

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

RESEARCH PAPERS

of Wrocław University of Economics

Nr 328

Taksonomia 23

**Klasyfikacja i analiza danych –
teoria i zastosowania**

Redaktorzy naukowci

Krzysztof Jajuga, Marek Walesiak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2014

Redaktor Wydawnictwa: Barbara Majewska

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Beata Mazur

Projekt okładki: Beata Dębska

Publikacja jest dostępna w Internecie na stronach:

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

w Dolnośląskiej Bibliotece Cyfrowej www.dbc.wroc.pl,

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 2014

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

ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

| | |
|---|-----|
| Wstęp | 11 |
| Małgorzata Rószkiewicz , Wykorzystanie metaanalizy w budowaniu modelu pomiarowego w przypadku braku niezmienniczości zasad pomiaru na przykładzie pomiaru zadowolenia z życia..... | 13 |
| Elżbieta Sobczak , Harmonijność inteligentnego rozwoju regionów Unii Europejskiej | 21 |
| Ewa Roszkowska, Renata Karwowska , Analiza porównawcza województw Polski ze względu na poziom zrównoważonego rozwoju w roku 2010..... | 30 |
| Tadeusz Kufel, Magdalena Osińska, Marcin Błażejowski, Paweł Kufel , Analiza porównawcza wybranych filtrów w analizie synchronizacji cyklu koniunkturalnego..... | 41 |
| Marcin Salamaga , Próba konstrukcji tablic „wymierania scenicznego” spektakli operowych na przykładzie Metropolitan Opera..... | 51 |
| Iwona Foryś , Wykorzystanie analizy dyskryminacyjnej do typowania rynków podobnych w procesie wyceny nieruchomości niemieszkalnych | 59 |
| Jerzy Korzeniewski , Selekcja zmiennych w klasyfikacji – propozycja algorytmu | 69 |
| Sabina Denkowska , Testowanie wielokrotne przy weryfikacji wieloczynnikowych modeli proporcjonalnego hazardu Coxa..... | 76 |
| Ewa Chodakowska , Teoria równań strukturalnych w klasyfikacji zmiennych jawnych i ukrytych według charakteru ich wzajemnych oddziaływań | 85 |
| Iwona Konarzewska , Model PCA dla rynku akcji – studium przypadku | 94 |
| Katarzyna Wójcik, Janusz Tuchowski , Dobór optymalnego zestawu słów istotnych w opiniach konsumentów na potrzeby ich automatycznej analizy | 106 |
| Aleksandra Łuczak , Zastosowanie metody AHP-LP do oceny ważności determinant rozwoju społeczno-gospodarczego w jednostkach administracyjnych | 116 |
| Aleksandra Witkowska, Marek Witkowski , Klasyfikacja pozycyjna banków spółdzielczych według stanu ich kondycji finansowej w ujęciu dynamicznym | 126 |
| Adam Depta , Zastosowanie analizy korespondencji do oceny jakości życia ludności na podstawie kwestionariusza SF-36v2 | 135 |
| Marek Lubicz, Maciej Zięba, Konrad Pawelczyk, Adam Rzechonek, Marek Marciniak, Jerzy Kołodziej , Indukcja reguł dla danych niekompletnych i niezbalansowanych: modele klasyfikatorów i próba ich zastosowania do predykcji ryzyka operacyjnego w torakochirurgii | 146 |

| | |
|---|-----|
| Małgorzata Misztal , Wybrane metody oceny jakości klasyfikatorów – przegląd i przykłady zastosowań..... | 156 |
| Anna M. Olszewska , Wykorzystanie wybranych metod taksonomicznych do oceny potencjału innowacyjnego województw | 167 |
| Iwona Bąk , Porównanie jakości grupowań powiatów województwa zachodniopomorskiego pod względem atrakcyjności turystycznej..... | 177 |
| Agnieszka Kozera, Joanna Stanisławska, Romana Głowicka-Wołoszyn , Segmentacja gospodarstw domowych według wydatków na turystykę zorganizowaną..... | 186 |
| Agnieszka Wałęga , Podejście syntetyczne w analizie spójności ekonomicznej gospodarstw domowych..... | 196 |
| Joanna Banaś, Małgorzata Machowska-Szewczyk, Bożena Mroczek , Zastosowanie analizy korespondencji do badania wpływu elektrowni wiatrowych na jakość życia ludności | 205 |
| Joanna Banaś, Krzysztof Małecki , Klasyfikacja punktów pomiarów ankietowych kierowców na granicy Szczecina z wykorzystaniem zmiennych symbolicznych..... | 214 |
| Aneta Becker , Wykorzystanie informacji granularnej w analizie wymagań rynku pracy..... | 222 |
| Katarzyna Cheba, Joanna Holub-Iwan , Wykorzystanie analizy korespondencji w segmentacji rynku usług medycznych..... | 230 |
| Adam Depta, Iwona Staniec , Identyfikacja czynników decydujących o jakości życia studentów łódzkich uczelni..... | 238 |
| Katarzyna Dębowska, Jarosław Kilon , Reguły asocjacyjne w analizie wyników badań metodą Delphi..... | 247 |
| Anna Domagała , O wykorzystaniu analizy głównych składowych w metodzie <i>Data Envelopment Analysis</i> | 254 |
| Alicja Grześkowiak , Analiza wykluczenia cyfrowego w Polsce w ujęciu indywidualnym i regionalnym..... | 264 |
| Anna M. Olszewska, Anna Gryko-Nikitin , Pomiar postrzegania jakości kształcenia uczelni wyższej na danych porządkowych z wykorzystaniem środowiska R..... | 273 |
| Karolina Paradysz , Hierarchiczna metoda grupowania powiatów jako podejście benchmarkowe w ocenie bezrobocia według BAEL-u w wybranych typach małych obszarów | 282 |
| Radosław Pietrzyk , Porównanie metod pomiaru efektywności zarządzania portfelami funduszy inwestycyjnych..... | 290 |
| Agnieszka Przedborska, Małgorzata Misztal , Wybrane metody statystyki wielowymiarowej w ocenie skuteczności terapeutycznej głębokiej stymulacji elektromagnetycznej u pacjentów z chorobą zwyrodnieniową stawów..... | 299 |

| | |
|---|-----|
| Wojciech Roszka, Marcin Szymkowiak , Podejście kalibracyjne w statystycznej integracji danych | 308 |
| Iwona Skrodzka , Zastosowanie wybranych metod klasyfikacji do analizy kapitału ludzkiego krajów Unii Europejskiej | 316 |
| Agnieszka Stanimir , Wielowymiarowa analiza czynników sprzyjających włączeniu społecznemu | 326 |
| Dorota Strózik, Tomasz Strózik , Przestrzenne zróżnicowanie poziomu życia w województwie wielkopolskim..... | 334 |
| Izabela Szamrej-Baran , Identyfikacja przyczyn ubóstwa energetycznego w Polsce przy wykorzystaniu modelowania miękkiego..... | 343 |
| Janusz Tuchowski, Katarzyna Wójcik , Klasyfikacja obiektów w systemie Krajowych Ram Kwalifikacji opisanych za pomocą ontologii | 353 |
| Aleksandra Matuszewska-Janica , Grupowanie krajów Unii Europejskiej ze względu na poziom feminizacji sektorów gospodarczych | 361 |
| Monika Rozkrut, Dominik Rozkrut , Identyfikacja strategii innowacyjnych przedsiębiorstw usługowych w Polsce | 369 |

Summaries

| | |
|---|-----|
| Małgorzata Rószkiewicz , The use of meta-analysis in building the measurement model in case of the absence of measurement invariance on the example of measuring of life satisfaction..... | 20 |
| Elżbieta Sobczak , Harmonious smart growth of European Union regions..... | 29 |
| Ewa Roszkowska, Renata Karwowska , The comparative analysis of Polish voivodeships with respect to sustainable development in 2010..... | 40 |
| Tadeusz Kufel, Magdalena Osińska, Marcin Błażejowski, Paweł Kufel , Comparative analysis of chosen filters in business cycles analysis | 50 |
| Marcin Salamaga , The attempt of construction of the life tables for opera works on the example of the Metropolitan Opera | 58 |
| Iwona Foryś , Using discriminant analysis to select similar markets in non-residential property valuation process..... | 68 |
| Jerzy Korzeniewski , Variable selection in classification – algorithm proposal | 75 |
| Sabina Denkowska , Multiple testing in the verification process of multifactorial Cox proportional hazards models | 84 |
| Ewa Chodakowska , The theory of structural equations modelling in the classification of observed variables and latent constructs according to the character of their relationship..... | 93 |
| Iwona Konarzewska , Modelling stock market by PCA factor model – case study | 105 |

| | |
|--|-----|
| Katarzyna Wójcik, Janusz Tuchowski , Selection of the optimal set of relevant words in consumers opinions in the context of the opinion mining .. | 115 |
| Aleksandra Łuczak , Application of AHP-LP to the evaluation of importance of determinants of socio-economic development in the administrative units | 125 |
| Aleksandra Witkowska, Marek Witkowski , A dynamic approach to the ranking of cooperative banks by their financial condition | 134 |
| Adam Depta , Application of correspondence analysis for the measurement of quality of life – questionnaire SF-36v2 based research | 145 |
| Marek Lubicz, Maciej Zięba, Konrad Pawelczyk, Adam Rzechonek, Marek Marciniak, Jerzy Kołodziej , Classification rules extraction for missing and imbalance data: models of classifiers and initial results in the rules-based thoracic surgery risk prediction..... | 155 |
| Małgorzata Misztal , Selected methods for assessing the performance of classifiers – an overview and examples of applications..... | 166 |
| Anna M. Olszewska , The application of selected quantitative methods to the evaluation of voivodeship innovation level potential..... | 176 |
| Iwona Bąk , The comparison of the quality of groupings of poviats of West Pomeranian Voivodeship in terms of tourism attractiveness | 185 |
| Agnieszka Kozera, Joanna Stanisławska, Romana Głowicka-Wołoszyn , Household segmentation with respect to the expenditure on organized tourism..... | 195 |
| Agnieszka Wałęga , Synthetic approach in the analysis of economic coherence of households | 204 |
| Joanna Banaś, Małgorzata Machowska-Szewczyk, Bożena Mroczek , Using the correspondence analysis to examine the impact of wind turbines on the quality of life..... | 213 |
| Joanna Banaś, Krzysztof Małecki , Classification of measurement survey points of drivers on the boundary of Szczecin using symbolic variables... | 221 |
| Aneta Becker , The use granular information in the analysis of the requirements of the labor market..... | 229 |
| Katarzyna Cheba, Joanna Hołub-Iwan , The application of the correspondence analysis of patients segmentation on the medical service market | 237 |
| Adam Depta, Iwona Staniec , Identification of the factors that determine the quality of students life at universities in Lodz..... | 246 |
| Katarzyna Dębkowska, Jarosław Kilon , Association rules in the analysis of research results the Delphi method | 253 |
| Anna Domagała , About using Principal Component Analysis in Data Envelopment Analysis | 263 |
| Alicja Grześkowiak , Analysis of the digital divide in Poland at the individual and regional level | 272 |

| | |
|---|-----|
| Anna M. Olszewska, Anna Gryko-Nikitin , Assessment of perception of quality of teaching at an institution of higher learning based on the ordinal data with the utilization of R environment..... | 281 |
| Karolina Paradysz , The hierarchical method of grouping poviats as a benchmark approach in the assessment of unemployment by BAEL in selected types of small areas | 289 |
| Radosław Pietrzyk , Comparison of methods of measuring the performance of investment funds portfolios..... | 298 |
| Agnieszka Przedborska, Małgorzata Misztal , Selected multivariate statistical analysis methods in the evaluation of efficacy of deep electromagnetic stimulation in patients with degenerative joint disease | 307 |
| Wojciech Roszka, Marcin Szymkowiak , A calibration approach in statistical data integration | 315 |
| Iwona Skrodzka , Application of some methods of classification to the analysis of human capital in the European Union..... | 325 |
| Agnieszka Stanimir , Multivariate analysis of social inclusion factors..... | 333 |
| Dorota Strózik, Tomasz Strózik , Spatial differentiation of the standard of living in Great Poland Voivodeship | 342 |
| Izabela Szamrej-Baran , Identification of fuel poverty causes in Poland using soft modelling | 352 |
| Janusz Tuchowski, Katarzyna Wójcik , Classification of objects in the National Classification Framework described by the ontology..... | 360 |
| Aleksandra Matuszewska-Janica , Clustering of European Union states taking into consideration the levels of feminization of economic sectors.. | 368 |
| Monika Rozkrut, Dominik Rozkrut , Identification of service sector innovation strategies in Poland..... | 379 |

Iwona Skrodzka

Uniwersytet w Białymstoku

ZASTOSOWANIE WYBRANYCH METOD KLASYFIKACJI DO ANALIZY KAPITAŁU LUDZKIEGO KRAJÓW UNII EUROPEJSKIEJ

Streszczenie: Celem artykułu jest analiza porównawcza wyników badań dotyczących kapitału ludzkiego krajów Unii Europejskiej, uzyskanych za pomocą dwóch metod: modelowania miękkiego oraz metody TOPSIS. Modelowanie miękkie umożliwia badanie powiązań między zmiennymi, które nie są bezpośrednio obserwowalne (tzw. zmienne ukryte), oraz oszacowanie wartości tych zmiennych. Oszacowane wartości mogą posłużyć do liniowego uporządkowania badanych obiektów. Metoda TOPSIS polega na wyznaczeniu odległości każdego obiektu wielocechowego od wzorca i antywzorca rozwoju, a następnie liniowym uporządkowaniu obiektów. W badaniach zostały wykorzystane dane statystyczne za rok 2010, dotyczące 27 krajów Unii Europejskiej.

Słowa kluczowe: kapitał ludzki, rozwój gospodarczy, modelowanie miękkie, metoda TOPSIS.

1. Wstęp

Znaczenie kapitału ludzkiego dla procesów wzrostu i rozwoju gospodarczego jest obecnie powszechnie uznawane przez wielu ekonomistów [Cichy, Małaga 2007, s. 20-49; Florczak 2007, s. 126-166]. Określenie jego stanu oraz perspektyw rozwoju stanowi istotny problem zarówno w teorii ekonomii, jak i praktyce gospodarczej. Kapitał ludzki to kategoria złożona. W szerokim ujęciu rozumiany jest jako zakumulowany zasób wiedzy, wykształcenia, zdolności, umiejętności, doświadczenia zawodowego, kwalifikacji, zdrowia zawarty w społeczeństwie [Domański 1993, s. 16-19; Marciniak 2000, s. 157-158; Florczak 2007, s. 112]. Wielowymiarowość kapitału ludzkiego determinuje wybór metody pomiaru. W badaniach empirycznych najczęściej stosuje się jedną z metod wielowymiarowej analizy statystycznej.

Celem artykułu jest analiza porównawcza wyników badań nad kapitałem ludzkim krajów Unii Europejskiej, uzyskanych za pomocą dwóch metod: modelowania miękkiego oraz metody TOPSIS.

2. Metoda modelowania miękkiego

Twórcą modelowania miękkiego jest H. Wold [1980]. W literaturze polskiej szczegółowy opis metody oraz jej uogólnienie można znaleźć w pracy J. Rogowskiego [1990], a przykłady zastosowań w pracach D. Mierzyńskiej [1999], D. Perło [2004] oraz I. Skrodzkiej [2012].

Model miękki umożliwia badanie powiązań między zmiennymi, które nie są bezpośrednio obserwowalne, tzw. zmienne ukryte. Składa się z dwóch podmodeli: wewnętrznego, definiującego relacje między zmiennymi ukrytymi, oraz zewnętrznego, definiującego zmienne ukryte za pomocą zbioru zmiennych obserwowalnych, zwanych indykatorami. Indykatory pozwalają na pośrednie obserwowanie zmiennych ukrytych i są dobierane na podstawie określonej teorii lub intuicji badacza. W modelowaniu miękkim zmienna ukryta może zostać zdefiniowana za pomocą indykatorów w oparciu o podejście indukcyjne (indykatory tworzą zmienne ukryte) lub w oparciu o podejście dedukcyjne (indykatory odzwierciedlają swoje pojęcia teoretyczne). W zależności od sposobu definiowania indykatory powinny wykazywać inne własności statystyczne. W podejściu dedukcyjnym powinny charakteryzować się wysoką korelacją między sobą, zaś w podejściu indukcyjnym powinny być nieskorelowane.

Estymacja parametrów modelu miękkiego odbywa się za pomocą częściowej metody najmniejszych kwadratów (*Partial Least Squares*). W wyniku zastosowania tej metody otrzymuje się oszacowania wartości zmiennych ukrytych. Zależą one nie tylko od relacji zewnętrznych, ale również od założonych w modelu wewnętrznym związków między zmiennymi ukrytymi. Oznacza to, że proces poznania uzależniony jest nie tylko od definicji danego pojęcia, ale również od opisu teoretycznego. Między innymi to właśnie wyróżnia prezentowaną metodę wśród większości powszechnie stosowanych metod wielowymiarowej analizy porównawczej¹.

Oceny jakości modelu można dokonać, posługując się współczynnikami determinacji, które są wyznaczane dla każdego równania modelu. Istotność parametrów modelu bada się w oparciu o odchylenia standardowe wyznaczone metodą cięć Tukeya². Jakość prognostyczną z kolei ocenia się za pomocą testu Stone'a-Geissera³ (w skrócie S-G).

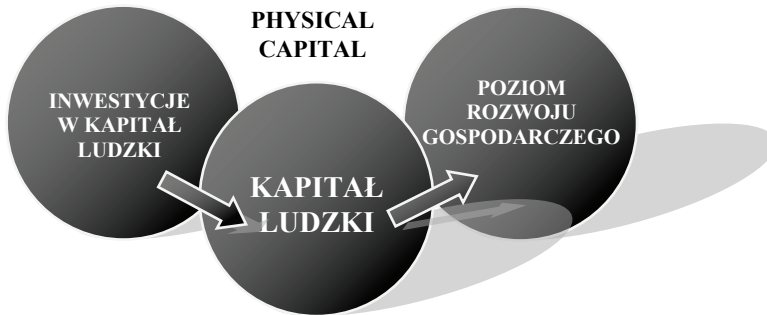
¹ Taką cechą mają również modele równań strukturalnych, których opis można znaleźć w [Konnarski 2009].

² Jeżeli wartość podwojonego odchylenia standardowego jest mniejsza niż wartość bezwzględna estymatora parametru, wówczas parametr jest istotny statystycznie (tzw. reguła „2s”).

³ Test został zaproponowany przez H. Wolda. Statystyki testu przyjmują wartości z przedziału $(-\infty, 1>$. Dla modelu idealnego wartość testu wynosi 1. Ujemne wartości wskazują na słabe własności modelu.

Etap 1. Specyfikacja modelu wewnętrznego – określenie zależności zachodzących między zmiennymi ukrytymi.

Koncepcję modelu wewnętrznego, który został wykorzystany w badaniach, prezentuje rysunek 1.



Rys. 1. Koncepcja modelu miękkiego

Źródło: opracowanie własne.

Model składa się z dwóch następujących równań:

$$KL_t = \alpha_1 IKL_{t-2} + \alpha_2 IKL_{t-1} + \alpha_3 IKL_t + \alpha_0 + \varepsilon, \quad (1)$$

$$PRG_t = \beta_1 KL_t + \beta_0 + \xi, \quad (2)$$

gdzie: KL – kapitał ludzki,
 IKL – inwestycje w kapitał ludzki,
 PRG – poziom rozwoju gospodarczego,
 $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \beta_0, \beta_1$ – parametry strukturalne,
 ε, ξ – składniki losowe,
 t – 2010⁴.

Etap 2. Specyfikacja modelu zewnętrznego – zdefiniowanie zmiennych ukrytych za pomocą zmiennych obserwowalnych (indykatorów).

Zbiór wszystkich indykatorów, które zostały zakwalifikowane do modelu, prezentuje tabela 1⁵.

Przy doborze zmiennych obserwowalnych kierowano się:

- uniwersalnością – uznana powszechnie waga i znaczenie indykatora,
- porównywalnością – indykatory zostały przedstawione w postaci wskaźników natężenia,
- zróżnicowaniem – współczynniki zmienności były większe od 10% [Hellwig i in. 1997, s. 25; Nowak 1990, s. 24-28].

⁴ Wybór okresu badania został podyktowany dostępnością danych statystycznych.

⁵ Dane pochodziły z baz statystycznych udostępnianych przez Eurostat oraz World Bank.

Tabela 1. Indykatory zmiennych ukrytych w modelu miękkim

| Zmienna ukryta | Indykator | Znaczenie indykatora | Rodzaj indykatora ⁶ | Źródło danych ⁷ |
|----------------|-----------|--|--------------------------------|----------------------------|
| KL | KL01 | Odsetek osób z wykształceniem wyższym. | S | E |
| | KL02 | Odsetek pracujących z wykształceniem wyższym. | S | E |
| | KL03 | Odsetek osób w wieku 25-64 lata uczestniczących w kształceniu ustawicznym. | S | E |
| | KL04 | Zasoby ludzkie w nauce i technologii na 100 tys. mieszkańców. | S | E |
| | KL05 | Liczba pracowników naukowo-badawczych w sektorze B+R na 1 mln mieszkańców. | S | WB |
| | KL06 | Liczba patentów na 1 mln mieszkańców. | S | WB |
| | KL07 | Odsetek osób oceniających swój stan zdrowia jako bardzo dobry. | S | E |
| | KL08 | Oczekiwana długość życia. | S | WB |
| | KL09 | Stopa śmiertelności noworodków. | D | E |
| IKL | IKL01 | Publiczne nakłady na edukację jako % PKB. | S | E |
| | IKL02 | Publiczne nakłady na edukację na 1 mieszkańca (PPS). | S | E |
| | IKL03 | Nakłady na zdrowie jako % PKB. | S | WB |
| | IKL04 | Nakłady na zdrowie na 1 mieszkańca (PPS). | S | WB |
| | IKL05 | Nakłady na B+R jako % PKB. | S | E |
| | IKL06 | Nakłady na B+R na 1 mieszkańca (PPS). | S | E |
| PRG | PRG01 | Produkt krajowy brutto na 1 mieszkańca (PPS). | S | WB |
| | PRG02 | Wartość dodana brutto na 1 pracującego (euro). | S | E |
| | PRG03 | Udział rolnictwa w tworzeniu wartości dodanej brutto. | D | E |
| | PRG04 | Udział sektora usług w tworzeniu wartości dodanej brutto. | S | E |
| | PEG05 | Stopa bezrobocia. | D | WB |

Źródło: opracowanie własne.

Etap 3. Estymacja parametrów modelu⁸.

Oszacowania relacji modelu wewnętrznego prezentują równania (3) oraz (4). Wyniki estymacji modelu zewnętrznego można znaleźć w [Skrodzka 2013].

$$\hat{KL}_t = 0,5612IKL_{t-2} + 0,2580IKL_{t-1} + 0,0922IKL_t + 5,0458 \quad (3)$$

(0,0196) (0,0317) (0,0456) (0,102)

$$\hat{PRG}_t = 0,7678KL_t - 4,6541 \quad (4)$$

(0,0291) (0,3532)

⁶ S – stymulanta, D – destymulanta.

⁷ WB – World Bank, E – Eurostat.

⁸ Do estymacji modelu wykorzystano program PLS, opracowany przez dr. hab. J. Rogowskiego, prof. UwB.

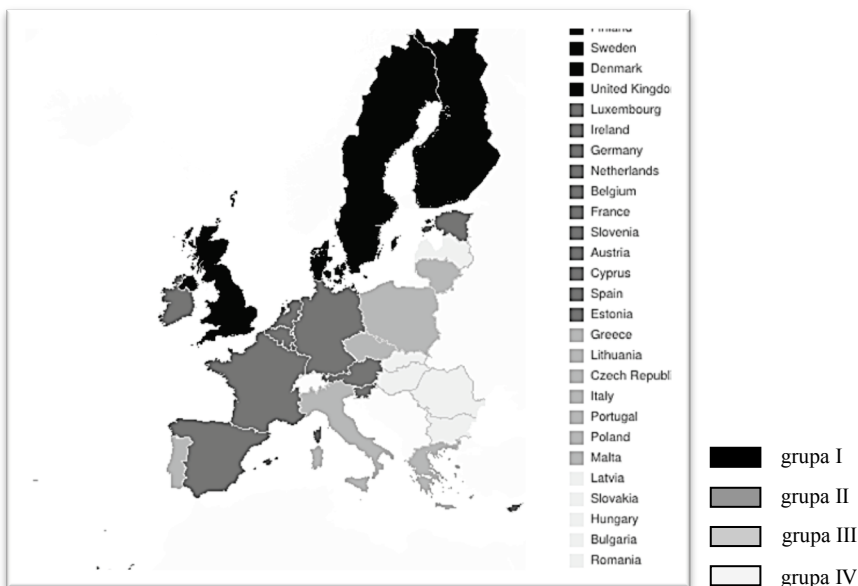
Etap 4. Weryfikacja statystyczna i merytoryczna modelu.

Wyniki estymacji są zgodne z oczekiwaniami. Uzyskano dodatnią zależność między kapitałem ludzkim a poziomem rozwoju gospodarczego oraz dodatnie zależności między inwestycjami w kapitał ludzki a zasobem kapitału ludzkiego. Wszystkie parametry modelu okazały się istotne statystycznie. Współczynnik determinacji dla równania (3) przyjął wartość 0,8, zaś dla równania (4) – wartość 0,6. Wartość testu S-G była dodatnia i kształtowała się na poziomie 0,36. Model zweryfikowano pozytywnie pod względem merytorycznym i statystycznym.

Etap 5. Uporządkowanie liniowe i klasyfikacja obiektów ze względu na wartość miernika syntetycznego, uzyskanego w wyniku modelowania.

Rysunek 2 przedstawia podział krajów Unii Europejskiej na cztery grupy typologiczne. Granice przedziałów zostały wyznaczone w oparciu o obliczone wartości: średniej arytmetycznej (\bar{z}) oraz odchylenia standardowego (s_z) miernika syntetycznego (z_i), uzyskanego w wyniku zastosowania modelowania miękkiego:

- grupa I (bardzo wysoki zasób kapitału ludzkiego): $z_i \geq \bar{z} + s_z$,
- grupa II (wysoki zasób kapitału ludzkiego): $\bar{z} \leq z_i < \bar{z} + s_z$,
- grupa III (średni i niski zasób kapitału ludzkiego): $\bar{z} - s_z \leq z_i < \bar{z}$,
- grupa IV (niski zasób kapitału ludzkiego): $z_i < \bar{z} - s_z$ [Nowak 1990, s. 92-93].



Rys. 2. Podział na grupy typologiczne krajów UE-27 ze względu na zasób kapitału ludzkiego w 2010 r. uzyskany metodą modelowania miękkiego

Źródło: opracowanie własne.

W grupie krajów o najwyższych zasobach kapitału ludzkiego znalazły się cztery kraje: Finlandia, Szwecja, Dania oraz Wielka Brytania. Najniższe zasoby kapitału ludzkiego były zlokalizowane na Łotwie, Słowacji, Węgrzech, w Bułgarii oraz Rumunii. Polska zajęła 21. pozycję w rankingu krajów UE-27.

3. Metoda TOPSIS

Metoda TOPSIS (*Technique for Order Preference by Similarity to an Ideal Solution*) należy do klasycznych metod porządkowania liniowego obiektów wielocechowych. W ramach metody wyznacza się odległości każdego obiektu wielocechowego od wzorca i antywzorca, a następnie dokonuje się liniowego uporządkowania obiektów. Po raz pierwszy metoda została zaprezentowana przez C.L. Hwang i K. Yoon [Hwang, Yoon 1981]. Można ją potraktować jako modyfikację metody taksonomicznej wzorca rozwoju Hellwiga [1968].

Etap 1. Wybór wskaźników (cech prostych) na podstawie przesłanek merytorycznych oraz statystycznych.

Wykorzystano zestaw wskaźników, który został zastosowany do zdefiniowania kapitału ludzkiego w modelu miękkim (zob. tabela 1).

Etap 2. Normalizacja wartości wskaźników

Zastosowano unitaryzację zerowaną zgodnie z poniższymi wzorami:

$$\begin{array}{ll} \text{dla stymulant} & \text{dla destymulant} \\ z_{ik} = \frac{x_{ik} - \min_i \{x_{ik}\}}{\max_i \{x_{ik}\} - \min_i \{x_{ik}\}}, & z_{ik} = \frac{\max_i \{x_{ik}\} - x_{ik}}{\max_i \{x_{ik}\} - \min_i \{x_{ik}\}}, \end{array} \quad (5)$$

gdzie i oznacza numer kraju ($i = 1, 2, \dots, n$), zaś k – numer wskaźnika ($k = 1, 2, \dots, m$).

Etap 3. Obliczenie odległości euklidesowej od wzorca $z_k^+ = [1, 1, \dots, 1]$ oraz antywzorca $z_k^- = [0, 0, \dots, 0]$, zgodnie ze wzorami

$$d_i^+ = \sqrt{\sum_{k=1}^m (z_{ik} - z_k^+)^2}, \quad d_i^- = \sqrt{\sum_{k=1}^m (z_{ik} - z_k^-)^2}, \quad (i = 1, 2, \dots, n = 27). \quad (6)$$

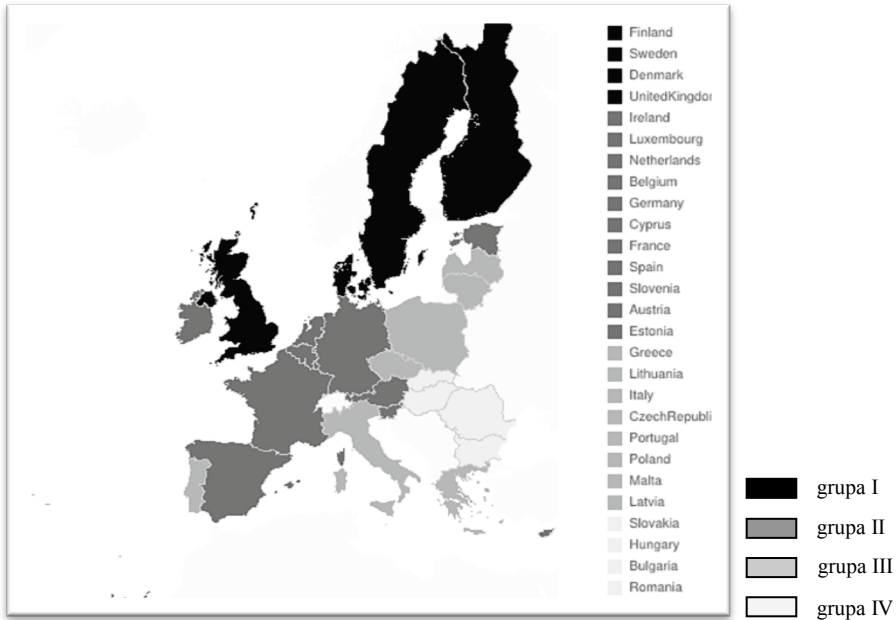
Etap 4. Wyznaczenie wartości syntetycznego miernika zgodnie ze wzorem:

$$q_i = \frac{d_i^-}{d_i^- + d_i^+} \quad (i = 1, 2, \dots, n = 27), \quad (7)$$

przy tym $0 \leq q_i \leq 1$. Wyższe wartości miernika świadczą o wyższej pozycji w rankingu i -tego obiektu.

Etap 5. Uporządkowanie liniowe i klasyfikacja obiektów ze względu na wartość syntetycznego miernika.

Rysunek 3 przedstawia podział krajów UE-27 na cztery grupy typologiczne. Grupy zostały wyodrębnione w oparciu o wartość średnią i odchylenie standardowe zmiennej syntetycznej.



Rys. 3. Podział na grupy typologiczne krajów UE-27 ze względu na zasób kapitału ludzkiego w 2010 r., uzyskany metodą TOPSIS

Źródło: opracowanie własne.

4. Porównanie wyników

Uporządkowanie krajów Unii Europejskiej uzyskane za pomocą modelowania miękkiego (MM) oraz metody TOPSIS prezentuje tabela 2. Zestawienie wskazuje na dużą zgodność otrzymanych rankingów. Potwierdza to również wartość współczynnika rang Spearmana, która kształtuje się na poziomie 0,99.

W tabeli 3 przedstawiono wynik klasyfikacji krajów UE-27. Uzyskany podział na grupy typologiczne również wykazuje bardzo dużą zgodność. Jedynym krajem, który w zależności od zastosowanej metody został inaczej sklasyfikowany, była Łotwa.

Tabela 2. Uporządkowanie krajów UE-27 ze względu na zasoby kapitału ludzkiego w 2010 roku – porównanie wyników

| Kraj | MM | TOPSIS | | Kraj | MM | TOPSIS |
|------------|----|--------|--|------------------|----|--------|
| Austria | 12 | 14 | | Łotwa | 23 | 23 |
| Belgia | 9 | 8 | | Malta | 22 | 22 |
| Bułgaria | 26 | 26 | | Niemcy | 7 | 9 |
| Cypr | 13 | 10 | | Polska | 21 | 21 |
| Dania | 3 | 3 | | Portugalia | 20 | 20 |
| Estonia | 15 | 15 | | Republika Czeska | 18 | 19 |
| Finlandia | 1 | 1 | | Rumunia | 27 | 27 |
| Francja | 10 | 11 | | Słowacja | 24 | 24 |
| Grecja | 16 | 16 | | Słowenia | 11 | 13 |
| Hiszpania | 14 | 12 | | Szwecja | 2 | 2 |
| Holandia | 8 | 7 | | Węgry | 25 | 25 |
| Irlandia | 6 | 5 | | Włochy | 19 | 18 |
| Litwa | 17 | 17 | | Wielka Brytania | 4 | 4 |
| Luksemburg | 5 | 6 | | | | |

Źródło: opracowanie własne.

Tabela 3. Podział krajów UE-27 na grupy typologiczne – porównanie wyników

| TOPSIS MM | Grupa I | Grupa II | Grupa III | Grupa IV |
|-----------|--|--|---|---|
| Grupa I | Finlandia, Szwecja, Dania, Wielka Brytania | | | |
| Grupa II | | Luksemburg, Irlandia, Niemcy, Holandia, Belgia, Francja, Słowenia, Austria, Cypr, Hiszpania, Estonia | | |
| Grupa III | | | Grecja, Litwa, Republika Czeska, Włochy, Portugalia, Polska, Malta | |
| Grupa IV | | | Łotwa | Słowacja, Węgry, Bułgaria, Rumunia |

Źródło: opracowanie własne.

Badania dotyczące zróżnicowania krajów Unii Europejskiej pod względem kapitału ludzkiego były prowadzone przez różnych autorów [por. Pawlas 2009; Stec i in. 2005; Wronowska 2009]. Trudno jednak dokonać porównania uzyskanych rezultatów ze względu na różne sposoby definiowania kapitału ludzkiego przez autorów, dobór różnych wskaźników kapitału ludzkiego oraz różnych okresów badawczych.

5. Podsumowanie

Artykuł miał na celu dokonanie analizy porównawczej wyników badań dotyczących kapitału ludzkiego krajów Unii Europejskiej, uzyskanych za pomocą dwóch metod: modelowania miękkiego oraz metody TOPSIS. Otrzymane rezultaty wskazały na to, iż obie metody przynoszą zbliżone uporządkowania krajów UE-27 pod względem zasobów kapitału ludzkiego oraz zbliżony podział krajów na grupy typologiczne. Wybór metody nie miał zatem większego wpływu na uzyskiwane rezultaty. W związku z tym decyzja, jaką metodę zastosować, powinna zależeć od celów szczegółowych badań. Jeżeli badania mają na celu dokonanie uporządkowania i grupowanie krajów, wygodniejsza i prostsza do zastosowania wydaje się metoda TOPSIS. Jeśli zaś jednym z celów badań jest analiza powiązań między zmienną a innymi kategoriami, użytecznym narzędziem jest wówczas modelowanie miękkie.

Literatura

- Cichy K., Malaga K. (2007), *Kapitał ludzki w modelach i teorii wzrostu gospodarczego*, [w:] M. Herbst (red.), *Kapitał ludzki i kapitał społeczny a rozwój regionalny*, Wydawnictwo Naukowe „Scholar”, Warszawa.
- Domański S.R. (1993), *Kapitał ludzki i wzrost gospodarczy*, PWN, Warszawa.
- Florczak W. (2007), *Kapitał ludzki a rozwój gospodarczy*, [w:] W. Welfe (red.), *Gospodarka oparta na wiedzy*, Polskie Wydawnictwo Ekonomiczne, Warszawa.
- Hellwig Z. (1968), *Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom rozwoju i strukturę wykwalifikowanych kadr*, „Przegląd Statystyczny” z. 4.
- Hellwig Z., Siedlecka U., Siedlecki J. (1997), *Taksonometryczne modele zmian struktury gospodarczej Polski*, IRiSS, Warszawa.
- Hwang C.L., Yoon K. (1981), *Multiple attribute decision making-methods and applications*, Springer-Verlag, Berlin – Heidelberg – New York.
- Konarski J. (2009), *Modele równań strukturalnych. Teoria i praktyka*, Wydawnictwo Naukowe PWN, Warszawa.
- Marciniak S. (2000), *Innowacje i rozwój gospodarczy*, Kolegium Nauk Społecznych i Administracji Politechniki Warszawskiej, Warszawa 2000.
- Mierzyńska D. (1999), *Modele miękkie w analizie porównawczej złożonych zjawisk społeczno-ekonomicznych*, rozprawa doktorska, Uniwersytet w Białymstoku, Białystok [maszynopis niepublikowany].

- Nowak E. (1990), *Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych*, PWE, Warszawa.
- Pawlas I. (2009), *Kapitał ludzki w krajach Unii Europejskiej w świetle badań taksonomicznych*, [w:] D. Kopycińska (red.), *Kapitał ludzki jako czynnik przewagi konkurencyjnej*, Katedra Mikroekonomii Uniwersytetu Szczecińskiego, Szczecin, s. 21-31.
- Perło D. (2004), *Źródła finansowania rozwoju regionalnego*, Wydawnictwo Wyższej Szkoły Ekonomicznej w Białymstoku, Białystok.
- Rogowski J. (1990), *Modele miękkie. Teoria i zastosowanie w badaniach ekonomicznych*, Wydawnictwo Filii UW w Białymstoku, Białystok.
- Skrodzka I. (2012), *Zastosowanie modelowania miękkiego do pomiaru kapitału ludzkiego*, rozprawa doktorska, Uniwersytet w Białymstoku, Białystok [maszynopis niepublikowany].
- Skrodzka I. (2013), *Spatial diversity of human capital in the European Union*, „Quantitative Methods in Economics”, Warszawa [w druku].
- Stec M., Janas A., Kuliński A. (2005), *Grupowanie państw Unii Europejskiej ze względu na zasoby kapitału ludzkiego i intelektualnego*, [w:] G. Woźniak (red.), *Nierówności społeczne a wzrost gospodarczy*, Uniwersytet Rzeszowski, Katedra Teorii Ekonomii, Zeszyt nr 6, Wydawnictwo Mitel, Rzeszów, s. 135-146.
- Wold H. (1980), *Soft Modelling: Intermediate between Traditional Model Building and Data Analysis*, Banach Centre Publication 6, Mathematical Statistics 1980.
- Wronowska G. (2009), *Kapitał ludzki w krajach Unii Europejskiej – analiza porównawcza*, [w:] D. Kopycińska (red.), *Kapitał ludzki jako czynnik przewagi konkurencyjnej*, Katedra Mikroekonomii Uniwersytetu Szczecińskiego, Szczecin, s. 32-45.

APPLICATION OF SOME METHODS OF CLASSIFICATION TO THE ANALYSIS OF HUMAN CAPITAL IN THE EUROPEAN UNION

Summary: The aim of this article is a comparative analysis of the results of research on human capital in the European Union achieved by two methods: soft modelling and TOPSIS. Soft modelling enables the study of relationships among unobserved variables (latent variables) and the estimation of values of these variables. The estimated values can be used to a linear ordering of the objects. TOPSIS method involves calculating the distance of each object from a positive ideal solution and a negative ideal solution and then a linear ordering of the objects. The studies used statistics for the year 2010 for the 27 countries of the European Union.

Keywords: human capital, economic development, soft modelling, TOPSIS method.