

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

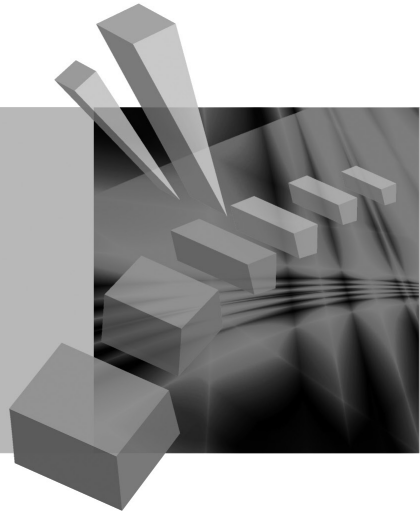
RESEARCH PAPERS

of Wrocław University of Economics

278

Taksonomia 20

Klasyfikacja i analiza danych – teoria i zastosowania



Redaktorzy naukowi

Krzysztof Jajuga

Marek Walesiak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2013

Redaktor Wydawnictwa: Aleksandra Śliwka

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Małgorzata Czupryńska

Projekt okładki: Beata Dębska

Publikacja jest dostępna w Internecie na stronach:

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

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

a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon

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

Informacje o naborze artykułów i zasadach recenzowania znajdują się

na stronie internetowej Wydawnictwa

www.wydawnictwo.ue.wroc.pl

Tytuł dofinansowany ze środków Narodowego Banku Polskiego

oraz ze środków Sekcji Klasyfikacji i Analizy danych PTS

Kopiowanie i powielanie w jakiegokolwiek formie

wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu

Wrocław 2013

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

ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

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

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

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

Summaries

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

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

Hanna Dudek, Joanna Landmesser

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

WPŁYW RELATYWNEJ DEPRYWACJI NA SUBIEKTYWNE POSTRZEGANIE DOCHODÓW

Streszczenie: W artykule przedstawiono analizę kształtowania się subiektywnej oceny własnej sytuacji dochodowej przez gospodarstwa domowe emerytów. W tym celu zastosowano metodę częściowo uogólnionych uporządkowanych modeli logitowych. Do determinant subiektywnego postrzegania dochodów zaliczono relatywną deprywację, wykształcenie, wiek, płeć, miejsce zamieszkania, fakt pozostawania w związku (formalnym lub nie) oraz dodatkowe źródło utrzymania z użytkowania gospodarstwa rolniczego. Analizę przeprowadzono na podstawie danych z badań budżetów gospodarstw domowych zrealizowanych przez GUS w 2009 r.

Słowa kluczowe: relatywna deprywacja, dochody gospodarstw domowych, modele logitowe.

1. Wstęp

Badania dotyczące subiektywnego postrzegania własnej sytuacji dochodowej stanowią stosunkowo nowy, lecz dynamicznie rozwijający się dział statystyki społecznej. Prace dotyczące analizy dobrobytu gospodarstw domowych w wielu rozwiniętych krajach nie ograniczają się jedynie do sfery konsumpcji i obiektywnej kondycji dochodowej, lecz uwzględniają także zagadnienie deprywacji obejmującej wiele aspektów poziomu życia, w tym deprywacji subiektywnej. Poznanie uwarunkowań satysfakcji z osiągniętych dochodów może pomóc w kształtowaniu polityki socjalnej mającej na celu złagodzenie skutków subiektywnego ubóstwa.

W pracy podjęto temat określenia determinant satysfakcji z osiągniętych dochodów przez gospodarstwa domowe emerytów. Ta grupa bowiem stanowi coraz liczniejszą część polskiego społeczeństwa i między innymi z tego względu wymaga starannych analiz. W celu wyjaśnienia subiektywnej oceny dochodów gospodarstw domowych emerytów uwzględniono relatywną deprywację dochodową¹. Jej poziom

¹ Pojęcie deprywacji oznacza stan wywołany brakiem możliwości zaspokojenia istotnej potrzeby. W szczególności deprywacja dochodowa dotyczy sytuacji, gdy powodem niezaspokojenia danej potrzeby jest sytuacja finansowa. Deprywację można pojmować w sposób relatywny (względny) lub absolutny. W ujęciu relatywnym poziom zaspokojenia potrzeb analizowanych jednostek odnosi się do poziomu ich zaspokojenia przez innych członków społeczeństwa [Panek 2011]. W ujęciu absolutnym ten aspekt nie jest uwzględniany.

dla danego gospodarstwa domowego określono na podstawie różnic między jego dochodami a dochodami osiąganymi przez inne gospodarstwa emerytów. Ponadto w charakterze dalszych determinant subiektywnej oceny dochodów rozpatrzono wiele cech socjodemograficznych, takich jak wiek, płeć oraz miejsce zamieszkania.

W badaniu zastosowano metodę częściowo uogólnionych uporządkowanych modeli logitowych. Analizę przeprowadzono na podstawie danych z badań budżetów gospodarstw domowych zrealizowanych przez GUS w 2009 r. Proponowana metodyka jest modyfikacją podejścia zaprezentowanego w publikacjach [Vera-Toscano i in. 2006; Ferrer-i-Carbonell 2005], gdzie do estymacji parametrów modelu objaśniającego satysfakcję z osiąganymi dochodów wykorzystano uporządkowane modele probitowe i logitowe, nie weryfikując dość silnych założeń nakładanych na te modele.

2. Dane

Wykorzystane w pracy dane pochodzą z badań budżetów gospodarstw domowych prowadzonych przez GUS w 2009 r. Analizę przeprowadzono na podstawie informacji z 10 472 gospodarstw domowych emerytów. Dane na temat subiektywnego postrzegania swojej sytuacji dochodowej przez te gospodarstwa odnoszą się do odpowiedzi na jedno z pytań uwzględnionych w kwestionariuszach stosowanych przez GUS: „Czy z aktualnym dochodem Pana(i) gospodarstwo domowe wiąże koniec z końcem: 1) z wielką trudnością, 2) z trudnością, 3) z pewną trudnością, 4) raczej łatwo, 5) łatwo, 6) bardzo łatwo”. W tabeli 1 przedstawiono podstawowe informacje dotyczące oceny własnych dochodów przez gospodarstwa domowe emerytów w 2009 r.

Tabela 1. Subiektywne postrzeganie własnych dochodów

| Z aktualnym dochodem gospodarstwo wiązało koniec z końcem | Kategoria | Odsetek gospodarstw domowych emerytów |
|---|-----------|---------------------------------------|
| Z wielką trudnością | $j = 1$ | 9,26% |
| Z trudnością | $j = 2$ | 18,5% |
| Z pewną trudnością | $j = 3$ | 44,69% |
| Raczej łatwo | $j = 4$ | 22,62% |
| Łatwo | $j = 5$ | 4,32% |
| Bardzo łatwo | $j = 6$ | 0,61% |

Źródło: obliczenia własne.

Ze względu na fakt, iż odsetek gospodarstw domowych emerytów oceniających, że z aktualnymi dochodami było im bardzo łatwo wiązać koniec z końcem, nie przekraczał 1%, to gospodarstwa te połączono z grupą tych, którym łatwo przychodziło wiązanie końca z końcem. Ostatecznie zatem rozpatrywano nie 6, lecz 5 kategorii odpowiedzi.

3. Metoda

W celu określenia wartości względnej (relatywnej) deprywacji dochodowej na początku rozpatrzono dochody ekwiwalentne, uwzględniając tzw. zmodyfikowaną skalę OECD, tzn. dochody rozporządzalne każdego gospodarstwa domowego podzielono przez odpowiadającą mu wartość skali ekwiwalentności, otrzymując dochody ekwiwalentne y_1, y_2, \dots, y_n , gdzie $n = 10\,472$ – liczba gospodarstw domowych emerytów w próbie GUS. Następnie uporządkowano niemalejąco tę cechę, uzyskując wartości $y_{(1)}, y_{(2)}, \dots, y_{(n)}$, spełniające zależność $y_{(1)} \leq y_{(2)} \leq \dots \leq y_{(n)}$. Dla każdego gospodarstwa o dochodzie ekwiwalentnym $y_{(i)}$ wyznaczono wartość funkcji relatywnej deprywacji dochodowej na podstawie wzoru:

$$d(y_{(i)}) = \frac{1}{n} \sum_{j=i+1}^n (y_{(j)} - y_{(i)}), \quad d(y_{(n)}) = 0 \quad [\text{D'Ambrosio, Frick 2007}]. \quad (1)$$

W celu określenia jej wpływu na subiektywne postrzeganie własnej sytuacji dochodowej przez gospodarstwa domowe rozpatrzono modele wyjaśniające kształtowanie się zmiennej wyrażonej na skali porządkowej. Ogólna postać tych modeli jest następująca:

$$u_i^* = \mathbf{x}_i \boldsymbol{\beta} + \varepsilon_i, \quad (2)$$

gdzie: u_i^* – zmienna nieobserwowalna odnosząca się do i -tej obserwacji,

$\boldsymbol{\beta}$ – wektor kolumnowy parametrów $\beta_0, \beta_1, \beta_2, \dots, \beta_K$,

\mathbf{x}_i – wektor wierszowy wartości zmiennych objaśniających dla i -tej obserwacji,

ε_i – składnik losowy dla i -tej obserwacji, $i = 1, 2, \dots, n$, n – liczba obserwacji.

W powyższym podejściu ciągła nieobserwowalna zmienna u^* jest dyskretyzowana przez zbiór wartości progowych $\delta_0, \delta_1, \dots, \delta_m$, co prowadzi do zmiennej skategoryzowanej u określonej jako: $u_i = j$, jeśli $\delta_j < u_i^* \leq \delta_{j+1}$, przy czym: $-\infty = \delta_0 < \delta_1 < \dots < \delta_m < \delta_{m+1} = \infty$. Prawdopodobieństwa $P(u_i = j | \mathbf{x}_i)$ w uporządkowanym modelu można określić jako: $P(u_i = j | \mathbf{x}_i) = F(\delta_j - \mathbf{x}_i \boldsymbol{\beta}) - F(\delta_{j-1} - \mathbf{x}_i \boldsymbol{\beta})$, dla $j = 1, 2, \dots, m$ ², gdzie F jest dystrybucją składnika losowego ε . W zastosowaniach zwykle wykorzystuje się uporządkowane modele logitowy i probitowy, w któ-

² W niniejszej pracy uwzględniono 5 kategorii odpowiedzi, stąd parametr $m = 5$.

rych odpowiednio $F(z) = \Lambda(z) = \frac{1}{1 + \exp(-z)}$ oraz $F(z) = \Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$, gdzie: $z_i = \mathbf{x}_i \boldsymbol{\beta}$. W modelach tych zakłada się, że:

$$P(u_i \leq j | \mathbf{x}_i) = F(\delta_j - \mathbf{x}_i \boldsymbol{\beta}), \quad (3)$$

czyli parametry przy zmiennych objaśniających nie zależą od kategorii j , $j = 1, 2, \dots, m$. Do weryfikacji warunku (3), zwanego założeniem równoległych regresji, w niniejszej pracy zastosowano test Branta. Jego idea polega na rozważeniu $m - 1$ modeli dwumianowych, w których zmienne objaśniane zdefiniowane są następująco³:

$$u_j^{**} = \begin{cases} 1, & \text{gdy } u > j, \\ 0, & \text{w przeciwnych przypadkach,} \end{cases} \quad (4)$$

gdzie $j = 1, 2, \dots, m - 1$.

Po oszacowaniu łącznej macierzy wariancji i kowariancji parametrów we wszystkich $m - 1$ modelach dwumianowych przeprowadza się testy Walda – łączny oraz indywidualne testy dla poszczególnych zmiennych objaśniających. Pierwszy z testów wykorzystywany jest do weryfikacji hipotezy zerowej mówiącej o równości odpowiednich parametrów we wszystkich $m - 1$ regresjach dwumianowych łącznie dla całego zestawu zmiennych objaśniających⁴. Odrzucenie hipotezy zerowej oznacza, że co najmniej dla jednej zmiennej objaśniającej parametry istotnie różnią się w przynajmniej dwóch modelach dwumianowych, w następstwie czego założenie regresji równoległych nie jest spełnione. Przeprowadzenie indywidualnych testów Walda umożliwia wskazanie zmiennych „odpowiedzialnych” za naruszenie tego założenia. Hipoteza zerowa wtedy mówi o równości parametrów przy poszczególniej zmiennej objaśniającej we wszystkich modelach dwumianowych. Jeśli co najmniej jedna z powyższych hipotez zostanie odrzucona, to model dla kategorii uporządkowanych nie powinien być stosowany. W takiej sytuacji można wykorzystać ideę uogólnionego modelu uporządkowanego, gdzie:

$$P(u_i \leq j | \mathbf{x}_i) = F(\delta_j - \mathbf{x}_i \boldsymbol{\beta}_j), \text{ dla } j = 1, 2, \dots, m - 1 \text{ [Greene, Hensher 2010]}. \quad (5)$$

W przypadku, gdy indywidualne testy Walda nie nakazują odrzucenia hipotez o równości parametrów niektórych ze zmiennych objaśniających, to możliwe jest

³ Opis testu Branta podano za publikacjami [Long 1997; Książek 2010; Dudek 2012].

⁴ Statystyka testowa w teście Branta ma asymptotyczny rozkład chi-kwadrat z $p(m - 2)$ stopniami swobody, gdzie p – liczba parametrów przy zmiennych objaśniających, m – liczba kategorii zmiennej porządkowej odpowiadającej nieobserwowalnej zmiennej objaśnianej.

zastosowanie częściowo uogólnionych modeli uporządkowanych⁵. W modelach tego typu we wzorze (5) parametry przy niektórych zmiennych objaśniających nie zależą od kategorii analizowanej zmiennej porządkowej. Takie podejście jest z jednej strony mniej restrykcyjne niż zastosowanie modelu uporządkowanego, z drugiej strony zaś – z uwagi na uwzględnienie w modelu mniejszej liczby parametrów – bardziej „oszczędne” oraz umożliwiające łatwiejszą interpretację parametrów strukturalnych modelu.

W pracy rozważono model (2), w którym zmienna u^* odnosi się do satysfakcji z osiąganego dochodu, jedna ze zmiennych objaśniających – do relatywnej deprivacji dochodowej, pozostałe zmienne objaśniające zaś do różnych cech socjodemograficznych.

4. Wyniki

Oszacowano wiele modeli wyjaśniających kształtowanie satysfakcji z osiągniętych dochodów. Przy wyborze zmiennych objaśniających kierowano się względami merytorycznymi i statystycznymi. Do porównania modeli z różnym zestawem zmiennych objaśniających w modelach zastosowano kryteria informacyjne Akaike’a i Schwarza. Ostatecznie otrzymano oszacowanie dla uporządkowanego modelu logitowego⁶, którego wyniki zaprezentowano w tab. 2. W charakterze ilościowych zmiennych objaśniających uwzględniono relatywną deprivację dochodową (oznaczenie: deprivacja), wiek osoby odniesienia⁷ oraz wiek osoby odniesienia do kwadratu. Zmienne jakościowe są zmiennymi binarnymi przyjmującymi wartość 1 dla wariantu wskazanego nazwą oraz 0 w pozostałych przypadkach.

Na początku rozpatrywano uporządkowane modele logitowe. Po wykonaniu łącznego testu Branta stwierdzono, że założenie regresji równoległych powinno być odrzucone. Analiza wyników indywidualnego testu Branta wykazała, że „odpowiedzialność” za naruszenie tego założenia ponoszą cztery zmienne wymienione na początku tab. 3.

Dla zestawu zmiennych prezentowanego w tab. 3 wartość statystyki testowej w łącznym teście Branta wyniosła 263,25, co świadczy o odrzuceniu założenia regresji równoległych (wartość krytyczna, odczytana z tablic rozkładu chi-kwadrat, dla 39 stopni swobody i poziomu istotności 0,05 jest równa 54,57). Z informacji przedstawionych w tab. 3 wynika, że dla zmiennych odnoszących się do względnej

⁵ W anglojęzycznej literaturze przedmiotu modele takie są nazywane mianem *partial proportional odds models for ordinal response variables* [Peterson, Harrell 1990; Williams 2006].

⁶ Wyniki prezentowane w tab. 2-5 uzyskano na podstawie danych dotyczących gospodarstw składających się z co najwyżej 7 osób, których wartości dochodów ekwiwalentnych oraz wartości wskaźnika deprivacji kształtowały się w granicach $[Q_1 - 2 * IQR; Q_3 + 2 * IQR]$, gdzie Q_1 i Q_3 – 1. i 3. kwartyl, IQR – rozstęp ćwiartkowy. Pozostałe obserwacje, stanowiące ok. 2,7% próby, uznano za odstające i wyłączone z analizy.

⁷ Osoba odniesienia to osoba, która ukończyła 16 lat i osiągnęła najwyższy dochód spośród wszystkich członków gospodarstwa domowego.

Tabela 2. Wyniki estymacji parametrów uporządkowanego modelu logitowego

| Zmienna | Ocena parametru | Błąd standardowy | Zmienna | Ocena parametru | Błąd standardowy |
|------------------------------------|-----------------|------------------|-------------------|-----------------|------------------|
| Deprywacja | -0,0046 | 0,0001 | kobieta | -0,2832 | 0,0341 |
| Miasto powyżej 100 tys. mieszk. | -0,3418 | 0,0518 | region* centralny | -0,3813 | 0,0507 |
| Miasto do 100 tys. mieszkańców | -0,1360 | 0,0480 | region południowy | 0,1428 | 0,0512 |
| Singiel | -0,2889 | 0,0478 | region wschodni | 0,2010 | 0,0554 |
| Dodatkowe źródło utrż. z rolnictwa | 0,3293 | 0,0865 | stała_1 | -7,8540 | 0,5142 |
| Wykształcenie wyższe | 0,5148 | 0,0672 | stała_2 | -6,2104 | 0,5117 |
| Wykształcenie średnie | 0,1013 | 0,0456 | stała_3 | -3,6698 | 0,5091 |
| Wiek | -0,1038 | 0,0001 | stała_4 | -1,3015 | 0,5099 |
| Wiek ² | 0,0009 | 0,0000 | – | – | – |

* W bazie danych pozyskanej z GUS, stanowiącej podstawę analizy, uwzględniono podział na 6 regionów: południowo-zachodni, północno-zachodni, północny, centralny, południowy i wschodni.

Źródło: obliczenia własne w programie Stata v. 10.

Tabela 3. Wyniki testu Branta dla uporządkowanego modelu logitowego

| Zmienna | Wartość statystyki testowej | Wartość p | Zmienna | Wartość statystyki testowej | Wartość p |
|---|-----------------------------|-------------|-------------------|-----------------------------|-------------|
| Deprywacja | 93,03 | < 0,01 | wiek | 0,37 | 0,947 |
| Miasto powyżej 100 tys. mieszk. | 33,16 | < 0,01 | wiek ² | 0,24 | 0,971 |
| Miasto do 100 tys. mieszkańców | 10,73 | 0,013 | kobieta | 4,91 | 0,179 |
| Singiel | 30,82 | < 0,01 | region centralny | 7,54 | 0,057 |
| Dodatkowe źródło utrzymania z gospodarstwa rolniczego | 1,66 | 0,645 | region południowy | 3,62 | 0,305 |
| Wykształcenie wyższe | 1,13 | 0,769 | region wschodni | 5,79 | 0,122 |
| Wykształcenie średnie | 0,67 | 0,879 | – | – | – |

Źródło: obliczenia własne w programie Stata v. 10. Oznaczenie „<0,01” odnosi się do wartości prawdopodobieństwa testowego mniejszej niż 0,01.

deprywacji, miejsca zamieszkania oraz niepozostawania osoby odniesienia w związku (formalnym lub nieformalnym) odrzucono hipotezę, że parametry przy każdej z tych zmiennych nie zależą od kategorii. Dlatego też wykorzystano częściowo uogólniony model uporządkowany. Wyniki estymacji⁸ tych parametrów modelu, których wartości nie zależały od przynależności do kategorii zmiennej porządkowej, odnoszącej się do subiektywnej oceny dochodów, zaprezentowano w tab. 4.

Oszacowania pozostałych parametrów częściowo uogólnionego uporządkowanego modelu logitowego, które różnią się w poszczególnych kategoriach, zamieszczono w tab. 5.

⁸ W tabeli 4 zamieszczono oszacowania parametrów przy tych zmiennych, dla których wartość p w tab. 3 przekraczała 0,05.

Tabela 4. Wyniki estymacji parametrów częściowo uogólnionego uporządkowanego modelu logitowego

| Zmienna | Ocena parametru | Błąd standardowy | Zmienna | Ocena parametru | Błąd standardowy |
|---|-----------------|------------------|------------------|-----------------|------------------|
| Wykształcenie wyższe | 0,3898 | 0,0667 | kobieta | -0,2244 | 0,0455 |
| Wykształcenie średnie | 0,0877 | 0,0455 | region centralny | -0,3798 | 0,0502 |
| Wiek | -0,0984 | 0,0156 | region | 0,1430 | 0,0512 |
| Wiek ² | 0,0008 | 0,0001 | region wschodni | 0,2055 | 0,0559 |
| Dodatkowe źródło utrzymania z gospodarstwa rolniczego | 0,3087 | 0,0893 | - | - | - |

Źródło: obliczenia własne w programie Stata v. 10.

Tabela 5. Wyniki estymacji parametrów częściowo uogólnionego uporządkowanego modelu logitowego, cd.

| Zmienna | Dla $u > 1$ | Dla $u > 2$ | Dla $u > 3$ | Dla $u > 4$ |
|----------------------------------|------------------|------------------|------------------|------------------|
| Deprywacja | -0,0042 (0,0002) | -0,0041 (0,0001) | -0,0055 (0,0002) | -0,0077 (0,0005) |
| Miasto pow. 100 tys. mieszkańców | -0,5517 (0,0963) | -0,5119 (0,0670) | -0,2184 (0,0667) | 0,1901 (0,1387) |
| Miasto do 100 tys. mieszkańców | -0,2580 (0,0917) | -0,2205 (0,0627) | -0,0178 (0,0636) | 0,0486 (0,1439) |
| Singiel | -0,5393 (0,0828) | -0,4137 (0,0574) | -0,1278 (0,0594) | 0,0782 (0,1119) |
| Stała | 7,7295 (0,5287) | 5,9682 (0,5206) | 3,5833 (0,5186) | 1,2712 (0,5319) |

Źródło: obliczenia własne w Stata v. 10. W nawiasach podano błędy standardowe szacunku parametrów.

Na podstawie wyników estymacji można stwierdzić, że przy założeniu *ceteris paribus*:

- im wyższa względna deprywacja dochodowa, tym większe prawdopodobieństwo tego, że gospodarstwo oceniało swoją sytuację dochodową jako bardzo trudną;
- prawdopodobieństwo, że z aktualnym dochodem można było związać koniec z końcem łatwo lub bardzo łatwo, było większe dla gospodarstw, w których osoba odniesienia (głowa gospodarstwa domowego) legitymowała się wyższym lub średnim wykształceniem niż w sytuacji braku takiego wykształcenia;
- jeśli osoba odniesienia nie pozostawała w związku partnerskim (formalnym lub nie), to prawdopodobieństwo bardzo złej oceny własnej sytuacji dochodowej było większe niż w przypadku osoby będącej w takim związku;
- prawdopodobieństwo dobrej lub bardzo dobrej satysfakcji z osiągniętych dochodów gospodarstw w sytuacji, w której osoba odniesienia była kobietą, było niższe niż w przypadku mężczyzn;

- prawdopodobieństwo dużej satysfakcji z osiągniętych dochodów gospodarstw na początku zmniejszało się z wiekiem, po czym po przekroczeniu 60 roku życia rosło;
- jeśli gospodarstwo domowe posiadało dodatkowe źródło utrzymania z użytkowania gospodarstwa rolniczego, to prawdopodobieństwo bardzo dobrej i dobrej oceny własnej sytuacji dochodowej było większe niż w przypadku gospodarstwa niemającego dodatkowego źródła tego rodzaju;
- gospodarstwa zamieszkujące miasta częściej oceniały swą sytuację jako bardzo trudną niż gospodarstwa zamieszkujące na wsi;
- w stosunku do gospodarstw z rejonów południowo-zachodniego, północno-zachodniego i północnego (tzw. ziemie odzyskane) prawdopodobieństwo oceny swojej sytuacji dochodowej jako bardzo trudnej było większe w rejonie centralnym i mniejsze w rejonie południowym i wschodnim.

5. Wnioski

Satysfakcja z osiągniętych dochodów przez gospodarstwa domowe emerytów w 2009 r. zależała od ich relatywnej deprivacji. Fakt ten oznacza, że na postrzeganie swojej sytuacji dochodowej wpływało porównywanie własnej, rzeczywistej kondycji finansowej z kondycją innych gospodarstw domowych emerytów w Polsce. Ponadto, podobnie jak w pracach [Ferrer-i-Carbonell 2005; D'Ambrosio, Frick 2007; Vera-Toscano i in. 2006], stwierdzono, że w charakterze determinant subiektywnej oceny dochodów należy uwzględnić wiek, płeć, wykształcenie, fakt pozostawania w związku partnerskim osoby odniesienia oraz miejsce zamieszkania.

W kwestii zależności satysfakcji z osiągniętych dochodów od miejsca zamieszkania zaobserwowano interesujące relacje. Stwierdzono, że gospodarstwa z rejonów w przybliżeniu należących do tzw. Polski B częściej niż gospodarstwa z tzw. Polski A deklarowały dobrą lub bardzo dobrą ocenę swojej sytuacji dochodowej. Generalnie ludność Podlasia, Lubelszczyzny, Podkarpacia i Małopolski ma silne poczucie własnej tożsamości oraz jest przywiązana do tradycyjnych wartości. Fakt ten prawdopodobnie wpływa na pozytywne postrzeganie swojej sytuacji dochodowej. Podobnie mieszkańcom wsi częściej łatwo lub bardzo łatwo było związać koniec z końcem niż mieszkańcom miast, gdyż w mniejszych społecznościach na ogół obserwuje się silniejsze więzi społeczne, co może przekładać się na bardziej optymistyczną ocenę swojego położenia, w szczególności na ocenę własnej kondycji finansowej.

Podobnie jak w literaturze przedmiotu do opisu kształtowania się własnego postrzegania swojej kondycji dochodowej przez gospodarstwa domowe rozpatrzono możliwość wykorzystania modeli dla zmiennych wielomianowych uporządkowanych. Stwierdzono, że „zwykły” uporządkowany model logitowy nie powinien być użyty do analizy sytuacji w rozpatrywanej próbie. Dlatego też zastosowano metodę częściowo uogólnionych uporządkowanych modeli logitowych dopuszczających

zależność części parametrów strukturalnych modelu od przynależności do kategorii zmiennej porządkowej określającej stopień satysfakcji z osiągniętych dochodów.

Literatura

- D'Ambrosio C., Frick J., *Income satisfaction and relative deprivation: An empirical link*, "Social Indicators Research" 2007, vol. 81, nr 3, s. 497-519.
- Dudek H., *Subiektywne skale ekwiwalentności – analiza na podstawie danych o satysfakcji z osiągniętych dochodów*, [w:] K. Jajuga i M. Walesiak (red.), *Klasyfikacja i analiza danych*, Taksonomia 14, Wydawnictwo AE we Wrocławiu, Wrocław 2012, s. 153-162.
- Ferrer-i-Carbonell A., *Income and well-being: An empirical analysis of the comparison income effect*, "Journal of Public Economics" 2005, vol. 89, s. 997-1019.
- Greene W.H., Hensher D.A., *Modeling Ordered Choices: a Primer*, Cambridge University Press, Cambridge 2010.
- Książek M., *Modele zmiennych wielomianowych uporządkowanych*, [w:] M. Gruszczyński (red.), *Mikroekonometria*, Wolters Kluwer Polska, Warszawa 2010, s. 103-152.
- Long J.S., *Regression Models for Categorical and Limited Dependent Variables*, Sage Publications, Thousand Oaks, 1997.
- Panek T., *Ubóstwo, wykluczenie społeczne i nierówności*, Oficyna Wydawnicza Szkoły Głównej Handlowej, Warszawa 2011.
- Peterson B., Harrell F.E. Jr., *Partial proportional odds models for ordinal response variables*, „Journal of the Royal Statistical Society. Series C” 1990, vol. 39, s. 205-217.
- Vera-Toscano E., Ateca-Amestoy V., Serrano-Del-Rosal R., *Building financial satisfaction*, "Social Indicators Research" 2006, vol. 77, no. 2, s. 211-243.
- Williams R., *Generalized ordered logit/partial proportional odds models for ordinal dependent variables*, „The Stata Journal” 2006, vol. 6, no. 1, s. 58-82.

IMPACT OF THE RELATIVE DEPRIVATION ON SUBJECTIVE INCOME SATISFACTION

Summary: The article presents an analysis of income satisfaction of households of retirees. For this purpose, the method of partial generalized ordered logit models was used. It was found that relative deprivation, education, age, gender, place of residence, the fact of being in a relationship (formal or not) and additional source of income from the use of agricultural farm were determinants of income satisfaction. The empirical investigation was based on data from household budget surveys carried out by Central Statistical Office in 2009.

Keywords: relative deprivation, incomes of households, logit models.