

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 poziomu 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

Beata Bieszk-Stolorz, Iwona Markowicz

Uniwersytet Szczeciński

WYKORZYSTANIE MODELI PROPORCJONALNEGO I NIEPROPORCJONALNEGO HAZARDU COXA DO BADANIA SZANSY PODJĘCIA PRACY W ZALEŻNOŚCI OD RODZAJU BEZROBOCIA

Streszczenie: Celem artykułu jest analiza szans szybkiego znalezienia zatrudnienia przy wykorzystaniu modelu nieproporcjonalnego hazardu Coxa. W modelu tym zakłada się, że wpływ zmiennej objaśniającej na hazard zależy od czasu. Wstępne badania wskazały na to, że ilorazy hazardu nie zawsze są jednakowe dla bezrobotnych o różnym czasie poszukiwania pracy. Wyniki pozwoliły na określenie wpływu płci, wieku i wykształcenia na szansę szybkiego podjęcia zatrudnienia przez osoby z grup o określonym rodzaju bezrobocia: krótkookresowym (do 3 miesięcy), średniookresowym (od 3 do 6 miesięcy), długookresowym (od 6 do 12 miesięcy) i długotrwałym (12 miesięcy i więcej).

Słowa kluczowe: model hazardu Coxa, badanie proporcjonalności hazardu, bezrobocie.

1. Wstęp

W 2010 r. z Powiatowego Urzędu Pracy w Szczecinie zostało wyrejestrowanych prawie 18 tys. osób bezrobotnych. Osoby te były zróżnicowane pod względem czasu poszukiwania pracy. Podzielono je na grupy według rodzaju bezrobocia. Szanse na podjęcie pracy zostały ocenione na tle innych przyczyn wyrejestrowania.

Celem badawczym artykułu jest analiza wpływu wybranych cech (płeć, wiek, wykształcenie) na szybkie znalezienie zatrudnienia przy wykorzystaniu modelu nieproporcjonalnego hazardu Coxa. W modelu tym zakłada się, że wpływ zmiennej objaśniającej na hazard zależy od czasu, czyli ilorazy hazardu zmieniają się w czasie. Przy zastosowaniu modelu proporcjonalnego hazardu ilorazy hazardu dla dwóch jednostek różniących się wartością analizowanej cechy są uśrednieniem dla całego badanego okresu. Wstępne badania wskazały na to, że ilorazy hazardu nie zawsze są jednakowe dla bezrobotnych o różnym czasie poszukiwania pracy. Wyniki pozwoliły na określenie wpływu płci, wieku i wykształcenia na szansę podjęcia zatrudnienia przez osoby z grup o określonym rodzaju bezrobocia: krótkookresowym (do 3 miesięcy), średniookresowym (od 3 do 6 miesięcy), długookresowym (od 6 do 12 mie-

sięcy) i długotrwałym (12 miesięcy i więcej). Celem poznawczym jest porównanie wyników szacowania parametrów modelu proporcjonalnego i nieproporcjonalnego hazardu.

W artykule przedstawiono wyniki ostatniego etapu badań, prowadzonych w ramach projektu badawczego MNiSW N N111 273538 (lata 2010-2012).

2. Model proporcjonalnego hazardu Coxa

Badania przeprowadzono na 17 793 obserwacjach dotyczących czasu pozostawiania w rejestrze bezrobotnych Powiatowego Urzędu Pracy (PUP) w Szczecinie. Osoby te zostały wyrejestrowane w 2010 r. Liczne powody wyrejestrowania podzielono na dwie grupy: podjęcie pracy i pozostałe (niepodjęcie pracy). Strukturę badanych według tych grup oraz według wybranych cech dotychczasowych bezrobotnych przedstawiono w tab. 1.

Tabela 1. Struktura osób wyrejestrowanych z PUP w Szczecinie w 2010 r. według przyczyny wyrejestrowania (podjęcie lub niepodjęcie pracy) oraz według płci, grupy wieku i wykształcenia

Grupy	Podjęcie pracy	Niepodjęcie pracy
	liczba osób	
Kobiety	3 960	6 385
Mężczyźni	3 475	3 973
W_1 (18-24 lata)	1 166	2 408
W_2 (25-34 lata)	3 008	3 577
W_3 (35-44 lata)	1 398	1 690
W_4 (45-54 lata)	1 342	1 665
W_5 (55-59 lat)	464	802
W_6 (60-64 lata)	57	216
S_1 (co najwyżej gimnazjalne)	1 313	3 890
S_2 (zasadnicze zawodowe)	1 542	2 291
S_3 (średnie ogólnokształcące)	790	1 238
S_4 (średnie zawodowe)	1 516	1 704
S_5 (wyższe)	2 274	1 235
Razem	7 435	10 358

Źródło: obliczenia własne na podstawie danych z PUP w Szczecinie.

W pierwszym etapie analizy wykorzystano model proporcjonalnego hazardu Coxa [Cox, Oakes 1984]:

$$h(t, X) = h_0(t) \exp\left(\sum_{i=1}^p \beta_i X_i\right), \quad (1)$$

gdzie: $X = (X_1, \dots, X_p)$.

W badaniu zmiennymi objaśniającymi są *Płeć*, *Wiek* i *Wykształcenie*. Płeć (P) jest zmienną dychotomiczną zdefiniowaną następująco: kobiety – 1, mężczyźni – 0. Wiek i wykształcenie są zmiennymi kategoryzowanymi i w badaniu przekształcono je na zmienne dychotomiczne. Wiek badanych osób podzielono na 6 przedziałów (W_i), a wykształcenie na 5 poziomów (S_i). Utworzono w ten sposób zmienne sztuczne, kodowane 0-1. Budując modele, z każdej grupy zmiennych wybrano grupę referencyjną. W przypadku płci są to mężczyźni, dla wieku – od 18 do 25 lat (W_1), a dla wykształcenia – co najwyżej gimnazjalne (S_1).

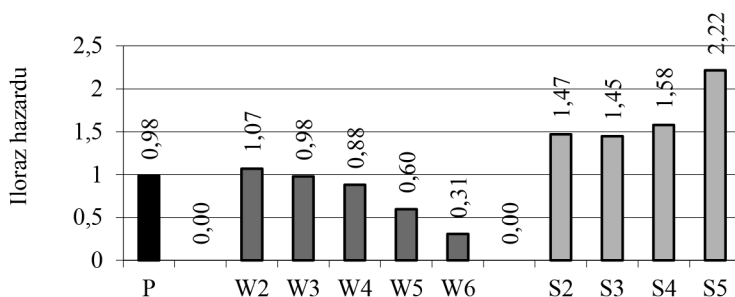
Wyniki estymacji parametrów modeli proporcjonalnego hazardu przedstawiono w tab. 2. Zmiennymi objaśniającymi w modelu 1 jest płeć, w 2 – grupy wieku, a w 3 – wykształcenie. Do zbadania wpływu poszczególnych zmiennych objaśniających na hazard zastosowano trzy odrębne modele. Analogiczne wyniki można otrzymać, budując jeden model uwzględniający wszystkie analizowane zmienne i interakcje między nimi.

Tabela 2. Modele proporcjonalnego hazardu Coxa – wyniki estymacji

Zmienna	Ocena parametru	Błąd standardowy	Statystyka Walda	Poziom p
Model 1				
P	-0,0156	0,0234	0,4472	0,5037
Model 2				
W_2	0,0655	0,0346	3,5841	0,0583
W_3	-0,0198	0,0398	0,2466	0,6195
W_4	-0,1274	0,0404	9,9705	0,0016
W_5	-0,5190	0,0554	87,6868	0,0000
W_6	-1,1817	0,1358	75,7508	0,0000
Model 3				
S_2	0,3860	0,0376	105,6303	0,0000
S_3	0,3703	0,0450	67,6033	0,0000
S_4	0,4565	0,0377	146,3259	0,0000
S_5	0,7963	0,0347	526,4844	0,0000

Źródło: obliczenia własne na podstawie danych z PUP w Szczecinie.

Brak istotności parametru przy zmiennej P (równocześnie całego modelu) świadczy o tym, że szanse szybkiego podjęcia zatrudnienia przez kobiety i mężczyzn w 2010 r. nie różniły się istotnie (wartość ilorazu hazardu bliska 1; rys. 1).



Rys. 1. Ilorazy hazardu – modele proporcjonalnego hazardu Coxa

Źródło: opracowanie własne na podstawie danych z PUP w Szczecinie.

W modelu 2 szanse szybkiego znalezienia pracy przez osoby w wieku od 25 do 34 lat i od 35 do 44 lat były zbliżone do szans osób najmłodszych (brak istotności parametrów przy zmiennych W_2 i W_3). Dla trzech ostatnich grup wieku szanse te były coraz mniejsze w stosunku do grupy referencyjnej. Ilorazy hazardu osób z poszczególnych grup wykształcenia w stosunku do osób z wykształceniem co najwyżej gimnazjalnym istotnie przekraczały wartość 1. Najszybciej (ponad 2 razy) pracę podejmowali bezrobotni z wykształceniem wyższym.

3. Model nieproporcjonalnego hazardu Coxa¹

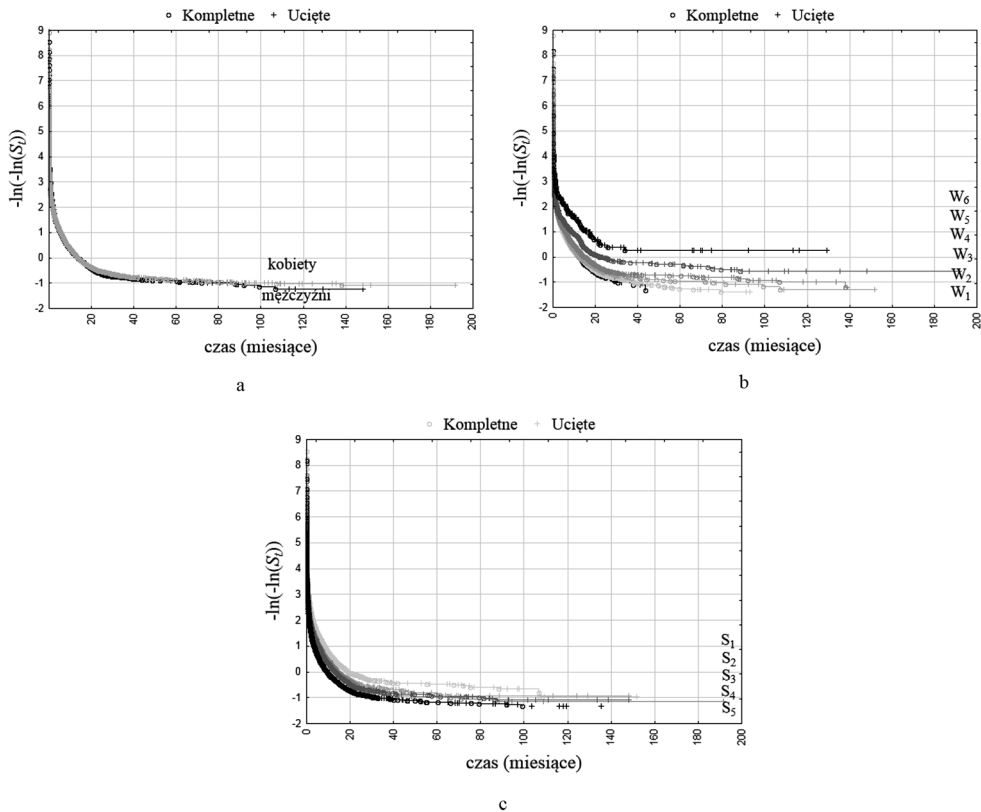
Drugim etapem badań było sprawdzenie zmienności w czasie ilorazów szans na szybkie podjęcie pracy i ich analiza. Do wstępnej oceny braku proporcjonalności hazardu można wykorzystać metody graficzne². Jedną z tych metod opiera się na analizie odległości między przekształconymi krzywymi przeżycia analizowanych grup. W przypadku proporcjonalności przekształcenie wartości S_i do postaci $-\ln(-\ln(S_i))$ powoduje, że krzywe przeżycia³ są do siebie równoległe, a odległość między nimi jest równa parametrowi β_i . Metodę graficzną zilustrowano na rys. 2a-c.

Na podstawie zastosowanej metody graficznej można przypuszczać, że brak proporcjonalności hazardu ma miejsce jedynie w przypadku zmiennej objaśniającej wiek.

¹ Szerzej na temat modeli nieproporcjonalnego hazardu Coxa w pracy [Therneau, Grambsch 2000].

² Szerzej na temat metod graficznych badania proporcjonalności w pracy [Bieszk-Stolorz, Markowicz 2012, s. 37-56].

³ W pracy krzywe przeżycia oszacowano, wykorzystując estymator Kaplana-Meiera [1958].



Rys. 2. Przekształcone estymatory Kaplana-Meiera (prawdopodobieństwo pozostania w rejestrze według płci – a, grup wieku – b i grup wykształcenia – c)

Źródło: opracowanie własne na podstawie danych z PUP w Szczecinie.

W celu dokładnego zbadania, czy wpływ zmiennych objaśniających X_1, \dots, X_p na hazard zmienia się w czasie, można wykorzystać model postaci:

$$h(t, X) = h_0(t) \exp \left(\sum_{i=1}^p \beta_i X_i + \sum_{i=1}^p \sum_{j=1}^k \delta_j X_i \times g_j(t) \right), \text{ dla } X = (X_1, \dots, X_p), \quad (2)$$

$$\text{gdzie: } g_j(t) = \begin{cases} 1 & \text{dla } t \in \langle t_j, t_{j+1} \rangle \\ 0 & \text{dla } t \notin \langle t_j, t_{j+1} \rangle \end{cases}, \text{ dla } j \in \{1, 2, \dots, k-1\} \quad (3)$$

$$\text{oraz } g_k(t) = \begin{cases} 1 & \text{dla } t \geq t_k \\ 0 & \text{dla } t < t_k \end{cases}.$$

Oszacowano parametry trzech modeli ze zmiennymi objaśniającymi płeć, grupy wieku i grupy wykształcenia. W modelach tych występują dwojakiego rodzaju parametry. Parametry β_i określają wpływ zmiennych endogenicznych na hazard, a parametry δ_j wskazują, czy wpływ ten zmienia się w czasie. Przyjmując ustalone oznaczenia zmiennych, model (2) można zapisać następująco:

$$h(t, P) = h_0(t) \exp \left(\beta_1 P + \sum_{j=1}^3 \delta_j P \times g_j(t) \right), \quad (4)$$

$$h(t, W) = h_0(t) \exp \left(\sum_{i=2}^6 \beta_i W_i + \sum_{i=2}^6 \sum_{j=1}^3 \delta_j W_i \times g_j(t) \right), \quad (5)$$

gdzie: $W = (W_2, W_3, W_4, W_5, W_6)$,

$$h(t, S) = h_0(t) \exp \left(\sum_{i=2}^5 \beta_i S_i + \sum_{i=2}^5 \sum_{j=1}^3 \delta_j S_i \times g_j(t) \right), \quad (6)$$

gdzie: $S = (S_2, S_3, S_4, S_5)$.

W przeprowadzonym badaniu czas podzielono na cztery przedziały, ustalając następujące punkty podziału: 3 miesiące, 6 i 12 miesięcy. Podział taki pozwolił na określenie wpływu płci, wieku i wykształcenia na szansę szybkiego (intensywność [Balicki 2006, s. 31]) podjęcia zatrudnienia przez osoby z grup o określonym rodzaju bezrobocia: krótkookresowym (do 3 miesięcy), średniookresowym (od 3 do 6 miesięcy), długookresowym (od 6 do 12 miesięcy) i długotrwałym (12 miesięcy i więcej). Wzory (3) w tym przypadku przyjmują postać:

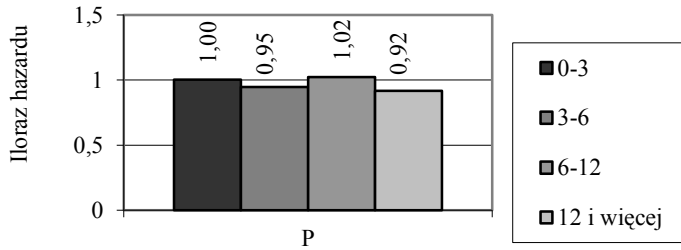
$$g_1(t) = \begin{cases} 1 & \text{dla } t \in \langle 3, 6 \rangle \\ 0 & \text{dla } t \notin \langle 3, 6 \rangle \end{cases}, \quad g_2(t) = \begin{cases} 1 & \text{dla } t \in \langle 6, 12 \rangle \\ 0 & \text{dla } t \notin \langle 6, 12 \rangle \end{cases}, \quad g_3(t) = \begin{cases} 1 & \text{dla } t \geq 12 \\ 0 & \text{dla } t < 12 \end{cases}. \quad (7)$$

Wyniki szacowania parametrów trzech modeli nieproporcjonalnego hazardu przedstawiono w tab. 3. W modelu pierwszym wszystkie parametry były nieistotne statystycznie, co potwierdza, że szanse kobiet i mężczyzn na szybkie podjęcie zatrudnienia były zbliżone i ich ilorazy nie zmieniały się w czasie (rys. 3). Wpływ wieku bezrobotnych na hazard zmieniał się w czasie (rys. 4) w każdej z grup wieku, przy czym największe szanse mieli krótkookresowo bezrobotni. Ponadto dla każdego rodzaju bezrobocia szanse podjęcia pracy malały wraz z wiekiem. Wyniki estymacji parametrów modelu 3 potwierdziły różnice między grupami wykształcenia i nie pozwoliły na odrzucenie hipotezy o proporcjonalności hazardu. Oznacza to, że wpływ wykształcenia na szanse szybkiego wejścia na rynek pracy nie zmieniał się w czasie (rys. 5).

Tabela 3. Modele nieproporcjonalnego hazardu Coxa – wyniki estymacji

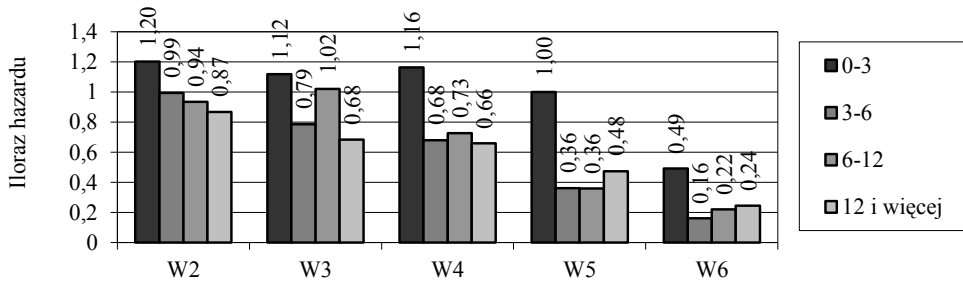
Zmienne	Ocena parametru	Błąd standardowy	Statystyka Walda	Poziom p
Model 1				
P	0,0044	0,0363	0,0150	0,9024
$Px(3 \leq t < 6)$	-0,0597	0,0633	0,8912	0,3452
$Px(6 \leq t < 12)$	0,0185	0,0596	0,0961	0,7566
$Px(t \geq 12)$	-0,0912	0,0727	1,5729	0,2098
Model 2				
W2	0,1842	0,0521	12,4871	0,0004
W3	0,1113	0,0610	3,3300	0,0680
W4	0,1520	0,0609	6,2302	0,0126
W5	0,0007	0,0820	0,0001	0,9934
W6	-0,7088	0,2131	11,0663	0,0009
$W2x(3 \leq t < 6)$	-0,1905	0,0881	4,6722	0,0307
$W2x(6 \leq t < 12)$	-0,2504	0,0892	7,8765	0,0050
$W2x(t \geq 12)$	-0,3254	0,1241	6,8789	0,0087
$W3x(3 \leq t < 6)$	-0,3511	0,1050	11,1731	0,0008
$W3x(6 \leq t < 12)$	-0,0904	0,1003	0,8134	0,3671
$W3x(t \geq 12)$	-0,4923	0,1397	12,4111	0,0004
$W4x(3 \leq t < 6)$	-0,5373	0,1076	24,9433	0,0000
$W4x(6 \leq t < 12)$	-0,4700	0,1030	20,8067	0,0000
$W4x(t \geq 12)$	-0,5693	0,1357	17,6100	0,0000
$W5x(3 \leq t < 6)$	-1,0161	0,1641	38,3319	0,0000
$W5x(6 \leq t < 12)$	-1,0238	0,1499	46,6298	0,0000
$W5x(t \geq 12)$	-0,7450	0,1618	21,1984	0,0000
$W6x(3 \leq t < 6)$	-1,1151	0,4643	5,7688	0,0163
$W6x(6 \leq t < 12)$	-0,7989	0,3641	4,8141	0,0282
$W6x(t \geq 12)$	-0,6981	0,3443	4,1114	0,0426
Model 3				
S2	0,4849	0,0566	73,4634	0,0000
S3	0,3909	0,0689	32,1988	0,0000
S4	0,4912	0,0585	70,5356	0,0000
S5	0,8266	0,0533	240,2207	0,0000
$S2x(3 \leq t < 6)$	-0,1374	0,1017	1,8246	0,1768
$S2x(6 \leq t < 12)$	-0,1728	0,0957	3,2616	0,0709
$S2x(t \geq 12)$	-0,2478	0,1205	4,2290	0,0397
$S3x(3 \leq t < 6)$	-0,0338	0,1227	0,0760	0,7828
$S3x(6 \leq t < 12)$	-0,0923	0,1163	0,6297	0,4275
$S3x(t \geq 12)$	0,0281	0,1374	0,0419	0,8377
$S4x(3 \leq t < 6)$	-0,1607	0,1049	2,3454	0,1257
$S4x(6 \leq t < 12)$	0,0096	0,0952	0,0102	0,9197
$S4x(t \geq 12)$	-0,0912	0,1164	0,6137	0,4334
$S5x(3 \leq t < 6)$	0,1426	0,0926	2,3724	0,1235
$S5x(6 \leq t < 12)$	-0,0967	0,0891	1,1780	0,2778
$S5x(t \geq 12)$	-0,2910	0,1100	6,9985	0,0082

Źródło: obliczenia własne na podstawie danych z PUP w Szczecinie.



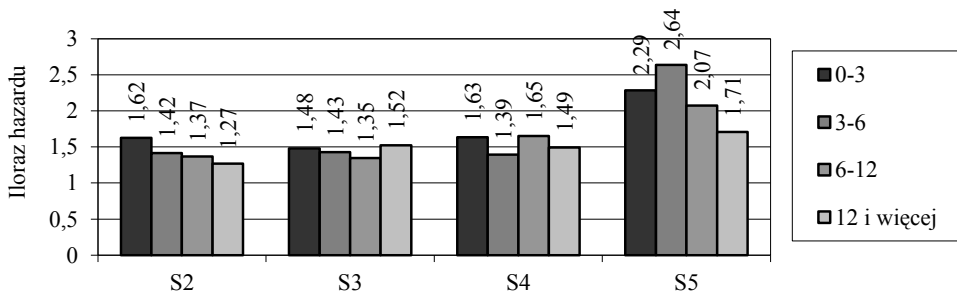
Rys. 3. Ilorazy hazardu według rodzaju bezrobocia dla płci

Źródło: opracowanie własne na podstawie danych z PUP w Szczecinie.



Rys. 4. Ilorazy hazardu według rodzaju bezrobocia dla grup wieku

Źródło: opracowanie własne na podstawie danych z PUP w Szczecinie.



Rys. 5. Ilorazy hazardu według rodzaju bezrobocia dla grup wykształcenia

Źródło: opracowanie własne na podstawie danych z PUP w Szczecinie.

4. Wnioski

Na podstawie przeprowadzonych badań można sformułować wnioski badawcze i poznawcze. Wyniki analizy danych wskazały na to, że szanse na podjęcie pracy

w Szczecinie w 2010 r.: nie różniły się istotnie dla kobiet i mężczyzn, malały wraz z wiekiem i rosły wraz z poziomem wykształcenia⁴. Badanie proporcjonalności hazardu, zarówno metodą graficzną, jak i przy wykorzystaniu modeli nieproporcjonalnego hazardu, wykazało, że wpływ płci i wykształcenia na hazard nie zmieniał się czasie, natomiast był zróżnicowany w przypadku zmiennej objaśniającej wiek. Oznacza to, że jedynie przynależność do grupy wieku wpływała na różnice w szansach podjęcia pracy w zależności od rodzaju bezrobocia. Im dłuższy czas poszukiwania pracy i wyższy wiek osoby zarejestrowanej w PUP, tym mniejsze były szanse na jej znalezienie. Zarówno przedstawione w niniejszym artykule, jak i wcześniejsze badania auterek pozwoliły na sformułowanie wniosku poznawczego. W przypadku stwierdzenia nieproporcjonalności hazardu istnieje możliwość zastosowania obu modeli regresji Coxa: proporcjonalnego i nieproporcjonalnego hazardu. Do badacza należy wybór jednego z nich. Należy jednak pamiętać o tym, że pierwszy z modeli dostarcza wyników uśrednionych w czasie. Stanowi więc przybliżenie wartości opisujących badane zjawisko.

Literatura

- Balicki A., *Analiza przeżycia i tablice wymieralności*, PWE, Warszawa 2006.
- Bednarski T., *Ocena przydatności danych Bael dla charakterystyki rozkładu czasu poszukiwania pracy na przykładzie danych z lat 2001-2002*, „Studia Ekonomiczne” nr 4, 2005.
- Bieszk-Stolorz B., Markowicz I., *Modele regresji Coxa w analizie bezrobocia*, CeDeWu, Warszawa 2012.
- Cox D.R., Oakes D., *Analysis of Survival Data*, Chapman and Hall, London 1984.
- Daras T., Jerzak M., *Wpływ cech społeczno-demograficznych osób bezrobotnych na możliwość znalezienia pracy, badanie na podstawie danych BAEL w latach 1993-2003*, „Materiały i Studia”, z. 189, NBP, Warszawa 2005.
- Foley M.C., *Determinants of unemployment duration In Russia*, Working Paper nr 81, University of Michigan Business School, 1997, <http://www.wdi.umich.edu>.
- Kaplan E.L., Meier P., *Nonparametric estimation from incomplete observations*, “Journal of the American Statistical Association” 1958, 53, s. 457-481.
- Malarska A., *Bezrobocie w Polsce w ujęciu regionalnym. Studium statystyczne*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 2000.
- Marcassa S., *Unemployment Duration of Spouses: Evidence From France*, THEMA Working Papers 2012-31, Université de Cergy-Pontoise 2011.
- Therneau T.M., Grambsch P.M., *Modelling Survival Data: Extending the Cox Model*, Springer, New York 2000.

⁴ Zastosowanie metod analizy przeżycia w badaniu bezrobocia można znaleźć np. w pracach: [Bednarski 2005; Daras, Jerzak 2005; Foley 1997; Malarska 2000; Marcassa 2011].

USING PROPORTIONAL AND NON PROPORTIONAL COX HAZARD MODELS TO RESEARCH THE CHANCES FOR TAKING UP A JOB ACCORDING TO THE TYPE OF UNEMPLOYMENT

Summary: The aim of the article is to analyse the chances for founding the job quickly with using the non proportional Cox hazard model. In this model it is assumed that the influence of explanatory variable on hazard depends on time. Preliminary research showed that the hazard ratios had not always been the same for the unemployed with different time of looking for a job. The results allowed to determine the influence of gender, age and education on the chance of finding the job quickly by the persons from groups with determined type of unemployment: short-term (to 3 months), middle-term (from 3 to 6 months), long-term (from 6 to 12 months) and long-lasting (12 months and more).

Keywords: Cox hazards model, hazards proportion research, unemployment.