

STATYSTYKA W PRAKTYCE

Beata KASPRZYK

Wykorzystanie modeli ekonometrycznych do badania satysfakcji z wynagrodzenia

Streszczenie. *Artykuł prezentuje wyniki badania mającego na celu podjęcie próby statystycznej analizy subiektywnego postrzegania wynagrodzeń jako głównego komponentu dochodów gospodarstwa domowego, przy wykorzystaniu określonego modelowania ekonometrycznego. Badano, czy możliwe jest wskazanie pewnych determinant społeczno-demograficznych, a jeśli tak, to w jakim stopniu mogą one wpływać na subiektywne odczucie satysfakcji z uzyskiwanego wynagrodzenia. Badanie przeprowadzono w 2015 r. na podstawie danych zebranych metodą ankietową dla próby losowej gospodarstw domowych w woj. podkarpackim, gdzie wynagrodzenia należą do najniższych w kraju. Porównano modele dwumianowe, które wykorzystano do objaśniania zmiennej jakościowej w zależności od poziomu zmiennych egzogenicznych (jakościowych i ilościowych).*

Modele regresji logitowej i probitowej pozwoliły określić prawdopodobieństwo odnoszenia sukcesu jako szansy odpowiedzi pozytywnej, czyli (w odniesieniu do przedmiotowego zakresu badania) stanu satysfakcji z uzyskiwanego wynagrodzenia.

Słowa kluczowe: teoria płacy, wynagrodzenia, gospodarstwo domowe, model logitowy.

JEL: C25, C83, D31

Pierwszą teorię dotyczącą wynagrodzeń, nazwaną teorią ceny sprawiedliwej, stworzył w średniowieczu św. Tomasz z Akwinu. Wskazał on zasadniczy wyznacznik — zasadę *instum premium* (podobnie jak w przypadku wymiany dóbr), oznaczającą, że „sprawiedliwe wynagrodzenie” musi być zapłatą za „wykonane dzieło” (Zajac, 2005, s. 113 i 114). Odtąd panował pogląd, że płaca powinna odpowiadać kosztom utrzymania pracownika i jego rodziny. Dalszy rozwój teorii

plac należy wiązać z Adamem Smithem, który w 1778 r. w dziele *Badania nad naturą i przyczynami bogactwa narodów* stworzył teorię wartości opartej na pracy. Twierdził on, iż *placa robocza musi co najmniej wystarczać na utrzymanie, w większości wypadków musi ona być nawet nieco wyższa* (Smith, 1954, s. 87 i 88). Uważał też, że należy różnicować wynagrodzenia; można go uznać za prekursora klasyfikacji czynników różnicujących wysokość plac między zawodami (Smith, 1954, s. 43). Inny klasyczny ekonomista David Ricardo twierdził, że praca to towar, placa jest ceną, a wartość pracy stanowią minimalne koszty utrzymania robotnika.

Kolejni ekonomiści reprezentowali różne nurty związane z teorią plac. Według merkantylistów państwo powinno zapewnić niskie ceny żywności, aby móc jak najmniej płacić robotnikom. Taka polityka miała gwarantować utrzymanie niskich kosztów pracy, a niska placa miała uczyć pracowitości i wiązać pracującego z miejscem pracy. Monetaryści natomiast byli przeciwni jakimkolwiek regulacjom rynku i interwencji państwowej, m.in. amerykański ekonomista Milton Friedman twierdził, że o wysokości plac powinien decydować wyłącznie mechanizm rynkowy jako mechanizm samoregulujący się (Friedman, 1993).

Z kolei Frederick W. Taylor dowodził, że jednym z ważnych czynników motywacji człowieka pracującego jest pieniądz, a ludzie na pierwszym miejscu stawiają własne korzyści i dążą do maksymalizacji dochodów. Zmusza to pracodawców do ustalania wysokości płacy w zależności od wyników pracy — im są one lepsze, tym wyższe jest wynagrodzenie. Podobne stanowisko w kwestii wynagradzania pracowników zależnie od efektywności ich pracy prezentowało wielu przedstawicieli klasycznego nurtu zarządzania, m.in. Harrington Emerson, Henry L. Gantt (twórca zadaniowo-premiewego systemu plac) i Henri Fayol. Istnieją również inne teorie plac oparte np. na teorii potrzeb (Abraham Maslow i Frederic Herzberg — za: Jacukowicz, 2003, s. 179).

Teorie w tym zakresie tworzone także z uwzględnieniem aspektów psychologicznych. Herzberg opracował dwuczynnikową teorię motywacji, dowodząc, że czynniki takie, jak: zadowolenie z pracy, osobiste osiągnięcia, uznanie przełożonych czy możliwość rozwoju motywują do osiągnięcia lepszych wyników (czynniki te nazwał motywatorami), natomiast złe stosunki ze współpracownikami i przełożonymi, niska placa oraz niesprzyjające warunki pracy powodują niezadowolenie. Według tej teorii wynagrodzenie odgrywa dwie role. Po pierwsze, jeśli pracownik uważa, że jego podstawowa pensja jest za niska, będzie niezadowolony z pracy, a brak czynników dających satysfakcję nie będzie motywował go do osiągnięcia dobrych wyników. Po drugie, każdą dodatkowo premiewaną część wynagrodzenia, postrzeganą jako zapłata za osiągnięcie lub uznanie osiągniętych wyników, należy traktować jako czynnik dający satysfakcję pracownikowi (Atkinson, Banker i Kaplan, 1995, s. 536—538).

W ubiegłym stuleciu teorie plac i systemy wynagrodzeń ewoluowały. Obecnie wynagrodzenie definiuje się jako zapłatę za pracę wykonaną przez pracownika na rzecz pracodawcy z tytułu zawartej umowy (Borkowska, 1999). Przyjmuje się, że w polityce wynagrodzeń powinna obowiązywać zasada ścisłego powiązania płacy z wynikami pracy, czyli wynagradzania proporcjonalnego do rangi stanowiska, wkładu pracy i osiąganej wydajności. Zgodnie z tym podejściem stworzono: normy pracy, akordowe systemy placowe, kategoryzację stanowisk pracy oraz taryfowy system zaszerogowań pracowników zależny od kwalifikacji. Sys-

temy płac są określone przez ścisłe reguły dotyczące przedmiotu pracy, metod jej wykonywania i należnego wynagrodzenia. Zmiany globalizacyjne, w tym rosnąca konkurencja i coraz większy udział kosztów pracy w łącznym koszcie produkcji, powodują jednak konieczność wprowadzania nowych rozwiązań w dziedzinie wynagrodzeń. Współcześnie wyzwaniem jest połączenie prawidłowego wynagradzania pracowników ze wzrostem produktywności i zysku oraz obniżką kosztów (Cieślak i Kucharczyk, 2006, s. 495—510). Nie mniej istotne jest subiektywne odczucie związane z wysokością uzyskiwanego wynagrodzenia, gdyż stanowi ono, z jednej strony, element bytu i dobrostanu jednostki i jej rodziny (gospodarstwa domowego), a z drugiej — główny motywator efektywnej pracy.

WYNAGRODZENIA W POLSCE I WOJEWÓDZTWIE PODKARPACKIM

Opierając się na danych statystycznych GUS, zauważa się, że przeciętne miesięczne wynagrodzenie brutto w latach 2005—2016 w gospodarce narodowej ulegało zmianom. Na wykresie przedstawiono wysokość wynagrodzeń w Polsce i woj. podkarpackim — nominalnych oraz skorygowanych wskaźnikiem inflacji w celu zapewnienia porównywalności (według cen stałych z 2005 r.)¹.

Analizie poddano wartości wynagrodzenia realnego, w związku z czym wyniki modelowania badanej zmiennej nie uwzględniają zmian inflacyjnych, lecz autonomiczną i bezpośrednią zmianę przeciętnej płacy. Przyjmując dla analizowanych lat zmienną czasową t jako zmienną niezależną, oszacowano liniowe funkcje trendu. Model trendu, uwzględniający wysokość ogólnopolskich wynagrodzeń, ma postać²:

$$y_{\text{wynagrodz.}(t)} = 2509,68 + 75,66 t$$
$$(44,37) \quad p = 0,00 \quad (6,03) \quad p = 0,00 \quad (1)$$

$$R^2 = 0,94 \quad Se = 72,09 \quad F(1,10) = 157,5 \quad p = 0,0000$$

gdzie:

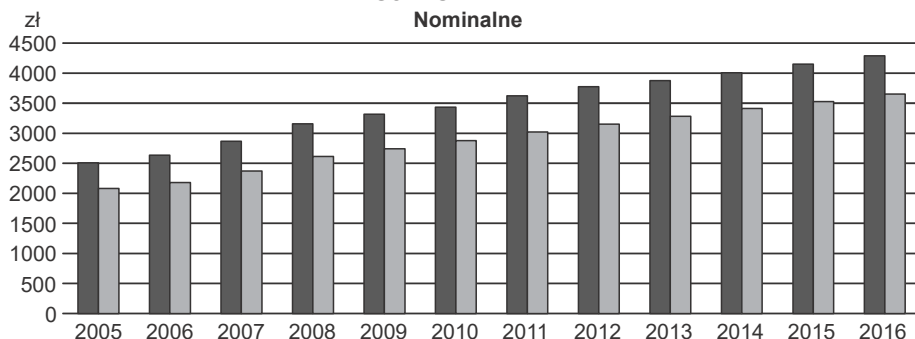
(I) — model dla danych ogólnopolskich,
 $t = 1, 2, \dots, 12$.

Z uwagi na poprawność struktury stochastycznej modelu, w tym wysoką wartość współczynnika determinacji ($R^2 = 0,94$), można przyjąć, że przeciętne realne miesięczne wynagrodzenie w Polsce wzrastało z roku na rok średnio o 75,66 zł, a średnie tempo zmian wynagrodzeń w latach 2005—2016 było dodatnie i wynosiło 3,04%.

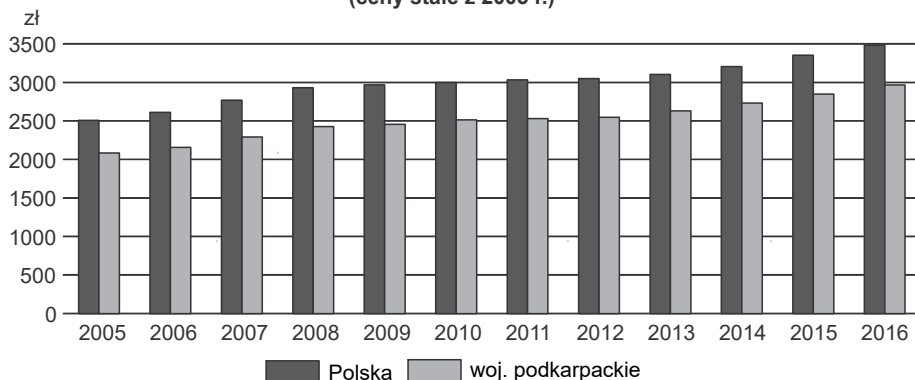
¹ Przyjęto wskaźnik cen towarów i usług konsumpcyjnych CPI dla roku 2005=100 (http://www.stat.gov.pl/gus/5840_1632_PLK_HTML.htm).

² W nawiasach pod równaniem podano błędy standardowe ocen parametrów oraz poziom istotności p dla statystyki t -Studenta. W najniższym wierszu przedstawiono wartości parametrów struktury stochastycznej modelu, kolejno: współczynnika determinacji R^2 , standardowego błędu modelu Se , statystyki Fishera F i poziomu istotności p .

PRZECIĘTNE MIESIĘCZNE WYNAGRODZENIE BRUTTO^a W POLSCE I WOJ. PODKARPACKIM



Skorygowane wskaźnikiem wzrostu cen towarów i usług konsumpcyjnych (ceny stałe z 2005 r.)



^a Bez podmiotów gospodarczych o liczbie pracujących do 9 osób.

Źródło: opracowanie własne na podstawie danych GUS.

Analiza porównawcza przeciętnych miesięcznych wynagrodzeń w woj. podkarpackim w relacji do wynagrodzeń ogólnopolskich przedstawia się mniej korzystnie. Realne wynagrodzenia w omawianych latach były w woj. podkarpackim niższe (od 2081,76 zł w 2005 r. do 2968,09 zł w 2016 r., czyli o ok. 425—517 zł mniej w odniesieniu do płac ogólnopolskich — por. wykres). Zmiany badanego zjawiska dla woj. podkarpackiego można opisać liniowym modelem trendu:

$$y_{\text{wynagrodz.}(II)} = 2052,49 + 71,17 t$$

$$(33,85) p = 0,00 \quad (4,60) p = 0,00 \quad (1)$$

$$R^2 = 0,96 \quad Se = 55,0 \quad F(1,10) = 239,38 \quad p = 0,0000$$

gdzie:

(II) — model dla danych ogólnopolskich,
 $t = 1, 2, \dots, 12$.

Przeciętne realne miesięczne wynagrodzenie brutto w woj. podkarpackim rokrocznie wzrastało o 71,17 zł, przy średnim tempie wzrostu wynoszącym 3,3%. Z analizy danych empirycznych wynika, że rokroczny wzrost przeciętnego poziomu wynagrodzeń w woj. podkarpackim był nieco wolniejszy niż w skali ogólnopolskiej.

Realny (mimo że niezbyt wysoki) wzrost przeciętnego wynagrodzenia oznacza zwiększenie możliwości dochodowych osób pracujących. W omawianym okresie trend wzrostowy był zauważalny, ale zróżnicowany w poszczególnych województwach. Najwyższe wynagrodzenie rokrocznie odnotowywano w woj. mazowieckim (ok. 25% wyższe niż przeciętne w kraju). Relatywnie wysokie były również płace w województwach: dolnośląskim, pomorskim i śląskim, natomiast najniższe — w podkarpackim i warmińsko-mazurskim; różnica między województwami o najwyższej i najniższej wysokości wynagrodzeń była znaczna (GUS, 2016b; GUS, 2016c).

ZAŁOŻENIA TEORETYCZNE MODELI DWUMIANOWYCH

Modele dwumianowe, czyli probitowe i logitowe, są wykorzystywane do opisu zmiennych jakościowych (Chow, 1995; Ostasiewicz, 1998; Gruszczyński, 2010). Można je stosować w przypadkach, gdy zmienna zależna ma charakter *stricte* dychotomiczny, czyli przyjmuje wartość 1, gdy zachodzi zdarzenie pożądane, zaś wartość 0, gdy takie zdarzenie nie zachodzi. Główne założenie polega na powiązaniu prawdopodobieństwa warunkowego zmiennej dychotomicznej tylko dla jednego z dwóch możliwych wyników z określonymi zmiennymi objaśniającymi. Modele tego typu opisują prawdopodobieństwo przyjęcia przez zmienną Y wartości 1 w odniesieniu do zmiennych egzogenicznych X_k , przy czym zmienne egzogeniczne mogą być zmiennymi jakościowymi i/lub ilościowymi.

Dla określonego przypadku i prawdopodobieństwo przyjmowania wartości 1 (pożądane zjawisko występuje) lub 0 (zjawisko nie występuje) wynosi zatem:

$$P(y_i = 1) = p_i \quad P(y_i = 0) = 1 - p_i \quad (3)$$

Zakłada się, że prawdopodobieństwo p_i jest funkcją wektora zmiennych objaśniających X oraz wektora parametrów β (jako ich iloczynu skalarnego $x_i^T \beta$) (Cramer, 2004):

$$p_i = P(y_i = 1) = F(x_i^T \beta) \quad (4)$$

Funkcja F może być funkcją typu liniowego (5), probitowego (6) i logitowego (7) (Judge, Hill, Griffiths, Lütkepohl i Lee, 1985; Stanisz, 2007), a modele prawdopodobieństwa mają postać:

- model liniowy:

$$p_i = F(x_i^T \beta) = x_i^T \beta \quad (5)$$

- probitowy³:

$$p_i = F(x_i^T \boldsymbol{\beta}) = \int_{-\infty}^{x_i^T \boldsymbol{\beta}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt \quad (6)$$

- logitowy⁴:

$$p_i = F(x_i^T \boldsymbol{\beta}) = \frac{1}{1 + \exp(-x_i^T \boldsymbol{\beta})} \quad (7)$$

Parametry (współczynniki) modelu regresji probitowej lub logistycznej dobiera się metodą największej wiarygodności lub uogólnioną metodą najmniejszych kwadratów (Gruszczyński i Podgórska, 1996, s. 139—141). Dla znanych wartości $y_i, x_{1i}, \dots, x_{ki}, i = 1, 2, \dots, n$ należy tak oszacować parametry $\beta_0, \beta_1, \dots, \beta_k$, by zapewniały maksymalną wartość logarytmu funkcji wiarygodności.

Logistyczny model regresji można rozpisać dokładniej:

$$P(Y = 1 | x_1, \dots, x_k) = \frac{\exp^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + \exp^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}} \quad (8)$$

gdzie β_1, \dots, β_k — współczynniki regresji logistycznej (Maddala, 2008; Stanisław, 2007).

Logity, obliczane jako $\ln(p_i/(1-p_i))$, stanowią wartości funkcji odwrotnej do funkcji F :

$$F^{-1}(p_i) = \ln \frac{p_i}{1-p_i} = \beta_0 + \sum_{i=1}^k \beta_i x_i \quad (9)$$

Logit wyznacza się jako logarytm ilorazu szans przyjęcia oraz nieprzyjęcia przez zmienną Y wartości równej 1. Jeśli szanse są jednakowe ($p_i = 0,5$), to logit równa się 0; dla $p_i > 0,5$ jest ujemny, a dla $p_i < 0,5$ — dodatni.

Oszacowane parametry modeli logitowego i probitowego pozwalają wyznaczyć wartości teoretyczne zmiennej objaśnianej, tj. logitów i probitów, dla każdej obserwacji i . Pozwala to obliczyć prognozę *ex post* wartości y_i dla każdej obserwacji według określonych zasad. W przypadku zbilansowanej próby (dla

³ Probity są wartościami funkcji odwrotnej do F powiększonymi o stałą 5. Wartości probitów wyznacza się zgodnie z zapisem: $Pr = F^{-1}(p_i) + 5$. Stałą 5 dodaje się w celu uniknięcia wartości ujemnych (Zeliaś, Pawełek i Wanat, 2003, s. 310).

⁴ Wykres funkcji logistycznej przypomina rozciągniętą literę S. Taki kształt umożliwia opis zjawisk, w których zmiany wartości do pewnej wartości progowej praktycznie nie zmieniają prawdopodobieństwa; po osiągnięciu wartości progowej prawdopodobieństwo gwałtownie wzrasta do 1 i utrzymuje się na tym poziomie.

zmiennej Y liczba jedynek nieznacznie różni się od liczby zer) prognozuje się następująco: $\hat{y}_i = 1$, jeśli $0,5 < \hat{p}_i \leq 1$ i $\hat{y}_i = 0$, jeśli $0 < \hat{p}_i \leq 0,5$. W przypadku gdy wartość graniczną δ ustala się jako udział jedynek w próbie: $y_i = 1$, jeśli $\delta < \hat{p}_i \leq 1$ i $y_i = 0$, jeśli $0 < \hat{p}_i \leq \delta$ (Gruszczynski, 2010, s. 73).

W większości zastosowań praktycznych zależność prawdopodobieństwa od zmiennych objaśniających jest nieliniowa. Modele probitowe i logitowe są do siebie podobne i w praktyce wykorzystuje się jeden z nich (Judge i in., 1985).

Poprawność oszacowanych modeli dwumianowych można sprawdzić za pomocą różnych miar, np. statystyki ilorazu wiarygodności, badając łączną istotność wszystkich zmiennych objaśniających⁵. Statystyka testowa LR ma dla dużych prób rozkład χ^2 (o liczbie stopni swobody równej liczbie zmiennych objaśniających); wyraża się ją wzorem (Gruszczynski, 2001; Greene, 2000):

$$LR = -2(\ln \hat{L}_R - \ln \hat{L}_{UR}) \quad (10)$$

gdzie:

\hat{L}_R — maksymalna wartość logarytmu funkcji wiarygodności dla modelu zawierającego jedynie wyraz wolny,

\hat{L}_{UR} — wartość funkcji wiarygodności dla pełnego modelu.

Miary oceniające zgodność modelu z danymi empirycznymi konstruuje się również na zasadzie odpowiedników klasycznego współczynnika determinacji dla modelu liniowego szacowanego metodą najmniejszych kwadratów. Taką miarą jest współczynnik determinacji (pseudo- R^2 McFaddena), który porównuje model pełny z modelem zredukowanym tylko dla wyrazu wolnego. Wzór oparty jest na wartości funkcji wiarygodności i wyrażony następująco⁶:

$$R_{McFadden}^2 = 1 - \frac{\ln L_{UR}}{\ln L_R} \quad (11)$$

Wielu praktyków uważa, że o jakości modelu decyduje trafność prognoz uzyskiwanych na jego podstawie, stąd do określenia zgodności modelu z danymi wykorzystuje się również mierniki dokładności prognoz (Judge i in., 1985). Trafność prognoz *ex post* dogodnie przedstawia się za pomocą tzw. tablicy trafności klasyfikacji przypadków, zliczając odpowiednie liczebności n_{ij} następująco (tabl. 1):

⁵ Hipoteza zerowa zakłada, że wszystkie parametry modelu poza wyrazem wolnym są równe 0. Hipoteza zerowa ma wtedy postać $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$, a hipoteza alternatywna — postać H_1 : co najmniej jeden parametr $\beta_j \neq 0$, $j = 1, 2, \dots, k$.

⁶ Pseudo- R^2 może służyć do porównań pomiędzy modelami w przypadku tej samej zmiennej. Wartość tego współczynnika mieści się w przedziale (0, 1), gdzie wartości bliskie 1 oznaczają doskonałe dopasowanie modelu, a wartość 0 — zupełny brak dopasowania.

TABL. 1. TRAFNOŚĆ KLASYFIKACJI PRZYPADKÓW

Przypadki empiryczne	Przypadki prognozowane		Łącznie
	$\hat{Y} = 1$	$\hat{Y} = 0$	
$Y = 1$	n_{11}	n_{10}	n_1
$Y = 0$	n_{01}	n_{00}	n_0
Razem	$n_{\cdot 1}$	$n_{\cdot 0}$	n

Ź r ó d ł o: Gruszczyński (2010), s. 73 i 74.

gdzie:

n_{11} — liczba przypadków, dla których wartość rzeczywista i przewidywana są równe 1,

n_{10} — liczba przypadków, dla których wartość rzeczywista wynosi 1, a przewidywana 0,

n_{01} — liczba przypadków, dla których wartość rzeczywista wynosi 0, a przewidywana 1,

n_{00} — liczba przypadków, dla których wartość rzeczywista i przewidywana są równe 0.

Tablica trafności pozwala na wyznaczenie miar dopasowania modelu tzw. R^2 zliczeniowego (*count R^2*) oraz ilorazu szans (*IS*). Zliczeniowy (ogólny) R^2 wyraża udział przypadków poprawnie prognozowanych w łącznej liczbie przypadków⁷ (Kufel, 2011):

$$R^2(\text{count}) = \frac{n_{00} + n_{11}}{N} \cdot 100 \quad (12)$$

Inną miarą dopasowania jest iloraz szans, liczony jako stosunek iloczynu poprawnie sklasyfikowanych przypadków do iloczynu przypadków sklasyfikowanych niepoprawnie:

$$IS = \frac{n_{11} \cdot n_{00}}{n_{01} \cdot n_{10}} \quad (13)$$

Jeśli $IS > 1$, oznacza to, że klasyfikacja na podstawie modelu jest trafniejsza od klasyfikacji czysto przypadkowej.

MATERIAŁ BADAWCZY

Badając rynek pracy, można podjąć próbę scharakteryzowania i wyznaczenia czynników wpływających na główny barometr tego rynku, jakim są wynagrodzenia. Jednym z najbardziej użytecznych narzędzi statystycznych temu służących jest model ekonometryczny, który bezpośrednio identyfikuje czynniki istotne

⁷ Im wartość tej miary jest bliższa 1, tym lepsze dopasowanie modelu do danych empirycznych. Model ten dobrze sprawdza się w prognozowaniu badanego zjawiska, gdy zliczeniowy $R^2 > 50\%$, co oznacza, że klasyfikacja na podstawie modelu jest trafniejsza od przypadkowej (Maddala, 2008).

statystycznie oraz określa kierunek i wielkość ich wpływu na zmienną objaśnianą. Indywidualne dane dotyczące zmiennych dla gospodarstw domowych w wielu przypadkach pozwalają na implementację modeli dwumianowych. W analizie ekonometrycznej możliwe jest powiązanie wyników dotyczących badanej zmiennej zależnej (Y) z innymi zmiennymi charakteryzującymi dany przypadek (X) przy użyciu prawdopodobieństwa, pod warunkiem odpowiednio dużej liczby obserwacji i zastosowania odpowiedniego modelu ekonometrycznego.

W modelach logitowych i probitowych zastosowanych w omawianym badaniu wykorzystano materiał uzyskany w badaniu ankietowym. Ankietę, o charakterze anonimowym, przeprowadzono w 2015 r. wśród ludności woj. podkarpackiego. Ostateczna liczebność próby wyniosła $n = 1010$ poprawnie udzielonych odpowiedzi⁸. Stosownie do celu badań w wywiadzie ankietowym sformułowano pytanie: *Czy wynagrodzenie (przeciętne miesięczne netto) uzyskiwane z tytułu pracy jest dla Pana/i satysfakcjonujące?* i przyjęto tylko dwa warianty odpowiedzi: „tak” lub „nie”. Równocześnie zebrano informacje dotyczące podstawowych cech ekonomiczno-społecznych i demograficznych respondentów.

Subiektywny empiryczny rozkład ocen w zakresie zadowolenia i braku zadowolenia z uzyskiwanego wynagrodzenia nie był jednakowy. Pozytywnych odpowiedzi (zadowolenie z wynagrodzenia) udzieliło zaledwie 30% respondentów ($n_{y=1} = 303$); negatywnych było $n_{y=0} = 707$. Znaczna większość badanych osób deklarowała zatem niezadowolenie z wynagrodzenia uzyskanego z tytułu pracy.

Celami badania były:

- ścisła identyfikacja istotnych statystycznie czynników wpływających na satysfakcję/brak satysfakcji z wynagrodzenia;
- zastosowanie i porównanie wyników modelowania logitowego i probitowego, pozwalających określić prawdopodobieństwo sukcesu jako szansy odpowiedzi pozytywnej zależnie od wartości zmiennych objaśniających.

Przyjęto, że zmienna Y określa respondenta (głowę gospodarstwa domowego) następująco:

- $y_i = 1$, jeżeli i -ta osoba deklaruje zadowolenie (wystąpiło zdarzenie pożądane, traktowane jako sukces);
- $y_j = 0$, jeżeli i -ta jednostka jest niezadowolona z przeciętnego miesięcznego wynagrodzenia netto uzyskanego z tytułu pracy (brak zdarzenia pożadanego, traktowane jako porażka).

Wejściowy zbiór zmiennych niezależnych (X) obejmował zmienne charakteryzujące następujące uwarunkowania funkcjonowania gospodarstwa domowego:

⁸ Badanie oparto na próbie wylosowanej z bazy rejestru TERYT (Krajowego Rejestru Urzędowego Podziału Terytorialnego Kraju). Wielkość próby (zasięg terytorialny — woj. podkarpackie) ustalono na $n = 1400$ gospodarstw domowych, z poziomem błędu $d = 2,45\%$ (przy poziomie ufności 0,95). Niewzięcie udziału w badaniu i braki odpowiedzi odnotowano w przypadku 28% ogółu wylosowanych gospodarstw domowych. Większość prowadzących badania, a także statystyków podziela pogląd Särndala i Lundströma, że współcześnie braki odpowiedzi są normalną (choć niepożądaną) cechą badań ankietowych (Särndal i Lundström 2006, s. 9), a procent braków odpowiedzi zależy od wielu czynników (Szreder, 2015, s. 4—11).

Omawiane w artykule badanie ankietowe przeprowadzono techniką wywiadu bezpośredniego, z wykorzystaniem specjalnie przygotowanego kwestionariusza, po uzyskaniu zgody respondenta na wypełnienie kwestionariusza. Respondentem była osoba wnosząca największy wkład w całkowity miesięczny dochód gospodarstwa domowego, tj. głowa gospodarstwa domowego (nazywana też osobą odniesienia).

przynależność do ściśle określonej kwintylowej grupy dochodowej, liczba osób w gospodarstwie domowym, w tym liczba osób osiągających dochody, liczba dzieci na utrzymaniu oraz struktura wydatków. Ponadto analizie poddano typowe dla tego typu badań cechy gospodarstwa domowego: *pleć, wiek, wykształcenie, miejsce zamieszkania, formę zatrudnienia głowy gospodarstwa* oraz *skład rodzinny* (GUS, 2016a; Panek i Czapiński, 2015). Wprowadzono także nowe potencjalne zmienne objaśniające o charakterze jakościowym: *sposób gospodarowania w gospodarstwie domowym, subiektywna aktualna ocena materialnego poziomu życia, posiadanie oszczędności i subiektywna ocena ogólnej jakości życia*. Warianty odpowiedzi w tej grupie zmiennych objaśniających zestawiono poniżej (zestawienie 1).

ZESTAWIENIE 1. ZMIENNE JAKOŚCIOWE — WARIANTY ODPOWIEDZI

Pytanie w kwestionariuszu ankietowym	Warianty odpowiedzi
Sposób gospodarowania pieniędzmi w Pana/i gospodarstwie domowym jest	bardzo skromny skromny średni — w miarę wystarczający dobry — wystarcza na wiele potrzeb bardzo dobry — pozwala na pewien luksus
Jak Pan/i ocenia aktualny materialny poziom życia członków gospodarstwa domowego	zły (raczej zły) przeciętny dobry bardzo dobry
Czy gospodarstwo domowe posiada oszczędności	tak nie
Czy zadowolony/a jest Pan/i z aktualnej ogólnej jakości życia gospodarstwa domowego (biorąc pod uwagę sytuację materialną, pracę, zdrowie, mieszkanie, styl życia itd.)	tak nie

U w a g a. Respondent wskazywał tylko jeden wariant odpowiedzi.

Ź r ó d ł o: opracowanie własne.

W badaniach dotyczących płac z pewnością należałoby uwzględnić jeszcze inne cechy, które z jednej strony określają uzyskiwanie wynagrodzenia przez pracowników, a z drugiej — ich ocenę. Wynagrodzenie powinno odpowiadać w szczególności rodzajowi pracy i kwalifikacjom wymaganym przy jej wykonywaniu oraz uwzględniać ilość i jakość świadczonych prac. Ponadto indywidualna wysokość wynagrodzenia zależy od wielu innych czynników bezpośrednio mierzalnych, takich jak lata i staż pracy, a także od tych trudniej mierzalnych, jak np. nakłady finansowe na zdobycie danego zawodu, osobiste zaangażowanie czy podejście do pracy. Wysokość zarobków jest zindywidualizowana, dlatego wynagrodzenie pracowników zatrudnionych na takim samym stanowisku może się znacznie różnić. Wpływa na to wiele czynników. Z pewnością można zaobserwować zależności pomiędzy wynagrodzeniem a wielkością firmy, wykształceniem, branżą, regionem oraz stanowiskiem, doświadczeniem i charakterem pracy. Pewne czynniki, np. posiadane kompetencje, cechy charakteru (m.in. zaradność, pracowitość czy umiejętność osiągania celów), można określić jako personalne. Oprócz wymiaru ekonomicznego w analizie wynagrodzeń należałoby wziąć pod uwagę także elementy sfery socjologicznej czy psychologicznej.

Należy podkreślić, że w omawianym badaniu uwzględniono i analizie poddano jedynie wybrane identyfikowalne cechy, które można było w miarę poprawnie kwantyfikować, oraz informacje zawarte w empirycznym materiale badawczym. Wybór do modeli zmiennych objaśniających z przyjętego wstępnie zbioru poprzedzono sprawdzeniem, czy występują istotne zależności między potencjalnymi zmiennymi niezależnymi a zmienną zależną. Weryfikację hipotez o niezależności stochastycznej dla badanych cech dokonano za pomocą testu niezależności chi-kwadrat Pearsona (tabl. 2).

TABL. 2. WYNIKI TESTU NIEZALEŻNOŚCI CHI-KWADRAT ZMIENNYCH OBJAŚNIAJĄCYCH WZGLĘDEM ZMIENNEJ Y

Zmienne objaśniające	Wartość testu χ^2 Pearsona	<i>p-value</i>
Grupa dochodowa	124,7	0,000
Ocena poziomu życia	296,7	0,000
Oszczędności	159,6	0,000
Jakość życia	316,2	0,000
Liczba osób osiągających dochody	21,6	0,102
Liczba osób ogółem w gospodarstwie	20,2	0,002
Wykształcenie	35,3	0,000
Wiek	11,4	0,022
Miejsce zamieszkania	21,2	0,101
Forma zatrudnienia	20,1	0,000
Skład rodzinny gospodarstwa	21,4	0,103
Płeć	1,1	0,301
Liczba dzieci	5,9	0,441
Wydatki (żywność)	7,7	0,151
Wydatki (mieszkanie)	1,5	0,483
Wydatki (inne)	1,6	0,448

U w a g a. *p-value* oznacza najmniejszy poziom istotności, przy którym można odrzucić hipotezę zerową.

Ź r ó d ł o: opracowanie własne z wykorzystaniem pakietu *Statistica*.

Wyniki testu wskazują, że w większości przypadków należy odrzucić hipotezę o niezależności badanych cech, co oznacza, że większość przyjętych zmiennych objaśniających wpływa statystycznie istotnie na zmienną zależną *Y* ($p < 0,05$). Brak powiązania dotyczył następujących cech: *pleć*, *liczba dzieci* i kategorie wydatków⁹. Dodatkowo, w celu sprawdzenia i uniknięcia współliniowości danych dla zmiennych *X* wyznaczono macierz współzależności; zastosowano również test niezależności chi-kwadrat Pearsona (tabl. 3).

⁹ Zmienna o charakterze porządkowym dla wydatków: żywność, utrzymanie mieszkania (domu), pozostałe (inne) wydatki, od najbardziej do najmniej obciążających budżet gospodarstwa domowego.

**TABL. 3. WYNIKI TESTU NIEZALEŻNOŚCI CHI-KWADRAT
WYBRANYCH ZMIENNYCH OBJAŚNIAJĄCYCH**

Zmienne objaśniające	Aktualny poziom życia	Sposób gospodarowania	Oszczędności	Jakość życia
<i>Wykształcenie</i>	90,63 ($p = 0,000$) 0,17	97,74 ($p = 0,000$) 0,18	44,30 ($p = 0,000$) 0,21	33,26 ($p = 0,000$) 0,18
<i>Płeć</i>	14,48 ($p = 0,006$) 0,12	13,01 ($p = 0,011$) 0,11	7,63 ($p = 0,006$) 0,09	4,34 ($p = 0,000$) 0,07
<i>Wiek</i>	47,38 ($p = 0,000$) 0,11	55,67 ($p = 0,000$) 0,12	12,61 ($p = 0,013$) 0,12	12,09 ($p = 0,017$) 0,11
<i>Miejsce zamieszkania</i>	21,98 ($p = 0,038$) 0,08	26,37 ($p = 0,009$) 0,09	10,44 ($p = 0,016$) 0,10	5,82 ($p = 0,121$) 0,08
<i>Forma zatrudnienia</i>	88,77 ($p = 0,000$) 0,15	82,09 ($p = 0,000$) 0,14	18,29 ($p = 0,001$) 0,14	37,76 ($p = 0,000$) 0,19
<i>Skład rodzinny gospodarstwa</i>	74,63 ($p = 0,000$) 0,16	63,94 ($p = 0,000$) 0,15	9,94 ($p = 0,019$) 0,09	15,94 ($p = 0,001$) 0,13
<i>Grupa kwintylowa</i>	191,09 ($p = 0,000$) 0,22	202,87 ($p = 0,000$) 0,22	102,87 ($p = 0,000$) 0,32	104,26 ($p = 0,000$) 0,32
<i>Liczba osób ogółem</i>	96,03 ($p = 0,000$) 0,15	87,33 ($p = 0,000$) 0,15	12,6 ($p = 0,179$) 0,11	19,29 ($p = 0,023$) 0,14
<i>Liczba osób osiągających dochody w gospodarstwie domowym</i>	58,78 ($p = 0,000$) 0,13	47,86 ($p = 0,003$) 0,11	19,21 ($p = 0,004$) 0,14	6,94 ($p = 0,326$) 0,08

U w a g a. Wartości statystyki oraz poziom p podano w pierwszym wierszu. Pod wartościami zmiennych (w drugim wierszu) przedstawiono wynik pomiaru natężenia siły zależności stochastycznej za pomocą współczynnika V Cramera.

Ź r ó d ł o: jak przy tabl. 2.

Wyniki wskazują, że wybrane zmienne objaśniające były w większości istotnie zależne ($p < 0,05$). Odrzucono hipotezę zerową o stochastycznej niezależności badanych zmiennych, co oznacza, że istnieją przesłanki do uznania badanych zależności za istotne. Brak zależności istotnych statystycznie dotyczył zmiennych: *liczba osób osiągających dochody a jakość życia* ($p = 0,326$), *miejsce zamieszkania a jakość życia* ($p = 0,121$) oraz *liczba osób ogółem a oszczędności* ($p = 0,179$). W przypadku prawie wszystkich rozpatrywanych cech odrzucono hipotezę o niezależności stochastycznej, dlatego zasadna była ocena siły zależności za pomocą współczynnika V Cramera (Steczkowski i Zeliaś, 1997; Domański, 1990). Wartości współczynnika V Cramera w zakresie 0,07—0,32 pozwalają przypuszczać, że między zmiennymi występuje związek, jednakże jego siła jest niewielka. Wniosek ten potwierdzają wartości drugiego współczynnika — φ Yule'a, nieco wyższe od wartości współczynnika Cramera¹⁰. Założenia wstępne, zweryfikowane pod kątem występowania zależności pomiędzy zmiennymi, pozwoliły na przejście do kolejnego etapu badania, polegającego na wyznaczaniu ocen parametrów przyjętych modeli.

¹⁰ Oba współczynniki przybierają wartości z przedziału (0, 1). Im wartość bliższa 1, tym silniejsza jest zaobserwowana zależność rozpatrywanych cech (Steczkowski i Zeliaś, 1997, s. 173—177).

WYNIKI ESTYMACJI MODELI DWUMIANOWYCH

Za determinanty subiektywnych ocen wynagrodzeń przyjęto jednostkowe uwarunkowania respondentów, wyrażone poprzez określone zmienne ilościowe i jakościowe. W modelowaniu użyto algorytmu iteracyjnego quasi-Newtona, a modelowania dokonywano aż do uzyskania poprawnych modeli ekonometrycznych, czyli do osiągnięcia kryterium zbieżności dla określonego zestawu zmiennych objaśniających. Do pierwszej grupy modeli logitowych i probitowych (MODELE 1) wprowadzono wstępnie przyjęty zbiór następujących zmiennych niezależnych — zestawienie 2:

ZESTAWIENIE 2. MODELE 1 — ZMIENNE OBJAŚNIAJĄCE

Zmienne objaśniające	Kategorie
Przynależność do dochodowej grupy kwintylowej ^a	$X_{1,1}$ — I (R) ^b $X_{1,2}$ — II $X_{1,3}$ — III $X_{1,4}$ — IV $X_{1,5}$ — V
Liczba osób ogółem w gospodarstwie domowym	X_2
Subiektywna ocena aktualnego materialnego poziomu życia	$X_{3,1}$ — zła, raczej zła (R) ^b $X_{3,2}$ — przeciętna $X_{3,3}$ — dobra $X_{3,4}$ — bardzo dobra
Oszczędności (posiada/nie posiada)	X_4 (posiada = 1)
Subiektywna ocena zadowolenia z ogólnej jakości życia (pozytywna/negatywna)	X_5 (pozytywna = 1)
Liczba osób osiągających dochody w gospodarstwie domowym	X_6
Sposób gospodarowania pieniędzmi w gospodarstwie domowym	$X_{7,1}$ — bardzo skromny, skromny (R) ^b $X_{7,2}$ — średni $X_{7,3}$ — dobry $X_{7,4}$ — bardzo dobry

^a Wyróżniono pięć kwintylowych grup dochodowych, obejmujących po 20% ogółu gospodarstw domowych. Klasyfikacji danego gospodarstwa dokonano na podstawie przeciętnego miesięcznego dochodu *per capita*: I grupa — gospodarstwa najuboższe, V grupa — gospodarstwa najbogatsze. ^b (R) — kategoria referencyjna.

Ź r ó d ł o: jak przy zestawieniu 1.

Na bazie zmiennej X_1 (przynależność do dochodowej grupy kwintylowej) skonstruowano pięć zmiennych dwumianowych (wartość 1, jeśli respondent reprezentował dany stan/kategorię, i wartość 0 w przeciwnym razie). W przypadku zmiennej X_3 (subiektywna ocena aktualnego materialnego poziomu życia) oraz X_7 (sposób gospodarowania pieniędzmi w gospodarstwie domowym) przyjęto cztery zmienne dwumianowe. Należy podkreślić wyraźną przewagę zmiennych o charakterze subiektywnym (X_3 , X_4 , X_5 i X_7 — respondent wskazuje własną ocenę w odniesieniu do mierzonych zjawisk). Wprowadzono także zmienne silnie zależne od zmiennej dychotomicznej, związane z dochodami (X_1 , X_6 i X_7) oraz zmienną ilościową o charakterze demograficznym (X_2).

W trakcie obliczeń metodą *a posteriori* z modeli eliminowano zmienne nieistotne statystycznie (przyjęto poziom istotności $\alpha = 0,05$). Zmienną, dla której

uzyskano wartość $p > 0,05$, eliminowano i dokonywano kolejnej estymacji, aż do uzyskania modelu, w którym wartości p były mniejsze lub równe α . W ten sposób uzyskano modele logitowy i probitowy zawierające $k = 10$ zmiennych niezależnych. W tabl. 4 przedstawiono wartości oszacowanych ocen parametrów modeli, błędy standardowe ocen oraz istotność statystyczną (weryfikowaną na podstawie statystyki t -Studenta).

TABL. 4. MODELE I: WYNIKI ESTYMACJI — OCENY PARAMETRÓW

Wyszczególnienie	Ocena parametru β_i		Błąd standardowy $S(\beta_i)$		Statystyka t -Studenta		p -value	
	modele							
	logitowy	probitowy	logitowy	probitowy	logitowy	probitowy	logitowy	probitowy
Stała β_0	-6,5712	-3,4840	1,0817	0,4840	-6,0749	-7,1994	0,0000	0,0000
$X_{1,3}$	0,7329	0,3754	0,3430	0,1917	2,1369	1,9584	0,0328	0,0500
$X_{1,4}$	1,3017	0,6907	0,3499	0,1963	3,7204	3,5183	0,0002	0,0001
$X_{1,5}$	1,9891	1,0924	0,3878	0,2171	5,1293	5,0320	0,0000	0,0000
X_2	-0,3990	-0,2231	0,1265	0,0723	-3,1529	-3,0845	0,0017	0,0021
$X_{3,2}$	2,1707	1,0285	1,0270	0,4432	2,1137	2,3205	0,0348	0,0205
$X_{3,3}$	3,1218	1,5881	1,0321	0,4468	3,0248	3,5544	0,0260	0,0004
$X_{3,4}$	3,5374	1,8469	1,0808	0,4830	3,2730	3,8241	0,0011	0,0001
X_4	0,5332	0,3078	0,2154	0,1223	2,4752	2,5159	0,0135	0,0120
X_5	1,8551	1,0574	0,2270	0,1263	8,1714	8,3716	0,0000	0,0000
X_6	0,4119	0,2264	0,0916	0,0519	0,4461	4,3661	0,0000	0,0000
Statystyka LR	$\chi^2 = 473,669$; $p = 0,0000$							

Źródło: jak przy tabl. 2.

Poprawność dopasowania modeli logitowego i probitowego wyrażona wartością testu ilorazu wiarygodności równa 473,7 ($p = 0,0000$) oznacza, że należy odrzucić hipotezę zerową o łącznej nieistotności ocen parametrów tych modeli; wyznaczone modele różnią się także istotnie od modelu tylko ze stałą β_0 . Oceniając wartości statystyki t -Studenta dla poszczególnych zmiennych, można wywnioskować, że zmienne, dla których oszacowano parametry, wywierają także osobno istotny statystycznie wpływ na zmienną zależną (przy $p < 0,05$). Wyniki modelowania wskazują, że w obu porównywanych modelach uzyskano oceny parametrów o tym samym kierunku (w modelu logitowym wartości parametrów w przypadku każdej zmiennej X są zawsze większe niż w modelu probitowym). Istotne jest, że w obu modelach uzyskano taką samą grupę statystycznie istotnych zmiennych objaśniających, co potwierdza zależności i związki pomiędzy badanymi zjawiskami.

Można uznać, że modele probitowy i logitowy są do siebie podobne. W rozwiązaniach praktycznych — głównie w celach interpretacyjnych — wykorzystuje się najczęściej jeden z nich (Judge i in., 1985). W tabl. 5 podano wyniki statystyki Walda oraz wartości jednostkowego ilorazu szans JIS wraz z przedziałami ufności JIS [-95%; +95%] dla modelu logitowego¹¹.

¹¹ Efekty krańcowe obliczone dla średnich wartości zmiennych objaśniających w modelach logitowym i probitowym okazały się prawie identyczne, stąd dalsze wnioskowanie interpretacyjne przedstawiono na podstawie jednego modelu — logitowego.

TABL. 5. MODEL I (logitowy): WYNIKI ESTYMACJI — ILORAZ SZANS

Wyszczególnienie	Statystyka Walda ^a	p-value	JIS	JIS (-95%)	JIS (+95%)
Stała β_0	36,9044	0,0000	0,0014	0,0002	0,0117
X_{1_3}	4,5663	0,0326	2,0812	1,0617	4,0796
X_{1_4}	13,8417	0,0002	3,6754	1,8498	7,3025
X_{1_5}	26,3093	0,0000	7,3093	3,4149	15,6450
X_2	9,9406	0,0016	0,6710	0,5235	0,8601
X_{3_2}	4,4677	0,0345	8,7641	1,1682	65,7529
X_{3_3}	9,1492	0,0025	22,6883	2,9937	171,9440
X_{3_4}	10,7126	0,0011	34,4773	4,1228	286,6474
X_4	6,1265	0,0133	1,7044	1,1168	2,6012
X_5	66,7726	0,0000	6,3921	4,0942	9,9797
X_6	19,7681	0,0000	1,5096	1,2587	1,8106

a Test Walda dany wzorem $\left(\frac{\beta_i}{s(\beta_i)}\right)^2$ weryfikuje hipotezę zerową, że i -ty parametr modelu jest równy 0. Statystyka testu Walda ma rozkład chi-kwadrat z liczbą stopni swobody równą 1.

Źródło: jak przy tabl. 2.

Wartości testu chi-kwadrat Walda potwierdzają, że zmienne objaśniające, dla których oszacowano parametry, wywierają istotny statystycznie wpływ na zmienną zależną (przy $p < 0,05$)¹². Jak wynika z modelowania, istotne statystycznie okazały się wszystkie zmienne z wyjątkiem X_7 (*sposób gospodarowania pieniędzmi w gospodarstwie domowym*). Wyniki te jednoznacznie wskazują, że zmienne uwzględnione w modelach mają statystycznie istotny wpływ na subiektywną ocenę zadowolenia/braku zadowolenia z uzyskiwanego wynagrodzenia. Wyniki estymacji potwierdzają podobieństwo modelu probitowego i logitowego. Na podstawie końcowych wyników modelowania można wysunąć wnioski na temat czynników wywierających istotny wpływ na ocenę wynagrodzenia oraz siłę tych oddziaływań. Znak wartości oszacowanych parametrów określa kierunek zależności między zmiennymi niezależnymi a zmienną zależną Y . Dodatkowo wartości uzyskano dla wszystkich zmiennych niezależnych z wyjątkiem X_2 (*liczba osób ogółem w gospodarstwie domowym*), co oznacza, że zmienne te powodują zwiększenie prawdopodobieństwa zaklasyfikowania danego gospodarstwa domowego do grupy zadowolonych z wynagrodzenia.

Bezpośrednia interpretacja wartości oszacowanych parametrów modelu dwumianowego nie jest możliwa, dlatego należy brać pod uwagę efekty krańcowe, tj. jednostkowe ilorazy szans JIS ¹³. Wzrost lub spadek ilorazu szans dla tej samej zmiennej niezależnej odnosi się do kategorii referencyjnej (R), jeśli taką przyjęto. Wyniki modelowania wskazują, że w przypadku wszystkich zmiennych (z wyjątkiem X_2) wartości JIS były większe od 1 (tabl. 5). Pierwszą istotną

¹² Wartości testu Walda oraz statystyki t -Studenta dla ocen parametrów służą do testowania istotności każdej zmiennej niezależnej oddzielnie.

¹³ Jednostkowy iloraz szans związany z daną zmienną wskazuje, ile razy rośnie lub maleje prawdopodobieństwo zaistnienia danego zdarzenia, jeżeli nastąpi jednostkowa zmiana zmiennej niezależnej. W przypadku zwiększenia się jednej ze zmiennych objaśniających o jednostkę — zakładając, że pozostałe zmienne pozostają na stałym poziomie (*ceteris paribus*) — iloraz szans zmienia się $\exp(\beta_i)$ razy; $\exp(\beta_i) > 1$ oznacza wzrost, a $\exp(\beta_i) < 1$ — spadek ilorazu szans (Śliwicki i Ręklawski, 2012, s. 25).

zmienną objaśniającą, a zarazem czynnikiem zidentyfikowanym przez model, była zmienna X_1 (*przynależność do dochodowej grupy kwintylowej*)¹⁴. Przykładowo, można wskazać, że osoby¹⁵ z III grupy kwintylowej miały o 108% większą szansę pozytywnej oceny wynagrodzenia niż osoby z grupy referencyjnej, czyli grupy I (20% gospodarstw domowych osiągających najniższe dochody). Szansa ta wyraźnie rosła wraz z przynależnością do kolejnych (wyższych) grup kwintylowych. Prawdopodobieństwo pozytywnej odpowiedzi dotyczącej wynagrodzenia wzrosło o 268% w przypadku IV grupy kwintylowej, a najwyższe było w V grupie kwintylowej (ponad 6-krotnie większe w stosunku do gospodarstw najbiedniejszych).

Kolejną statystycznie istotną zmienną zidentyfikowaną przez model była zmienna X_3 (*subiektywna ocena aktualnego materialnego poziomu życia*). Oszacowany dla $X_{3,2}$ jednostkowy iloraz szans oznacza ponad 7-krotnie większą szansę sukcesu (osiągnięcia satysfakcji z wynagrodzenia) osób, które oceniają swój poziom życia jako przeciętny, w porównaniu z osobami z grupy referencyjnej (oceniających swój poziom życia jako zły lub raczej zły). Także w tym przypadku prawdopodobieństwo pozytywnej oceny wynagrodzeń znacząco zwiększało się wraz z podnoszeniem oceny poziomu życia.

Kolejny statystycznie istotny predyktor stanowiła zmienna X_5 (*subiektywna ocena zadowolenia z ogólnej jakości życia*). W przypadku osób zadowolonych z ogólnej jakości życia szanse uzyskania $Y = 1$ były ponad 5-krotnie większe niż osób niezadowolonych z jakości życia.

Odnosnie do zmiennej X_4 (*oszczędności*) istnieje duża szansa (70%), że ocena uzyskiwanego wynagrodzenia będzie pozytywna w przypadku gospodarstw domowych posiadających oszczędności.

Biorąc pod uwagę uwarunkowania demograficzne, można wysnuć wniosek, że zwiększenie prawdopodobieństwa pozytywnej oceny idzie w parze ze wzrostem liczby osób osiągających dochody w gospodarstwie domowym. Zwiększenie się liczby osób wnoszących wkład w całkowity dochód gospodarstwa domowego powoduje wzrost szansy pozytywnego postrzegania wysokości wynagrodzenia o 51%. Sytuacja wygląda przeciwnie w przypadku wzrostu liczby osób w gospodarstwie domowym — szansa sukcesu spada o 33% wraz z kolejną osobą wchodzącą w skład gospodarstwa.

Na podstawie oszacowanych parametrów modelu (tu: logitowego) można wyznaczyć oczekiwane wartości prawdopodobieństwa zadowolenia z uzyskiwanego wynagrodzenia dla gospodarstwa domowego charakteryzującego się określonymi cechami. Wartości prawdopodobieństwa, z jakim wystąpi wariant zmiennej $Y = 1$ np. dla głowy gospodarstwa o cechach: *posiadanie oszczędności, wyłącznie jedna osoba osiąga dochody, dobra ocena aktualnego poziomu życia oraz pozytywna ocena ogólnej jakości życia*, w zależności od grupy kwintylowej oraz zmiany liczby osób ogółem w gospodarstwie, zestawiono w tabl. 6.

Oczekiwane prawdopodobieństwo wskazania stanu zadowolenia zauważalnie zwiększa się wraz ze wzrostem przynależności do wyższej grupy kwintylowej, natomiast wzrost liczby osób w danej grupie kwintylowej powoduje spadek

¹⁴ Zmienna dotycząca II grupy kwintylowej okazała się nieistotna statystycznie.

¹⁵ Stosowane terminy: *osoba, gospodarstwo domowe, jednostka i respondent* w warstwie interpretacyjnej zawsze odnoszą się do głowy gospodarstwa domowego.

prawdopodobieństwa, przy czym różnice te są najmniejsze w przypadku najbogatszych gospodarstw.

TABL. 6. OCZEKIWANE WARTOŚCI PRAWDOPODOBIEŃSTWA DLA $Y = 1$ WEDŁUG GRUP KWINTYLOWYCH^a I LICZBY OSÓB OGÓŁEM

Liczba osób w gospodarstwie domowym	Grupy kwintylowe			
	I	III	IV	V
1	0,24	0,42	0,56	0,97
2	0,19	0,33	0,46	0,96
3	0,14	0,25	0,37	0,94
4	0,10	0,18	0,28	0,91
5	0,07	0,13	0,21	0,87

a Zmienna dotycząca II grupy kwintylowej okazała się statystycznie nieistotna.

Źródło: jak przy zestawieniu 1.

Warianty i cechy badanych obiektów można sprawdzać w różnych zestawieniach symulacyjnych, prognozując w ten sposób sukces w przypadku każdego gospodarstwa domowego. Jeśli np. rozważymy przypadek gospodarstw nieposiadających oszczędności (pozostałe zmienne bez zmian), wartości prawdopodobieństwa będą maleć. I tak w 5-osobowym gospodarstwie domowym szansa uzyskania wariantu zmiennej $Y = 1$ wynosi odpowiednio: 0,04, 0,08, 0,13 i 0,80 w kolejnych grupach kwintylowych, począwszy od I. Jeżeli liczba osób osiągających jakiegokolwiek dochody np. wzrośnie do dwóch, to prawdopodobieństwo zadowolenia zwiększy się od 0,06 w I grupie kwintylowej do: 0,12 w II, 0,19 w III i 0,86 w V grupie kwintylowej.

W kolejnym modelowaniu (MODELE II) wprowadzono przede wszystkim klasyczne zmienne, wprost reprezentujące uwarunkowania społeczno-ekonomiczne gospodarstw domowych. Wejściowy zbiór obejmował takie zmienne objaśniające, jak: *grupa kwintylowa*, *płeć*, *wiek* oraz *wykształcenie głowy gospodarstwa*, *miejsce zamieszkania*, *forma zatrudnienia*, *skład rodzinny gospodarstwa domowego* oraz wybrane zmienne demograficzne. Wprowadzono nowe zmienne dwumianowe, odpowiednio przyporządkowane zmiennym niezależnym (zestawienie 3).

ZESTAWIENIE 3. MODELE II — ZMIENNE OBJAŚNIAJĄCE

Zmienne objaśniające	Kategorie
<i>Wykształcenie</i>	$X_{7,1}$ — podstawowe, zasadnicze (R) ^a $X_{7,2}$ — średnie $X_{7,3}$ — wyższe
<i>Wiek</i>	$X_{8,1}$ — do 40 lat (R) ^a $X_{8,2}$ — 41—50 $X_{8,3}$ — 51—60 $X_{8,4}$ — 61 i więcej lat
<i>Miejsce zamieszkania (miasto/wieś)</i>	X_9 (miasto = 1)
<i>Skład rodzinny gospodarstwa domowego</i>	$X_{10,1}$ — gospodarstwo jednorodzinne $X_{10,2}$ — gospodarstwo wielorodzinne $X_{10,3}$ — rodzina niepełna $X_{10,4}$ — gospodarstwo jednoosobowe (R) ^a

a (R) — kategoria referencyjna.

ZESTAWIENIE 3. MODELE II — ZMIENNE OBJAŚNIAJĄCE (dok.)

Zmienne objaśniające	Kategorie
Płeć (mężczyzna/kobieta)	X_{11} (mężczyzna = 1)
Forma zatrudnienia: zatrudnienie u pracodawcy/inna forma pracy (na własny rachunek, praca w gospodarstwie rolnym, emerytura/renta, źródła niezarobkowe)	$X_{12,1}$ (praca u pracodawcy = 1)

a (R) – kategoria referencyjna.

Ź r ó d ł o: jak przy zestawieniu 1.

Końcowe oceny parametrów modeli logitowego i probitowego, istotność statystyczną oraz jednostkowe efekty krańcowe oszacowane dla modelu logitowego przedstawiono w tabl. 7 i 8.

TABL. 7. MODELE II: WYNIKI ESTYMACJI — OCENY PARAMETRÓW

Wyszczególnienie	Ocena parametru β_i		Błąd standardowy $S(\beta_i)$		Statystyka t -Studenta		p -value	
	modele							
	logitowy	probitowy	logitowy	probitowy	logitowy	probitowy	logitowy	probitowy
Stała β_0	-3,7359	-2,2083	0,4651	0,2642	-8,0321	-8,3576	0,0000	0,0000
$X_{1,2}$	0,5541	0,3069	0,2884	0,1579	1,9214	1,9429	0,0500	0,0523
$X_{1,3}$	0,9773	0,5458	0,2766	0,1538	3,5332	3,5490	0,0004	0,0004
$X_{1,4}$	1,6292	0,9490	0,2717	0,1527	5,9957	6,2168	0,0020	0,0000
$X_{1,5}$	2,6460	1,5664	0,2846	0,1594	9,2971	9,8266	0,0000	0,0000
$X_{7,3}$	0,5323	0,3175	0,2173	0,1288	2,4489	2,4646	0,0145	0,0139
$X_{8,3}$	0,5744	0,3394	0,1871	0,1101	3,0707	3,0817	0,0022	0,0021
$X_{10,1}$	1,1351	0,6694	0,3219	0,1855	3,5260	3,6076	0,0004	0,0003
$X_{10,2}$	1,4798	0,8685	0,3824	0,2203	3,8702	3,9414	0,0001	0,0001
$X_{10,3}$	0,9642	0,5677	0,4440	0,2557	2,1716	2,2201	0,0301	0,0266
Statystyka LR	$\chi^2 = 174,615 \quad p = 0,0000$							

Ź r ó d ł o: jak przy tabl. 2.

W efekcie modelowania (MODELE II) uzyskano statystykę testu LR $\chi^2=174,615$ ($p = 0,0000$), co uzasadnia odrzucenie hipotezy zerowej o łącznej nieistotności parametrów obu modeli. Zmienne niezależne modeli logitowego i probitowego wywierają osobno istotny statystycznie wpływ na zmienną zależną przy poziomie istotności $p \leq 0,05$. W obu modelach statystycznie istotne okazały się te same zmienne: *wykształcenie*, *wiek głowy gospodarstwa domowego*, *przynależność do grupy dochodowej*¹⁶ i *skład rodzinny gospodarstwa domowego* (przy czym wartości ocen parametrów były inne). W przypadku zmiennych: *miejsce zamieszkania*, *forma zatrudnienia* oraz *płeć* poziom istotności był wyższy od założonego $\alpha = 0,05$, co poskutkowało wycofaniem ich z modeli. Użytkowano zatem podobieństwo w warstwie analitycznej i interpretacyjnej modeli probitowego i logitowego.

¹⁶ Zmienna X_1 występuje również w tym modelowaniu, ze względu na wysoki stopień zależności ze zmienną zależną Y , a także na lepszą jakość dopasowania modeli uzyskaną z włączeniem tej zmiennej.

TABL. 8. MODEL II (logitowy): WYNIKI ESTYMACJI — ILORAZ SZANS

Wyszczególnienie	Statystyka Walda	<i>p-value</i>	<i>JIS</i>	<i>JIS</i> (-95%)	<i>JIS</i> (+95%)
Stała β_0	64,5151	0,0000	0,0239	0,0096	0,0594
$X_{1,2}$	3,6918	0,0500	1,7403	0,9883	3,0648
$X_{1,3}$	12,4832	0,0004	2,6574	1,5442	4,5730
$X_{1,4}$	35,9489	0,0000	5,0998	2,9921	8,6922
$X_{1,5}$	86,4363	0,0000	14,0970	8,0646	24,6420
$X_{7,3}$	5,9972	0,0143	1,7028	1,1115	2,6085
$X_{8,3}$	9,4293	0,0021	1,7760	1,2304	2,5637
$X_{10,1}$	12,4325	0,0004	3,1115	1,6543	5,8524
$X_{10,2}$	14,9788	0,0001	4,3920	2,0740	9,3008
$X_{10,3}$	4,7160	0,0299	2,6227	1,0974	6,2680

Źródło: jak przy tabl. 2.

Na prawdopodobieństwo sukcesu najmocniej wpływa zmienna X_1 (*przynależność do dochodowej grupy kwintylowej*) — efekty krańcowe okazały się ponownie największe właśnie dla tej zmiennej¹⁷. Oszacowane prawdopodobieństwo pozytywnego postrzegania wysokości wynagrodzenia wzrasta o 74% w przypadku gospodarstw domowych zakwalifikowanych do II grupy, o 166% — do III, o 410% — do IV i aż 13-krotnie w V grupie kwintylowej w porównaniu z gospodarstwami z grupy referencyjnej, tj. osiągniętymi najniższe dochody.

Odnosnie do zmiennej X_7 (*wykształcenie*) można przyjąć, że w przypadku osób z wykształceniem wyższym prawdopodobieństwo sukcesu jest wyższe o 70% niż osób z wykształceniem podstawowym i zasadniczym. Istotną statystycznie zmienną objaśniającą w obu modelach dotyczy osób z grupy wiekowej 51—60 lat. Wyniki modelowania wskazują, że w ich przypadku szansa pozytywnej oceny wynagrodzenia jest o 78% większa niż w grupie referencyjnej (osoby w wieku do 40 lat). Szansa wyrażenia pozytywnej oceny zmienia się istotnie w zależności od składu rodzinnego gospodarstwa domowego. W relacji do gospodarstwa jednoosobowego prawdopodobieństwo sukcesu jest większe w przypadku gospodarstw jednorodzinnych, gospodarstw wielorodzinnych, a nawet rodzin niepełnych odpowiednio ponad 2-, 3- i 1,5-krotnie.

Inne szczegółowe zmienne kategoriałne okazały się statystycznie nieistotne (poziom istotności $p \geq 0,05$).

Biorąc pod uwagę oszacowane parametry modelu oraz grupę głównych cech społeczno-ekonomicznych, można również prognozować prawdopodobieństwo odpowiedzi wyrażającej satysfakcję z wynagrodzenia. Ze względu na statystyczną istotność tylko niektórych zmiennych możliwe było rozpatrywanie jedynie określonych odpowiadających im kategorii wyróżnionych dla gospodarstw domowych.

W tabl. 9 zestawiono wyniki oczekiwanych wartości prawdopodobieństwa ze względu na wykształcenie głowy gospodarstwa domowego (przyjęto podział na wykształcenie wyższe oraz inne: podstawowe, zasadnicze i średnie). Ich analiza pozwala zauważyć, że poziom wykształcenia znacznie różnicuje szansę udzie-

¹⁷ Efekty krańcowe *JIS* w modelach logitowym i probitowym okazały się prawie zbieżne, dlatego wnioskowanie interpretacyjne przedstawiono na podstawie modelu logitowego.

lenia odpowiedzi pozytywnej. Mniej korzystne są wyniki gospodarstw domowych z wykształceniem innym niż wyższe, niezależnie od przynależności do grupy kwintylowej.

TABL. 9. OCZEKIWANE WARTOŚCI PRAWDOPODOBIEŃSTWA DLA $Y = 1$ WEDŁUG WYKSZTAŁCENIA, SKŁADU RODZINNEGO I GRUP KWINTYLOWYCH

Grupa kwintylowa Wiek w latach	Gospodarstwo			
	jednorodzinne	wielorodzinne	niepełne	jednoosobowe
Wykształcenie wyższe				
I do 50	0,18	0,24	0,16	0,07
51—60	0,28	0,36	0,25	0,11
III do 50	0,25	0,32	0,22	0,10
51—60	0,37	0,46	0,33	0,16
IV do 50	0,39	0,48	0,35	0,17
51—60	0,53	0,62	0,49	0,27
V do 50	0,64	0,72	0,60	0,36
51—60	0,76	0,82	0,73	0,50
Wykształcenie podstawowe, zasadnicze i średnie				
I do 50	0,11	0,15	0,10	0,04
51—60	0,19	0,24	0,16	0,07
III do 50	0,16	0,22	0,14	0,06
51—60	0,26	0,33	0,23	0,10
IV do 50	0,27	0,35	0,24	0,11
51—60	0,40	0,49	0,36	0,18
V do 50	0,51	0,60	0,47	0,25
51—60	0,65	0,72	0,61	0,37

Źródło: jak przy zestawieniu 1.

Biorąc pod uwagę skład rodzinny gospodarstwa domowego, można przyjąć, że cecha ta również jest czynnikiem różnicującym prawdopodobieństwo sukcesu. Szansa zadowolenia z uzyskiwanego wynagrodzenia jest największa w gospodarstwach wielorodzinnych, a następnie jednorodzinnych, najmniejsza zaś w jednoosobowych. Równocześnie, im wyższa grupa kwintylowa i starsza grupa wiekowa, tym wartości prawdopodobieństwa są większe. W odniesieniu do wieku głowy gospodarstwa dokonano podziału na grupy do 50 lat oraz 51—60 lat (grupa wiekowa powyżej 60 lat okazała się nieistotna statystycznie). Na podstawie porównania obu grup można wysnuć wniosek, że niezależnie od osiągniętych dochodów czy składu rodzinnego gospodarstwa osoby ze starszej grupy wiekowej są bardziej niż osoby młodsze zadowolone z uzyskiwanego wynagrodzenia (wartości oszacowanego prawdopodobieństwa są wyższe). Szansa ta zdecydowanie rośnie, gdy równocześnie wykształcenie głowy gospodarstwa jest wyższe.

OCENA EKONOMETRYCZNA JAKOŚCI MODELI

Jakość dopasowania modeli logitowego i probitowego zbadano, testując statystycznie łączną istotność parametrów modeli za pomocą testu ilorazu wiarygodności statystyki chi-kwadrat, a także badając oddzielny wpływ każdej zmien-

nej na zmienną zależną za pomocą statystyki chi-kwadrat Walda. Pozostałe wartości parametrów „dobroci” dopasowania modeli przedstawiono w tabl. 10.

TABL. 10. WYNIKI WERYFIKACJI MODELI

Wyszczególnienie	MODELE I	MODELE II
0 (niezadowolenie z wynagrodzenia)	$n = 707$ (70%)	
1 (zadowolenie z wynagrodzenia)	$n = 303$ (30%)	
Funkcja straty — model pełny	380,138	529,666
-2 ln (wiarygodność): model z wyrazem wolnym	1233,946	1233,946
-2 ln (wiarygodność): model pełny	760,277	1059,331
Statystyka LR	473,669	174,615
Współczynnik R^2 McFaddena	0,38	0,14

Źródło: jak przy tabl. 2.

W przypadku MODELI I obliczone wartości statystyki testowej wyniosły $LR = 473,669$ ($p = 0,0000$), a dla MODELI II — $LR = 174,615$ ($p = 0,0000$). Wartości poziomu istotności $p = 0,000$ pozwalają odrzucić hipotezę zerową i przyjąć alternatywną, że co najmniej jedna zmienna objaśniająca w jednym i drugim modelowaniu jest istotna statystycznie, co oznacza równocześnie istotność oszacowanych modeli. Uzyskane wartości końcowe, tj. funkcje straty największej wiarygodności, o wartościach: 380,138 (MODELE I) i 529,666 (MODELE II), można traktować jako miary dopasowania modeli. Wartości te oznaczają mocną wytyczną wyboru modeli dla danych empirycznych wraz ze zmiennymi objaśniającymi.

Zastosowano również inną miarę dopasowania — współczynnik R^2 McFaddena. Jego wartości były stosunkowo niskie: 0,38 (MODELE I) i 0,14 (MODELE II), nie świadczy to jednak o złym dopasowaniu modeli, ponieważ w modelach prawdopodobieństwa wartość tej miary jest z reguły niska¹⁸. Z porównania parametrów oceny jakości modeli obu grup modelowania wynika, że bardziej poprawne ekonometrycznie są modele logitowy i probitowy z grupy MODELE I.

Jeszcze inną miarą jakości dopasowania modelu jest klasyfikacja indywidualnych przypadków, wskazująca liczbę lub frakcję poprawnie i niepoprawnie sklasyfikowanych (prognozowanych) przypadków. Klasyfikację indywidualnych przypadków do grup „zadowolone” i „niezadowolone” z uzyskiwanego wynagrodzenia w badanej próbie losowej przedstawiono w tabl. 11. Z uwagi na niebilansowaną wielkość próby przyjęto zasadę optymalnej wartości granicznej δ i ustalono wartość odcinającą na poziomie $\delta = 0,3$ ¹⁹. Prognozowane przypadki wyznaczono następująco:

$$\hat{y}_i = \begin{cases} 1 & \text{gdy } \hat{p}_i > 0,3 \\ 0 & \text{gdy } \hat{p}_i \leq 0,3 \end{cases}$$

¹⁸ Można wykazać, że przy dużych próbach, gdy prawdopodobieństwo p_i nie przyjmuje wartości ekstremalnych (np. mieszczą się one w przedziale od 0,2 do 0,8), współczynnik R^2 McFaddena ma wartość ograniczoną do przedziału wartości bliskich 0, a mimo tej niskiej wartości model może być poprawny (Cox i Wermuth, 1992).

¹⁹ W badanej próbie 303 respondentów na 1010 wyraziło zadowolenie z wynagrodzenia, zatem wartość graniczną δ ustalono jako udział jedynek w grupie $\delta = 303/1010 = 0,3$.

Na tej podstawie obliczono frakcję poprawnych i niepoprawnych prognoz dotyczących gospodarstw. Klasyfikacja dla modeli logitowych i probitowych w obu grupach była identyczna.

TABL. 11. KLASYFIKACJA PRZYPADKÓW — MODELE I VS MODELE II

Obserwowane faktyczne wartości zmiennej Y	Przewidywane wartości y			
	$\hat{Y} = 1$	$\hat{Y} = 0$	łącznie	% poprawnie sklasyfikowanych przypadków
MODELE I				
$Y = 1$	221	82	303	72,94
$Y = 0$	93	614	707	86,85
Razem	314	696	1010	x
Iloraz trafień (IS)	17,79			
Zliczeniowy R^2	82,67			
MODELE II				
$Y = 1$	116	187	303	38,28
$Y = 0$	68	639	707	90,38
Razem	184	826	1010	x
Iloraz trafień (IS)	5,83			
Zliczeniowy R^2	74,75			

Źródło: jak przy tabl. 2.

Liczba przypadków poprawnej predykcji stanowiła 82,67% (MODELE I) i 74,75% (MODELE II) całej zbiorowości badanych gospodarstw domowych. W pierwszym modelowaniu trafność prognozowania wartości $Y = 1$ wynosiła prawie 73% łącznej liczby przypadków, natomiast trafność prognozowania wartości $Y = 0$ (niezadowolonych z wynagrodzenia) — 87%. Inne frakcje uzyskano w grupie MODELE II, gdzie trafność prognozowania dla zmiennej objaśnianej $Y = 0$ była lepsza (ponad 90% przypadków). Ogólna klasyfikacja przypadków okazała się prawidłowa, co można uznać za satysfakcjonujący wynik. Również wartości ilorazu trafień (szans) wyniosły w obu przypadkach odpowiednio $IS = 17,79$ oraz $IS = 5,83$, co oznacza, że klasyfikacje uzyskane dzięki oszacowanym modelom były lepsze od klasyfikacji całkowicie przypadkowych.

Podsumowanie

Wynagrodzenie stanowi istotny czynnik dochodu, a zatem i zamożności gospodarstw domowych. Należy przypisać mu funkcję motywacyjną, która przejawia się w skłonności ludzi do podejmowania pracy, oraz funkcję dochodową (płaca jest bezpośrednim źródłem dochodu, który służy do zaspokajania potrzeb materialnych i niematerialnych jednostki oraz jej rodziny). Znaczenie wynagrodzenia ma szeroki kontekst ekonomiczno-społeczny. Jego wysokość określa pozycję ekonomiczną i społeczną, buduje prestiż, indywidualne poczucie ważności i osobistej samorealizacji. Subiektywne postrzeganie i ocena wysokości wynagrodzenia ma zarówno osobisty, jak i społeczno-ekonomiczny oraz psychologiczny kontekst. Subiektywna ocena wynagrodzenia uzyskiwanego w danym

czasie przez jednostkę może stanowić próbę pomiaru równoległego do czysto obiektywnego pomiaru w statystyce publicznej.

W niniejszym artykule przedstawiono propozycję zastosowania i porównania modeli dwumianowych w celu modelowania — na podstawie danych empirycznych uzyskanych z badania losowej próby gospodarstw domowych w woj. podkarpackim — subiektywnych ocen zadowolenia/braku satysfakcji z wynagrodzenia. Na bazie tego samego materiału empirycznego uzyskano dwie odrębne grupy poprawnych ekonometrycznie modeli logitowych i probitowych; w każdej przyjęto inną wstępną grupę zmiennych niezależnych charakteryzujących określone identyfikowalne uwarunkowania respondentów. Zaletą (a równocześnie koniecznością) wykorzystania tego typu modeli była możliwość ujęcia oprócz zmiennych ilościowych także zmiennych jakościowych. Efektem modelowania ekonometrycznego była ścisła identyfikacja istotnych statystycznie czynników wpływających na satysfakcję/brak satysfakcji z wynagrodzenia. Ocenie poddano natężenie oraz kierunek wpływu wybranych cech i czynników społeczno-demograficznych na subiektywną ocenę satysfakcji z wynagrodzenia wyrażaną przez głowę gospodarstwa domowego.

W rezultacie porównania oszacowanych parametrów modeli logitowego i probitowego (w obu zestawach zmiennych niezależnych) uzyskano jednakową interpretację wyników. Poprawna ekonometrycznie regresja logistyczna i probitowa pozwoliła na oszacowanie prawdopodobieństwa zadowolenia z uzyskiwanego wynagrodzenia w zależności od tych samych zmiennych objaśniających. W grupie MODELE I istotne statystycznie okazały się zmienne: *przynależność do dochodowej grupy kwintylowej*, *subiektywna ocena aktualnego materialnego poziomu życia*, *subiektywna ocena zadowolenia z ogólnej jakości życia*, *oszczędności* oraz *liczba osób osiągających dochody* i *liczba osób ogółem w gospodarstwie domowym*.

Najistotniejszą zmienną, a zarazem najważniejszym czynnikiem zidentyfikowanym przez modele, była *przynależność do dochodowej grupy kwintylowej*. Prawdopodobieństwo wyrażenia satysfakcji wynagrodzenia znacznie rosło wraz z wysokością dochodów (przynależnością do wyższej grupy kwintylowej). Kolejne istotnie ważne czynniki stanowiły: *subiektywna ocena aktualnego materialnego poziomu życia* oraz *subiektywna ocena zadowolenia z ogólnej jakości życia*. Szansa pozytywnej oceny wynagrodzenia zwiększała się wraz ze wzrostem oceny tych czynników. Taki sam kierunek wpływu miał inny statystycznie istotny czynnik — *liczba członków gospodarstwa domowego osiągających dochody*, natomiast odwrotny miała zmienna *stricte* demograficzna. Wzrost liczby osób w gospodarstwie domowym powodował spadek prawdopodobieństwa (najmniej odczuwalny w przypadku najbogatszych gospodarstw).

Drugie modelowanie (MODELE II) zarówno w modelu logitowym, jak i probitowym przyniosło zbieżne wyniki analityczne i interpretacje. Badanie wpływu klasycznych zmiennych niezależnych wykazało, że statystycznie istotnymi czynnikami były: *grupa kwintylowa*, *wykształcenie*, *wiek* i *skład rodzinny gospodarstwa domowego*. Znaczne różnice prawdopodobieństwa satysfakcji z wynagrodzenia powodował też poziom wykształcenia. W gospodarstwach domowych z wykształceniem podstawowym, zasadniczym lub średnim (niezależnie od przynależności do grupy kwintylowej czy składu rodzinnego gospodarstwa) było ono niższe niż w gospodarstwach, w których głowa miała wykształcenie wyższe. Równocześnie

przynależność gospodarstwa do wyższej grupy kwintylowej (IV, V) i starszy wiek powodował wzrost szansy pozytywnej oceny wynagrodzenia. Z modelowania ekonometrycznego wynika także, że szansa ta była większa w przypadku gospodarstw wielorodzinnych i jednorodzinnych w porównaniu z gospodarstwami niepełnymi i jednoosobowymi. Uwzględniając z kolei wiek głowy gospodarstwa, niezależnie od uzyskiwanych dochodów oraz składu rodzinnego, można stwierdzić, że bardziej zadowolone z wynagrodzenia były gospodarstwa ze starszej grupy wiekowej. Szansa sukcesu znacznie zwiększała się wraz ze wzrostem wykształcenia głowy gospodarstwa. Regresja logistyczno-probitowa nie wykazała statystycznie istotnych zależności takich cech, jak: *miejsce zamieszkania, forma zatrudnienia i płeć głowy gospodarstwa domowego*.

Szczególnym celem przeprowadzonego badania była możliwość zastosowania określonego narzędzia ekonometrycznego — tutaj: modeli logitowych i probitowych — w wybranym obszarze zagadnień ekonomiczno-społecznych. Modele dwumianowe pozwoliły, przy całej złożoności omawianego zjawiska związanego z wynagrodzeniem i jego subiektywną oceną, na uchwycenie zależności tej oceny od pewnych cech ekonomicznych, społecznych, demograficznych, socjologicznych i innych, co może znaleźć zastosowanie w polityce płacowo-dochodowej. Należy podkreślić, że nie jest to badanie kompletne i wyczerpujące, chociażby ze względu na jego ograniczony zakres przedmiotowy, czasowy i przestrzenny. W związku z tym należałoby w dalszych badaniach empirycznych tego typu uwzględnić pomiar, identyfikację i modelowanie kolejnych potencjalnych zmiennych w celu bardziej precyzyjnego i pełniejszego wyjaśnienia kształtowania się zjawisk ekonomicznych i społecznych w sferze polityki płacowej i dochodowej ludności.

dr Beata Kasprzyk — Uniwersytet Rzeszowski

LITERATURA

- Atkinson, A., Banker, R., Kaplan, R. (1995). *Management Accounting*. Englewood Cliffs (New York): Prentice Hall.
- Borkowska, S. (1999). *Wynagrodzenie godziwe*. Warszawa: Instytut Pracy i Spraw Socjalnych.
- Chow, G. C. (1995). *Ekonometria*. Warszawa: Wydawnictwo Naukowe PWN.
- Cieślak, I., Kucharczyk, M. (2006). Międzynarodowe relacje wynagrodzeń. Analiza na podstawie wartości kapitału ludzkiego. *Nierówności społeczne a wzrost gospodarczy*, (8).
- Cox, D. R., Wermuth, N. (1992). A comment on the coefficient of determination for binary responses. *The American Statistician*, 46.
- Cramer, J. S. (2004). The early origins of the logit model. *Studies in History and Philosophy of Biological and Biomedical Sciences*, (35).
- Domański, Cz. (1990). *Testy statystyczne*. Warszawa: PWE.
- Friedman, M. (1993). *Kapitalizm i wolność*. Warszawa: Wydawnictwo Naukowe PWN.
- Greene, W. H. (2000). *Econometric Analysis*. Upper Saddle River (New Jersey): Prentice Hall Inc.
- Gruszczyński, M. (2001). *Modele i prognozy zmiennych jakościowych w finansach i bankowości*. Warszawa: Oficyna Wydawnicza Szkoły Głównej Handlowej.
- Gruszczyński, M. (red.). (2010). *Mikroekonometria. Modele i metody analizy danych indywidualnych*. Warszawa: Wolters Kluwer Polska.
- Gruszczyński, M., Podgórska, M. (red.). (1996). *Ekonometria*. Warszawa: Oficyna Wydawnicza Szkoły Głównej Handlowej.

- GUS (2016a). *Budżety gospodarstw domowych w 2015 r.* Warszawa: GUS.
- GUS (2016b). *Regionalne zróżnicowanie wynagrodzeń*. Pobrane z: <http://stat.gov.pl/obszary-tematyczne/rynek-pracy/pracujacy-zatrudnieni-wynagrodzenia-koszty-pracy/regionalne-zroznicowanie-wynagrodzen,11,1.html>.
- GUS (2016c). *Regiony Polski*. Warszawa: GUS.
- Jacukowicz, Z. (2003). *Płaca i praca w warunkach przemian gospodarczych i globalizacji*. W: B. Balcerzak-Paradowska (red.), *Praca i polityka społeczna wobec wyzwań integracji*. Warszawa: IPISS.
- Judge, G. G., Hill, C., Griffiths, W. E., Lütkepohl, H., Lee, T., (1985). *The Theory and Practice of Econometrics*. New York: John Wiley&Sons.
- Kufel, T. (2011). *Ekonometria. Rozwiązywanie problemów z wykorzystywaniem programu Gretl*. Warszawa: PWN.
- Maddala, G. S. (2008). *Ekonometria*. Warszawa: Wydawnictwo Naukowe PWN.
- Ostasiewicz, W. (red.). (1998). *Statystyczne metody analizy danych*. Wrocław: Wydawnictwo Akademii Ekonomicznej we Wrocławiu.
- Panek, T., Czapiński, J. (2015). Warunki życia gospodarstw domowych. Dochody i sposób gospodarowania dochodami. Diagnostyka Społeczna 2015. Warunki i Jakość Życia Polaków — Raport. *Contemporary Economics*, 9(4), 36—50.
- Śliwicki, D., Ręklawski, M. (2012). Wykorzystanie modeli logitowych w analizie czynników aktywności zawodowej ludności, *Wiadomości Statystyczne*, (2), Warszawa.
- Smith, A. (1954). *Badania nad naturą i przyczynami bogactw narodów*, t. 2. Warszawa: PWN.
- Stanisz, A. (2007). *Przystępny kurs z zastosowaniem Statistica PL na przykładach z medycyny*, 3. Kraków: Statsoft.
- Särndal, C. E., Lundström, S. (2006). *Estimation in Surveys with Nonresponse*. Chichester: J. Wiley.
- Szreder, M. (2015). Zmiany w strukturze całkowitego błędu badania próbkowego. *Wiadomości Statystyczne*, (1).
- Steczkowski, J., Zeliaś, A. (1997). *Metody statystyczne w badaniu zjawisk jakościowych*. Kraków: Wydawnictwo Akademii Ekonomicznej w Krakowie.
- Zajac, K., (2005). Święty Tomasz z Akwinu. Jego zapatrywania socjalne i ekonomiczne. *Annales. Etyka w życiu gospodarczym*, (8).
- Zeliaś, A., Pawełek, B., Wanat, S. (2003). *Prognozowanie ekonomiczne*. Warszawa: Wydawnictwo Naukowe PWN.

Summary. *The article presents the results of the research attempting to make a statistical analysis of subjective perception of wages and salaries as the main component of household income, using a specific econometric modelling. It was examined whether certain socio-demographic determinants could be identified and, if so, to what extent they could influence the subjective sense of satisfaction with wage and salaries. The research was conducted on the basis of data collected using a survey method in 2015 for a random sample of households in Podkarpackie voivodship, where wages and salaries are among the lowest in the country. The binomial models used to explain the qualitative variable depending on the level of exogenous variables (qualitative and quantitative) were compared.*

Logit and probit regression models made it possible to determine the probability of success as a chance of a positive response, i.e. (in relation to the scope of the research) a state of satisfaction with obtained wages and salaries.

Keywords: theory of wages, wages and salaries, household, logit model.