istotność parametrów statystycznych modelu zweryfikowano na podstawie testu *t*-Studenta, przyjęto poziom istotności $\alpha = 0,05$.

TABL. 2. WYNIKI ESTYMACJI MODELU LOGITOWEGO DLA DANYCH OGÓŁEM

Wyszczególnienie	Zmienne niezależne				
	stała	oszczędności	liczba dorosłych	liczba dzieci	wykształcenie
Ocena	0,626387	1,30288	-0,96412	-0,76201	1,04530
Błąd standardowy	0,637568	0,32307	0,14038	0,15150	0,19133
Statystyka <i>t</i>	0,982463	4,03277	-6,86801	-5,02979	5,46342
Poziom <i>p</i>	0,326515	0,00007	0,00000	0,00000	0,00000
-95%	-0,627347	0,66758	-1,24017	-1,05992	0,66907
+95%	1,880122	1,93819	-0,68808	-0,46410	1,42154
Chi-kwadrat Walda	0,965233	16,26323	47,16953	25,29877	29,84890
Poziom <i>p</i>	0,325879	0,00006	0,00000	0,00000	0,00000
Iloraz szans	1,870839	3,67989	0,38132	0,46673	2,84426
-95%	0,534007	1,94951	0,28934	0,34648	1,95242
+95%	6,554302	6,94613	0,50254	0,62870	4,14348

Źródło: jak przy tabl. 1.

Parametry modelu dla zmiennych: „oszczędności”, „liczba dorosłych”, „liczba dzieci” i „wykształcenie” okazały się statystycznie istotne, zatem wymienione zmienne miały wpływ na prawdopodobieństwo zakwalifikowania gospodarstwa domowego do grupy gospodarstw ubogich/nieubogich⁵. Ujemne szacunki parametrów dla zmiennych „liczba dorosłych” i „liczba dzieci” wskazywały, że wzrost wartości tych zmiennych powodował zmniejszenie się prawdopodobień-

⁵ Nieistotny okazał się tylko wyraz wolny modelu logitowego ($p = 0,33$).

stwa zaklasyfikowania gospodarstwa do grupy gospodarstw nieubogich. Dodatnie wartości parametrów przy zmiennych „wykształcenie” i „oszczędności” potwierdzają, że wzrost poziomu tych zmiennych (dla „oszczędności” — wystąpienie) zwiększał prawdopodobieństwo zaklasyfikowania gospodarstwa domowego do grupy gospodarstw nieubogich⁶.

Interpretacja parametrów w modelu logitowym jest dość trudna, a zatem podlegają interpretacji wyniki ilorazów szans⁷. Oszacowany jednostkowy iloraz szans dla zmiennej „oszczędności” równy 3,68 potwierdził przypuszczenia, że gospodarstwa domowe posiadające oszczędności mają prawie czterokrotnie większe szanse na uniknięcie stanu ubóstwa. Z kolei wzrost poziomu wykształcenia (np. ze średniego na wyższe) prawie trzykrotnie zwiększa szanse na zakwalifikowanie gospodarstwa do grupy gospodarstw nieubogich — iloraz szans dla zmiennej „wykształcenie” wyniósł 2,84. Model logitowy dla danych ogółem uznano za wystarczająco dobrze dopasowany do danych empirycznych (zestawienie).

**ZESTAWIENIE WYNIKÓW WERYFIKACJI MODELI LOGITOWYCH
DLA GOSPODARSTW DOMOWYCH OGÓŁEM**

Liczba 0 (ubogie n_i)	Liczba 1 (nieubogie n_i)	Wartość końcowa funkcji straty	-2 log wiarygodności	Stała (wyłącznie)	R^2_{A-N}	χ^2	poziom p
126 (33,78%)	247 (66,22%)	161,2603	322,5206	477,1167	0,2935	154,5960	0,000

Źródło: jak przy tabl. 1.

Wartość statystyki testu zgodności wartości obserwowanych chi-kwadrat (χ^2) dla wyznaczonego modelu, wraz z poziomem prawdopodobieństwa testowego $p < 0,05$ wskazują, że model różni się istotnie od modelu tylko z wyrazem wolnym. Ponadto statystyka -2log wiarygodności wyniosła dla modelu regresji logistycznej 322,52, a dla modelu tylko z wyrazem wolnym — 477,12. Różnica między nimi — wartość końcowa, czyli funkcja straty (161,26) jest również miarą dopasowania modelu, stanowiąc wytyczną wyboru modelu.

Na podstawie otrzymanego modelu logitowego można dokonać oszacowania prawdopodobieństwa, że dane gospodarstwo domowe zostanie zaklasyfikowane jako „nieubogie”, w zależności od wyróżnionych w modelowaniu cech społeczno-demograficznych. Przykładowo, można przyjąć prawdopodobieństwo, że

⁶ Z wyznaczonych przedziałów ufności wynika, że w 95% możemy być pewni, iż wartości rzeczywiste zmiennej „oszczędności” znajdują się w przedziale (0,67; 1,94), zmiennej „liczba dorosłych” — (-1,24; -0,69), „liczba dzieci” — (-1,06; -0,46), a zmiennej „wykształcenie” w przedziale (0,67; 1,42).

⁷ Szansę należy rozumieć jako szacunek prawdopodobieństwa tego, że dane zdarzenie wystąpi do prawdopodobieństwa zdarzenia przeciwnego.

gospodarstwo domowe o następujących cechach: ma oszczędności, liczba dorosłych — 1, liczba dzieci — 2, wykształcenie średnie głowy gospodarstwa domowego — jest nieubogie („normalny” poziom życia). Wartość tego prawdopodobieństwa, obliczona na podstawie oszacowanego modelu, to 0,93. Jeżeli w tym samym gospodarstwie zmianie ulegnie liczba dzieci (wzrośnie do trojga), to prawdopodobieństwo braku ubóstwa spadnie do 0,86. W tym przypadku wzrost liczby dzieci zmniejszy szanse na funkcjonowanie gospodarstwa w sferze gospodarstw nieubogich.

Jeśli z kolei rozważymy inne gospodarstwo o cechach: brak oszczędności, liczba dorosłych — 2, liczba dzieci — 3, wykształcenie zasadnicze zawodowe, wówczas wartość rozważanego prawdopodobieństwa ($Y=1$) wyniesie 0,18. Zakładając, że liczba dzieci wzrośnie do czterech, to prawdopodobieństwo pozostawania poza sferą ubóstwa zmniejszy się do 0,09. Niewątpliwie takie gospodarstwo ma znamiona gospodarstwa ubogiego. Jeżeli jednak, dla tego samego gospodarstwa z dwójką dzieci, wykształcenie głowy gospodarstwa zwiększy się do poziomu średniego, to prawdopodobieństwo braku ubóstwa zwiększy się do 0,58, a jeżeli do poziomu wyższego, to wyniesie 0,79.

Jak wynika z symulacji istotną determinantą pozostawania poza lub w sferze ubóstwa jest poziom wykształcenia głowy gospodarstwa. Wniosek ten stanowił przesłankę do przeprowadzenia modelowania logitowego dla grup gospodarstw według poziomu wykształcenia głowy gospodarstwa (tabl. 3). Po symulacjach mających na celu dobór zmiennych do modelowania, przyjęto ostatecznie trzy zmienne niezależne, których parametry oszacowano wykorzystując modelowanie logitowe.

TABL. 3. WYNIKI ESTYMACJI MODELI LOGITOWYCH DLA GOSPODARSTW DOMOWYCH WZGLĘDEM WYKSZTAŁCENIA

Wyszczególnienie	Zmienne			
	stała	oszczędności	liczba osób ogółem	liczba dzieci
Wykształcenie zasadnicze ($n = 80$)				
Ocena	1,78323	1,41789	-0,73249	—
Błąd standardowy	0,76338	0,60109	0,21588	—
Statystyka t	2,33597	2,35886	-3,39300	—
Poziom p	0,02210	0,02087	0,00109	—
Chi-kwadrat Walda	5,45676	5,56420	11,51243	—
Poziom p	0,01950	0,01834	0,00069	—
Iloraz szans	5,94903	4,12841	0,48071	—
Wykształcenie średnie ogólne ($n = 171$)				
Ocena	3,21202	1,170623	-0,78360	-0,74464
Błąd standardowy	0,62035	0,448694	0,18200	0,18515
Statystyka t	5,17774	2,608954	-4,30549	-4,02174
Poziom p	0,00000	0,009906	0,00003	0,00009
Chi-kwadrat Walda	26,80902	6,806641	18,53722	16,17441
Poziom p	0,00000	0,009086	0,00002	0,00006
Iloraz szans	24,82909	3,224000	0,45676	0,47491

TABL. 3. WYNIKI ESTYMACJI MODELI LOGITOWYCH DLA GOSPODARSTW DOMOWYCH WZGLĘDEM WYKSZTAŁCENIA (dok.)

Wyszczególnienie	Zmienne			
	stała	oszczędności	liczba osób ogółem	liczba dzieci
Wykształcenie wyższe ($n = 103$)				
Ocena	5,281	—	-1,00889	—
Błąd standardowy	1,125	—	0,34519	—
Statystyka t	4,696	—	-2,92268	—
Poziom p	0,000	—	0,00428	—
Chi-kwadrat Walda	22,054	—	8,54204	—
Poziom p	0,000	—	0,00347	—
Iloraz szans	196,590	—	0,36462	—

Źródło: jak przy tabl. 1.

Z analizy wykluczono gospodarstwa domowe z podstawowym wykształceniem głowy gospodarstwa, gdyż wyznaczenie modelu było trudne, ze względu na małą liczbę obserwacji ($n = 19$). Parametry modelu logitowego (dla gospodarstw domowych z wykształceniem zasadniczym) były istotnie różne od zera zarówno dla zmiennych objaśniających: „oszczędności”, „liczba osób ogółem”, jak i dla wyrazu wolnego. Na uwagę zasługuje wysoki poziom jednostkowego ilorazu szans dla zmiennej „oszczędności”. Okazało się, że w przypadku posiadania przez gospodarstwo domowe oszczędności, prawdopodobieństwo zakwalifikowania do grupy gospodarstw nieubogich jest przeszło 4-krotnie większe (iloraz szans — 4,13) w stosunku do gospodarstw bez oszczędności.

W przypadku grupy gospodarstw domowych z wykształceniem średnim ogólnym, najlepsze dopasowanie logitowe otrzymano dla kombinacji zmiennych niezależnych: „oszczędności”, „liczba osób dorosłych” oraz „liczba dzieci”, gdzie zarówno parametry przy zmiennych, jak i wyraz wolny były istotnie różne od zera. Również w tej grupie gospodarstw prawdopodobieństwo pozostawania poza strefą ubóstwa było przeszło 3-krotnie większe w przypadku posiadania oszczędności.

W grupie gospodarstw domowych z wykształceniem wyższym tylko zmienna „liczba dorosłych” miała istotny wpływ na prawdopodobieństwo sklasyfikowania gospodarstwa jako nieubogie, istotny okazał się również wyraz wolny. Ujemne oceny parametrów dla zmiennych „liczba osób ogółem”, „liczba dorosłych” i „liczba dzieci” wskazują, że wzrost wartości tych zmiennych powodował zmniejszenie prawdopodobieństwa sklasyfikowania gospodarstwa do grupy gospodarstw nieubogich. Dodatnie wartości parametrów dla zmiennej „oszczędności” oznaczają, że wzrost poziomu tej zmiennej (wystąpienie) zwiększał prawdopodobieństwo pozostawania gospodarstwa poza sferą ubóstwa. Weryfikacja statystyczna modeli regresji logistycznej wykazała, że są one poprawnie dopasowane do danych empirycznych (tabl. 4).

**TABL. 4. WYNIKI WERYFIKACJI MODELI LOGITOWYCH
WEDŁUG POZIOMU WYKSZTAŁCENIA**

Wykształcenie	Liczba 0 (ubogie n_i)	Liczba 1 (nieubogie n_i)	Wartość końcowa funkcji straty	-2 log wiary- godności	R^2_{A-N}	χ^2	poziom p
Zasadnicze	49 (61,25%)	31 (38,75%)	43,00	86,003	0,2065	20,816	0,000
Średnie ogólne	61 (35,67%)	110 (64,33%)	90,61	181,219	0,1957	41,598	0,000
Wyższe	7 (6,80%)	96 (93,2%)	20,42	40,840	0,0910	10,317	0,001

Źródło: jak przy tabl. 1.

Wartości statystyki testu zgodności (χ^2) oraz wartości prawdopodobieństwa testowego p potwierdzają statystyczną istotność wyznaczonych modeli. Dopasowanie modeli logitowych wyznaczonych względem wykształcenia — zasadnicze i średnie ogólne — uznano za wystarczające (na podstawie R^2_{A-N}). Ostatecznie uznano, że przyjęte zmienne istotnie wpływały na poziom prawdopodobieństwa braku ubóstwa ($Y = 1$) w danej grupie gospodarstw według wykształcenia. Wysoka wartość końcowa funkcji straty (z wyjątkiem grupy gospodarstw z wykształceniem wyższym) wskazuje na poprawność modeli logitowych dla gospodarstw domowych z wykształceniem głowy gospodarstwa zasadniczym i średnim ogólnym.

W kolejnym etapie analizy logitowej za kryterium podziału gospodarstw domowych przyjęto miejsce zamieszkania gospodarstwa (wieś/miasto). Analiza logitowa dostarczyła informacji dotyczącej prawdopodobieństwa uznania gospodarstwa domowego za nieubogie/ubogie. Na podstawie prawie identycznej liczby prób wyznaczono osobno modele logitowe dla gospodarstw wiejskich i miejskich, oceniając prawdopodobieństwo teoretyczne pozostawania poza sferą ubóstwa (tabl. 5).

TABL. 5. WYNIKI ESTYMACJI MODELI LOGITOWYCH WEDŁUG MIEJSCA ZAMIESZKANIA

Wyszczególnienie	Stała	Oszczędności	Liczba dorosłych	Liczba dzieci	Wykształcenie
$n = 187$ (gospodarstwa zamieszkałe na wsi)					
Ocena	1,01762	1,311599	-0,96433	-0,63940	0,82027
Błąd standardowy	0,75441	0,436773	0,17821	0,19456	0,25471
$t(182)$	1,34891	3,002931	-5,41117	-3,28640	3,22040
Poziom p	0,17904	0,003050	0,00000	0,00122	0,00152
Chi-kwadrat Walda	1,81955	9,017597	29,28071	10,80040	10,37097
Poziom p	0,17738	0,002676	0,00000	0,00102	0,00128
Iloraz szans	2,76661	3,712105	0,38124	0,52761	2,27112

TABL. 5. WYNIKI ESTYMACJI MODELI LOGITOWYCH WEDŁUG MIEJSCA ZAMIESZKANIA (dok.)

Wyszczególnienie	Stała	Oszczędności	Liczba dorosłych	Liczba dzieci	Wykształcenie
$n = 186$ (gospodarstwa zamieszkałe w miastach)					
Ocena	-0,07242	1,287758	-0,90580	-0,85992	1,2850
Błąd standardowy	1,28389	0,489822	0,24273	0,24235	0,3380
t (181)	-0,05641	2,629033	-3,73176	-3,54829	3,8023
Poziom p	0,95508	0,009299	0,00025	0,00049	0,0002
Chi-kwadrat Walda	0,00318	6,911815	13,92603	12,59038	14,4577
Poziom p	0,95502	0,008567	0,00019	0,00039	0,0001
Iloraz szans	0,93014	3,624650	0,40422	0,42320	3,6147

Źródło: jak przy tabl. 1.

Wartość statystyki dobroci dopasowania modeli logitowych do danych empirycznych przedstawiono w tabl. 6.

TABL. 6. WYNIKI WERYFIKACJI MODELI LOGITOWYCH WEDŁUG MIEJSCA ZAMIESZKANIA

Miejsce zamieszkania	Liczba 0 (ubogie n_i)	Liczba 1 (nieubogie n_i)	Wartość końcowa funkcji straty	-2 log wiarygodności	R^2_{A-N}	χ^2	poziom p
Wieś	83 (22,25%)	104 (27,88%)	91,78	183,57	0,2816	73,306	0,000
Miasto	43 (11,53%)	143 (38,34%)	68,474	136,95	0,2535	64,193	0,000

Źródło: jak przy tabl. 1.

Wartość statystyki testu zgodności (χ^2) wraz z poziomem prawdopodobieństwa testowego potwierdzają statystyczną istotność modeli logitowych „wieś”, „miasto”. Wyniki modelowania generalnie nie różnią się dla porównywanych grup gospodarstw, większe różnice dotyczą ocen parametrów „liczba dzieci” oraz „wykształcenie”. Okazuje się, że wzrost liczby dzieci w gospodarstwach miejskich w większym stopniu wpływa na zmniejszenie prawdopodobieństwa zaklasyfikowania gospodarstwa do grupy gospodarstw nieubogich niż w przypadku gospodarstw wiejskich. Z kolei wyższe wartości parametrów dla zmiennej „wykształcenie” w przypadku gospodarstw miejskich wskazują, że zmienna „wykształcenie” w większym stopniu wpływa na prawdopodobieństwo zaklasyfikowania gospodarstwa do grupy gospodarstw nieubogich niż w przypadku gospodarstw wiejskich. Wniosek ten potwierdza jednostkowy iloraz szans wynoszący dla modelu „miasto” — 3,61, a dla modelu „wieś” — 2,27. W przypad-

ku zmiennej „oszczędności” parametry oraz ilorazy szans są podobne, co oznacza, że posiadanie oszczędności zwiększa prawdopodobieństwo zakwalifikowania gospodarstwa do grupy nieubogich (prawie 4-krotnie).

W ostatnim etapie analizy regresji logitowej uwzględniono także wpływ liczby osób w gospodarstwie na poziom prawdopodobieństwa pozostania przez gospodarstwo domowe poza sferą ubóstwa. Wyniki modelowania logitowego przedstawiono dla gospodarstw 5-osobowych (tabl. 7).

TABL. 7. WYNIKI ESTYMACJI MODELU LOGITOWEGO DLA GOSPODARSTW DOMOWYCH 5-OSOBOWYCH

Wyszczególnienie	Zmienne			
	stała	oszczędności	wykształcenie	zamieszkanie
Ocena	-7,4026	3,2982	1,276	0,7437
Błąd standardowy	3,7151	1,1796	0,743	0,5909
<i>t</i> (112)	-1,9926	2,7960	1,717	1,2586
Poziom <i>p</i>	0,0555	0,0089	0,096	0,2179
Chi-kwadrat Walda	-14,9898	0,8891	-0,241	-0,4630
Poziom <i>p</i>	0,1846	5,7073	2,793	1,9505
Iloraz szans	3,9704	7,8177	2,949	1,5842

Końcowa strata 14,8428

$\chi^2(3) = 17,3308$

$R^2_{A-N} = 0,3376$

$p = 0,0006$

Źródło: jak przy tabl. 1.

Istotne statystycznie okazały się wyraz wolny (na poziomie istotności 0,1) oraz parametry przy zmiennych „oszczędności” (na poziomie 0,05) i „wykształcenie” (na poziomie 0,1). Na prawdopodobieństwo pozostania poza sferą ubóstwa dla grupy gospodarstw 5-osobowych nie miało istotnego wpływu miejsce zamieszkania. Stosunkowo wysokie, dodatnie wartości parametrów w przypadku zmiennych „oszczędności” i „wykształcenie” pozwoliły wskazać cechy gospodarstw domowych decydujące o prawdopodobieństwie braku ubóstwa (normalnej konsumpcji). Ponadto w przypadku gospodarstw domowych 5-osobowych wartość ilorazu szans dla zmiennej „wykształcenie” wyniosła 2,95, a dla zmiennej „oszczędności” — 7,82. Najlepsze dopasowanie modelu logitowego względem liczby osób uzyskano dla grupy gospodarstw domowych 5-osobowych ($R^2_{A-N} = 0,33$) (tabl. 8). W tablicy tej przedstawiono również wyniki weryfikacji pozostałych modeli regresji logistycznej według liczby osób w gospodarstwie domowym.

**TABL. 8. WYNIKI WERYFIKACJI MODELI REGRESJI LOGISTYCZNEJ
WEDŁUG LICZBY OSÓB**

Wyszczególnienie	Liczba 0 (ubogie n_i)	Liczba 1 (nieubogie n_i)	Wartość końcowa funkcji straty	R^2_{A-N}	χ^2	poziom p
1-osobowe ($n = 44$)	42 (95,45%)	2 (4,54%)	7,511	0,0276	1,250	0,535
2-osobowe ($n = 68$)	58 (85,29%)	10 (14,71%)	22,132	0,1556	12,525	0,006
3-osobowe ($n = 68$)	18 (23,08%)	60 (76,92%)	35,555	0,1622	13,161	0,004
4-osobowe ($n = 112$)	47 (41,96%)	65 (58,04%)	62,944	0,1912	26,472	0,000
5-osobowe ($n = 34$)	18 (52,94%)	16 (47,06%)	14,840	0,3376	17,330	0,001
6 i więcej osobowe ($n = 37$) ..	31 (83,78%)	6 (6,22%)	13,316	0,1427	6,160	0,046

Źródło: jak przy tabl. 1.

Statystycznie istotny był jedynie model dla gospodarstw jednoosobowych. Pozostałe modele różniły się istotnie od modeli tylko z wyrazem wolnym.

Podsumowanie

Zastosowanie modelowania logitowego pozwoliło na określenie czynników powodujących popadnięcie w ubóstwo, a następnie wyznaczenie prawdopodobieństwa ubóstwa dla przykładowego gospodarstwa domowego. Identyfikacja ubogich gospodarstw domowych na podstawie modeli regresji logistycznej pozwoliła sformułować następujące wnioski:

- cechy gospodarstw domowych: wykształcenie głowy gospodarstwa domowego, miejsce zamieszkania, liczba dzieci, liczba dorosłych, czy posiadanie lub nie oszczędności determinuje prawdopodobieństwo klasyfikacji gospodarstwa jako ubogie/nieubogie;
- w zależności od kryterium podziału gospodarstw domowych uzyskuje się zróżnicowane modele logitowe, z różnymi zmiennymi niezależnymi oraz parametrami i ocenami dopasowania modelu;
- w opisanych modelach zmienne „liczba dorosłych” i „liczba dzieci” mają wpływ na prawdopodobieństwo sklasyfikowania gospodarstwa do grupy ubogich/nieubogich, wzrost liczby osób w gospodarstwie domowym powoduje wzrost prawdopodobieństwa zaklasyfikowania gospodarstwa domowego jako ubogie;
- istotną determinantą pozostawania danego gospodarstwa domowego w obrębie/poza sferą ubóstwa jest poziom wykształcenia głowy gospodarstwa. Wzrost poziomu wykształcenia znacząco przyczyniał się do zmniejszenia prawdopodobieństwa popadnięcia gospodarstwa domowego w sferę ubóstwa. Jednocześnie poziom wykształcenia głowy gospodarstwa domowego miał większy wpływ na brak ubóstwa dla gospodarstw miejskich niż wiejskich;
- cechą charakteryzującą gospodarstwa nieubogie był fakt posiadania oszczędności.

Wyniki badań pokazują, że w woj. podkarpackim większy zasięg ubóstwa występuje przy niskim poziomie wykształcenia, wielodzietności, zamieszkiwaniu na wsi i braku oszczędności. Pauperyzacji gospodarstw domowych przejawiającej się zwiększonym ryzykiem popadnięcia w sferę ubóstwa, towarzyszą różnorodne czynniki. W celu ograniczania skali ubóstwa, rozpoznanie tych czynników jest istotne w budowaniu skutecznej polityki społecznej.

dr Beata Kasprzyk, dr Barbara Fura — Uniwersytet Rzeszowski

LITERATURA

- Aldrich J. H., Nelson F. D. (1984), *Linear probability, logit, and probit models*, „Quantitative Applications in the Social Sciences”, No. 07—045, Sage Publications, Newbury-London-New Delhi
- Anselin-Dorantes C., Serrano-Padial R. (2010), *Labor market flexibility and poverty dynamics*, „Labor Economics”, No. 17
- Atkinson T., Cantillon B., Marlier E., Nolan B. (2002), *Social Indicators The EU and Social Inclusion*, Oxford University Press, Oxford
- Baczewski G. (2008), *Oblicza biedy w zjednoczonej Europie*, Wydawnictwo UMCS, Lublin
- Chow G. C. (1995), *Ekonometria*, Wydawnictwo Naukowe PWN, Warszawa
- Cramer J. S. (2004), *The early origins of the logit model*, „Studies in History and Philosophy of Biological and Biomedical Sciences”, No. 35
- Desai P. (1995), *Poverty, Famine and Economic Development*, Edward Elgar Publishing Ltd, Aldershot
- Gruszczyński M. (2001), *Modele i prognozy zmiennych jakościowych w finansach i bankowości*, „Monografie i Opracowania”, nr 490, Oficyna Wydawnicza SGH, Warszawa
- Kot S. M. (2002), *Ekonometryczne modele dobrobytu*, Wydawnictwo Naukowe PWN, Warszawa
- Kurowski P. (2009), *Badania nad poziomem i strukturą zmodyfikowanego minimum socjalnego w 2008 r.*, IPISS, Warszawa
- Maaden G., Simpson M. (1996), *A probit model of household broadband service subscription intentions: A regional analysis*, „Information Economics and Policy”, No. 8
- Rocznik Demograficzny 2007* (2008), GUS, Warszawa
- Sen A. (2002), *Rozwój i wolność*, Zys i S-ka, Poznań
- Veit-Wilson J. (1994), *Dignity not poverty. A minimum income standard from the UK*, Institute for Public Policy Research, London
- Warunki życia ludności w 1998—2005 r. (z lat 1999—2007)*, GUS, Warszawa
- Watson D. (2000), *In search of the poor*, „Journal of Economic Psychology”, No. 21
- Zheng B. (2001), *Statistical inference for poverty measures with relative poverty lines*, „Journal of Econometrics”, No. 101

SUMMARY

The article presents the methodology of establishing the poverty line which can be marked with utilization of measures: absolute, relative or subjective measures. The variety of methods relating to establishing the border of poverty is the source

of methodological disputes among economists and experts. On this background the analysis of factors determining the problem of poverty in the households of Podkarpackie voivodship is presented. The classification of households on poor/nonpoor is accomplished on basis of the social minimum in 2008 according to the family members number. However, the application of the analysis of the logit regress allowed to estimate and statistical verify the logit models which enabled pointing the probability of the household falling in poverty.

РЕЗЮМЕ

В статье обсуждается методология определения черты убожества на основе абсолютных, относительных и субъективных показателей мер убожества. Множественность методов касающихся определения границы убожества является причиной методологических споров среди экономистов и экспертов. На их фоне был представлен анализ факторов детерминирующих проблему убожества в домашних хозяйствах подкарпатского воеводства. Классификация домашних хозяйств на бедные/не бедные проводилась на основе размера социального (прожиточного) минимума в 2008 г. в зависимости от количества человек в домашнем хозяйстве. Использование анализа логит регрессии позволило статистически оценить и проверить принятые в обследовании модели, которые использовались в определении вероятности попадания домашнего хозяйства в убожество.