

*Iwona Markowicz**
*Beata Bieszk-Stolorz***

DETERMINANTY CZASU OCZEKIWANIA NA PRACĘ I ICH WZAJEMNE ODDZIAŁYWANIE

Streszczenie. W artykule przedstawiono wyniki badań dotyczące wpływu płci, wieku i wykształcenia bezrobotnych na czas oczekiwania na pracę. Ze względu na specyfikę badanego zjawiska, a szczególnie na występowanie danych cenzurowanych, do identyfikacji determinant oraz do zbadania siły ich wpływu na czas do podjęcia pracy wykorzystany zostanie model proporcjonalnego hazardu Coxa. Do modelu oprócz zmiennych niezależnych włączone zostaną ich iloczyn (interakcje), które umożliwią wyznaczenie wzajemnego oddziaływania badanych zmiennych. Badaniem objęto 5074 osoby zarejestrowane jako bezrobotne w Powiatowym Urzędzie Pracy w Szczecinie i wyrejestrowane w IV kwartale 2006 roku. Jeżeli wyrejestrowanie nastąpiło z powodu innego niż znalezienie pracy, to taka obserwacja będzie uznana za uciętą.

Słowa kluczowe: bezrobocie, analiza przeżycia, model proporcjonalnego hazardu Coxa.

I. WSTĘP

Analiza ekonomiczna rynku pracy obejmuje przede wszystkim czynniki determinujące popyt na pracę (zapotrzebowanie na pracę zgłaszane przez przedsiębiorstwa) oraz podaż pracy (ilość osób chętnych do pracy). Popyt na pracę i podaż pracy decydują o liczbie osób pracujących i bezrobotnych. Niedostateczny popyt w gospodarce jest podstawową przyczyną występowania bezrobocia. Głównym celem artykułu jest analiza determinant czasu oczekiwania na pracę osób bezrobotnych zarejestrowanych w Powiatowym Urzędzie Pracy w Szczecinie, które znalazły zatrudnienie w ostatnim kwartale 2006 roku. Celem artykułu jest również zaprezentowanie możliwości wykorzystania metod analizy przeżycia do badania zjawiska bezrobocia. Analiza przeżycia¹ jest to zbiór procedur statystycznych, dla których zmienną losową jest czas między określonymi zdarzeniami bądź czas procesu. Zdarzenie powoduje przejście jednostki z jednego stanu w drugi (np. śmierć osoby, awaria urządzenia, upadek firmy (Markowicz, Stolorz, 2006), wyrejestrowanie z PUP). Okres między stanem początkowym

* Dr, Katedra Ekonometrii i Statystyki, Wydział Nauk Ekonomicznych i Zarządzania, Uniwersytet Szczeciński.

** Dr, Katedra Ekonometrii i Statystyki, Wydział Nauk Ekonomicznych i Zarządzania, Uniwersytet Szczeciński.

¹ Więcej na temat analizy przeżycia w pracy Frątczak, Gach-Ciepiela, Babiker, 2005.

a momentem wystąpienia zdarzenia nazywamy czasem przeżycia. Wyznaczając prawdopodobieństwo, że jednostka przeżyje kolejne wartości czasu t można zdefiniować funkcję przeżycia. Funkcje takie utworzone dla dwóch lub więcej prób można porównywać. Przykładem zmiennej losowej w tym przypadku jest czas zarejestrowania bezrobotnego w PUP. Zbadano, jaki wpływ na długość czasu oczekiwania na pracę ma płeć, wiek i wykształcenie osoby bezrobotnej oraz jakie są współzależności między tymi zmiennymi.

II. CHARAKTERYSTYKA DANYCH WYKORZYSTANYCH W BADANIU

Badania przeprowadzono na podstawie danych udostępnionych przez Powiatowy Urząd Pracy w Szczecinie dotyczących zarejestrowanych bezrobotnych. Kohortę tworzą osoby wyrejestrowane w ostatnim kwartale 2006 roku. Funkcja przeżycia opisuje w tym przypadku czas od chwili zarejestrowania bezrobotnego w urzędzie do momentu znalezienia zatrudnienia. Celem badawczym jest ustalenie czy płeć osoby bezrobotnej, jej wykształcenie oraz wiek mają wpływ na długość czasu pozostawania w rejestrze do chwili znalezienia pracy. Osoby wyrejestrowane z innych powodów traktowane są jako obserwacje ucięte (por. Domański, Pruska, 2000) (cenzurowane) prawostronnie. Do obliczeń wykorzystano podział na grupy według wykształcenia i wieku stosowany przez Powiatowy Urząd Pracy. Do celów badawczych poszczególne grupy ponumerowano. Łącznie w analizie wykorzystano dane 5074 osób. Charakterystykę ilościową badanych grup i sposób numeracji przedstawiono w tabeli 1.

Tabela 1. Charakterystyka ilościowa badanych grup

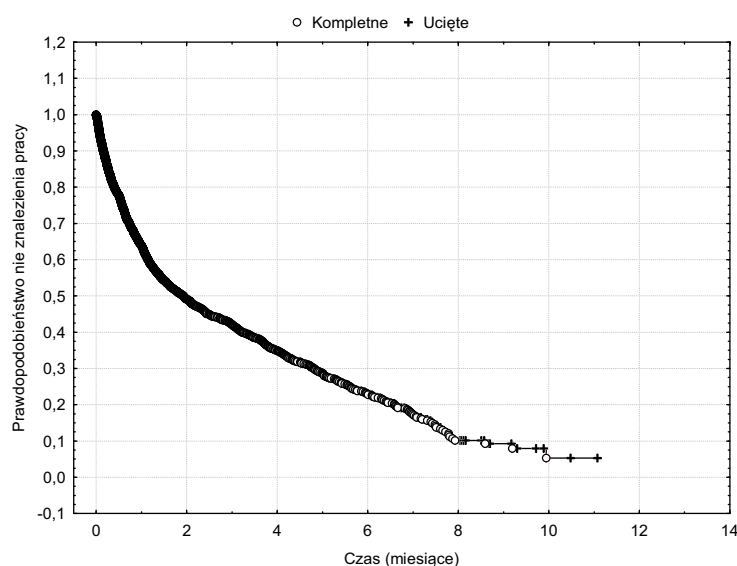
Cecha	Grupy (numeracja)		Liczebność grupy	Obserwacje ucięte		Obserwacje pełne	
				liczba	odsetek	liczba	odsetek
1	2	3	4	5	6	7	8
Płeć	kobiety	(1)	2469	1277	51,72	1192	48,28
	mężczyźni	(0)	2605	1652	63,42	953	36,58
Przedziały wieku	(18,25)	(1)	1219	703	57,67	516	42,33
	(25,35)	(2)	1718	920	53,55	798	46,45
	(35,45)	(3)	817	459	56,18	358	43,82
	(45,55)	(4)	1019	634	62,22	385	37,78
	(55,60)	(5)	249	172	69,08	77	30,92
	(60,65)	(6)	52	41	78,85	11	21,15

Tabela 1 (cd.)

1	2	3	4	5	6	7	8
Wykształcenie	brak lub niepełne podstawowe, podstawowe, gimnazjalne,	(1)	1755	1226	69,86	529	30,14
	zasadnicze zawodowe,	(2)	1055	628	59,53	427	40,47
	średnie ogólnokształcące,	(3)	488	273	55,94	215	44,06
	średnie zawodowe 4-letnie, średnie zawodowe, pomaturalne / policealne,	(4)	941	485	51,54	456	48,46
	wyższe (w tym licencjat)	(5)	835	317	37,96	518	62,04

Źródło: opracowanie własne na podstawie danych PUP w Szczecinie.

Funkcja przeżycia danej populacji otrzymana na podstawie estymatora Kaplana-Meiera jest przedstawiona na rysunku 1.



Rys.1. Funkcja przeżycia – estymator Kaplana-Meiera. Prawdopodobieństwo nie znalezienia pracy w czasie t dla osób wyrejestrowanych z PUP w IV kwartale 2006 roku

Źródło: Opracowanie własne – *Statistica*.

III. IDENTYFIKACJA DETERMINANT CZASU POZOSTAWANIA BEZROBOTNYM

W pierwszym etapie badania zweryfikowano hipotezę, że przyjęte zmienne (płeć, wykształcenie, wiek) istotnie statystycznie różnicują czas poszukiwania pracy w podgrupach. Czasy przeżycia można porównywać w dwóch lub więcej

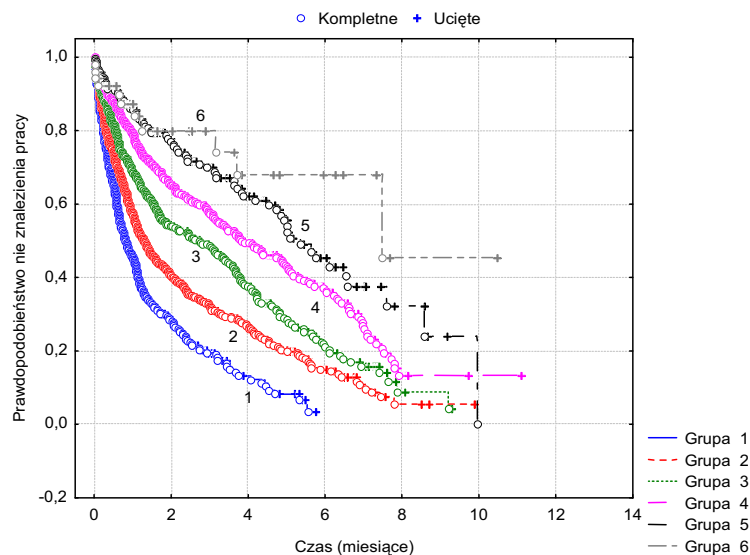
podgrupach. Ponieważ nie są znane ich rozkłady, należy stosować test nieparametryczny. Test taki jest oparty na porządku rangowym czasów przeżycia. Zastosowano test porównania przeżycia dla wielu prób, wykorzystując program *Statistica*. W przypadku dwóch podgrup (dla płci) test ten jest równoważny testowi Gehana (Gehan, 1965a, Gehan, 1965b) (uogólnienie testu Wilcoxona). Umożliwia on oszacowanie funkcji przeżycia dla każdej z podgrup i zbadanie istotności różnic między nimi. Wyniki przeprowadzonej weryfikacji zawiera tabela 2.

Tabela 2. Wyniki testu porównania czasu do znalezienia pracy badanych podgrup

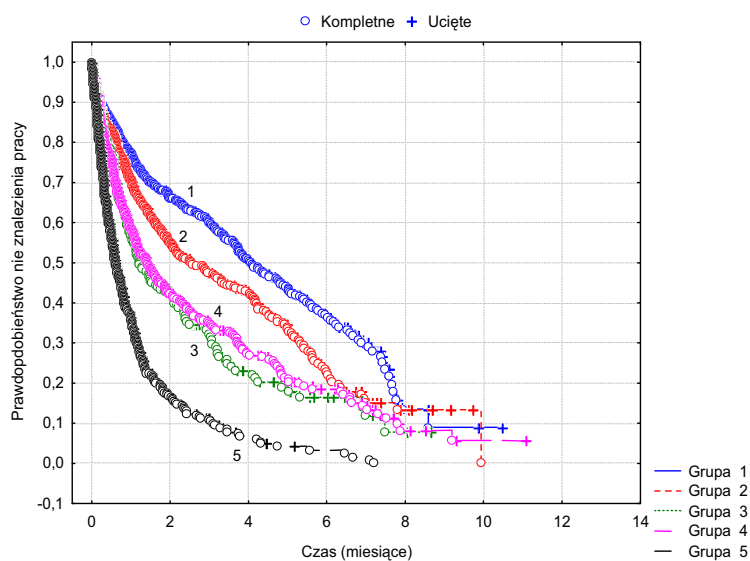
Cecha	Wariant	Wartość p
Płeć	kobiety	0,00949
	mężczyźni	
Przedziały wieku	<18,25)	0,00000
	<25,35)	
	<35,45)	
	<45,55)	
	<55,60)	
	<60,65)	
Wykształcenie	brak lub niepełne podstawowe, podstawowe, gimnazjalne,	0,00000
	zasadnicze zawodowe	
	średnie ogólnokształcące	
	średnie zawodowe 4-letnie, średnie zawodowe, pomaturalne / policealne	
	wyższe (w tym licencjat)	

Źródło: Obliczenia własne – *Statistica*.

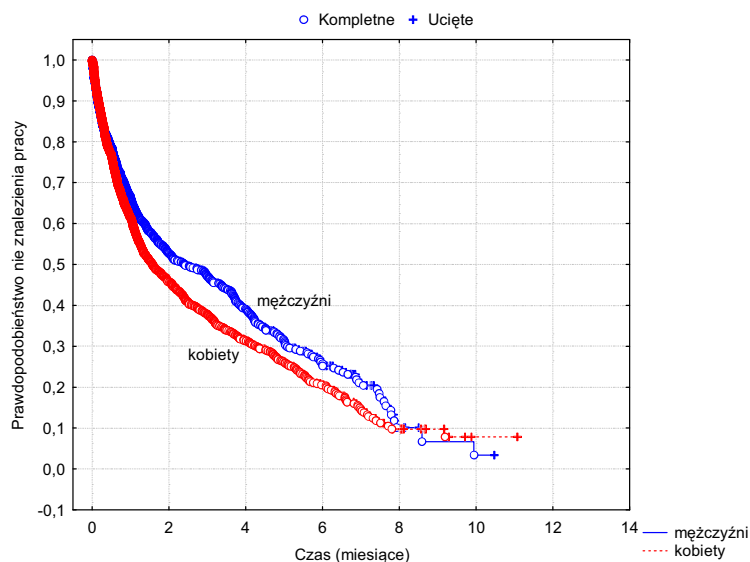
Przyjmując poziom istotności 0,05 można stwierdzić, że badane podgrupy bezrobotnych, wydzielone ze względu na płeć, przedział wieku i wykształcenie różnią się czasem przeżycia (tablica 1, wartość p). Zatem cechy te mają wpływ na intensywność znajdowania pracy. Prawdopodobieństwa przeżycia, oszacowane metodą Product-Limit-Estimation (PLE) Kaplana-Meiera (Lawless, 1982, s. 71–81) dla bezrobotnych według wieku, wykształcenia i płci są przedstawione odpowiednio na rysunkach 2–4.



Rys. 2. Funkcja przeżycia – estymator Kaplana-Meiera. Prawdopodobieństwo nie znalezienia pracy przez bezrobotnych według wieku.
 Źródło: Opracowanie własne – *Statistica*.



Rys. 3. Funkcja przeżycia – estymator Kaplana-Meiera. Prawdopodobieństwo nie znalezienia pracy przez bezrobotnych według wykształcenia.
 Źródło: Opracowanie własne – *Statistica*.



Rys. 4. Funkcja przeżycia – estymator Kaplana-Meiera. Prawdopodobieństwo nie znalezienia pracy przez bezrobotnych według płci.

Źródło: Opracowanie własne – *Statistica*.

IV. ANALIZA CZASU POZOSTAWANIA BEZROBOTNYM

Do identyfikacji determinant czasu pozostawania bez pracy wykorzystano model proporcjonalnego hazardu Coxa (Por. Frątczak, Gach-Ciepiela, Babiker, 2005, s. 111, Bednarski, 2005, *Metody...*, 2004, s.76, 185) (model regresji przedstawiony w pracy Cox, Oakes, 1984; zwany też modelem proporcjonalnych hazardów), który można zapisać następująco:

$$h(t : x_1, x_2, \dots, x_n) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n),$$

gdzie:

x_1, x_2, \dots, x_n – zmienne niezależne,

$h_0(t)$ – hazard odniesienia lub zerowa linia hazardu,

$\beta_1, \beta_2, \dots, \beta_n$ – współczynniki modelu,

t – czas obserwacji.

W modelu po oszacowaniu parametrów β_i do dalszej analizy najczęściej wykorzystuje się ilorazy hazardów określone wzorem $HR_i = e^{\beta_i}$. Ilorazy te umożliwiają zbadanie prawdopodobieństwa zajścia zdarzenia w danej grupie

w porównaniu z prawdopodobieństwem zajścia zdarzenia w innej grupie. Z tego też powodu wartość ta nazywana jest szansą względną (lub ryzykiem względnym).

Drugim etapem analizy była budowa modelu proporcjonalnego hazardu Coxa umożliwiającą ocenę jednoczesnego wpływu wielu zmiennych na czas przeżycia. Wykorzystano zero-jedynkowe kodowanie zmiennych zgodnie z procedurą przedstawioną przez Hosmera i Lemeshowa (1999, s. 120–121). W omawianej analizie zerem zakodowano: w przypadku wykształcenia – wykształcenie co najwyżej gimnazjalne, dla płci – mężczyźni i dla wieku – przedział $(18, 25)$ lat.

Skonstruowano dwa modele regresji. W pierwszym z nich wiek jest zmienną ciągłą, w drugim – zmienną kategoryzowaną. Wyniki przeprowadzonego modelowania zawierają tabele 3–4.

Istotność uzyskanych parametrów modelu Coxa weryfikuje się na podstawie statystyki Walda (Kleinbaum, Klein, 2005, s. 89). Istotność wskazuje wartość p poniżej przyjętego poziomu (w badaniu 0,05). Parametr przy zmiennej płeć w obu skonstruowanych modelach nie jest istotny. Interpretując ocenę parametru dla wieku, jako zmiennej ciągłej (w modelu pierwszym) należy stwierdzić, że wraz ze wzrostem wieku osoby bezrobotnej o jeden rok prawdopodobieństwo znalezienia pracy malało średnio o 3,3%. Dokładniejszych informacji o relacjach między poszczególnymi przedziałami wieku dostarcza model drugi. Wartości parametrów dla wykształcenia są zbliżone do siebie w obu oszacowanych modelach.

Tabela 3. Pierwszy model proporcjonalnego hazardu Coxa dla zmiennej ciągłej wiek i zmiennych kategoryzowanych: wykształcenie i płeć ($\chi^2 = 693,507, p = 0,0000$)

Cecha	Wariant	Parametr β	Ryzyko względne	Statystyka Walda	Wartość p
Wykształcenie	brak lub niepełne podstawowe, podstawowe, gimnazjalne,		1,00000		
	zasadnicze zawodowe	0,28559	1,33056	19,2479	0,00001
	średnie ogólnokształcące,	0,48886	1,63008	34,6282	0,00000
	średnie zawodowe 4-letnie, średnie zawodowe, pomaturalne / policealne,	0,56746	1,76378	77,2580	0,00000
	wyższe (w tym licencjat).	1,10189	3,00987	281,5147	0,00000
Płeć	mężczyźni		1,00000		
	kobiety	-0,07034	0,93207	2,4379	0,11845
Wiek	Zmienna ciągła	-0,03334	0,96721	232,3486	0,00000

Źródło: Obliczenia własne – *Statistica*.

Tabela 4. Drugi model proporcjonalnego hazardu Coxa dla zmiennych kategoryzowanych: wykształcenie, płeć i przedział wieku ($\chi^2 = 701,554$, $p = 0,0000$)

Cecha	Wariant	Parametr <i>beta</i>	Ryzyko względne	Statystyka Walda	Wartość <i>p</i>
Wykształcenie	brak lub niepełne podstawowe, podstawowe, gimnazjalne,		1,00000		
	zasadnicze zawodowe	0,28140	1,324982	18,5692	0,000016
	średnie ogólnokształcące,	0,48760	1,628401	34,3843	0,000000
	średnie zawodowe 4-letnie, średnie zawodowe, pomaturalne / policealne,	0,56548	1,760287	76,4566	0,000000
	wyższe (w tym licencjat).	1,17067	3,224167	312,4498	0,000000
Płeć	mężczyźni		1,00000		
	kobiety	-0,07337	0,929254	2,6243	0,10524
Wiek	<18,25)		1,00000		
	<25,35)	-0,47145	0,624097	66,2730	0,000000
	<35,45)	-0,69590	0,498627	95,7387	0,000000
	<45,55)	-0,97480	0,377269	187,9034	0,000000
	<55,60)	-1,29371	0,274250	106,8040	0,000000
	<60,65)	-1,64999	0,192051	28,6835	0,000000

Źródło: Obliczenia własne – *Statistica*.

Stosując zero-jedynkowe kodowanie wariantów zmiennych do szacowania wieloczynnikowego modelu regresji Coxa, poza wyznaczeniem szansy względnej znalezienia pracy w porównaniu z kategorią zakodowaną jako zero. Istnieje również możliwość obliczenia szansy względnej między pozostałymi kategoriami badanej cechy jako stosunek funkcji proporcjonalnego hazardu dla porównywanych kategorii danej zmiennej, przy założeniu stałości pozostałych zmiennych objaśniających (por. Hosmer, Lemeshow, 1999, s. 123–124). Otrzymane wartości szansy względnej dla wykształcenia i przedziałów wieku zaprezentowano w tabelach 5–6.

Tabela 5. Szansa względna znalezienia pracy według wykształcenia – wyznaczona na podstawie drugiego modelu regresji Coxa

Szansa względna podjęcia pracy bezrobotnych o wykształceniu	w stosunku do bezrobotnych o wykształceniu			
	1	2	3	4
2	1,32498			
3	1,62840	1,22899		
4	1,76028	1,08099	0,87957	
5	3,22416	1,83161	1,69438	1,92637

Źródło: Obliczenia własne – *Statistica*.

Tabela 6. Szansa względna znalezienia pracy według przedziału wieku – wyznaczona na podstawie drugiego modelu regresji Coxa

Szansa względna podjęcia pracy bezrobotnych w wieku	w stosunku do bezrobotnych w wieku				
	(18, 25)	(25, 35)	(35, 45)	(45, 55)	(55, 60)
(25, 35)	0,62409				
(35, 45)	0,49862	0,79895			
(45, 55)	0,37726	0,75661	0,94700		
(55, 60)	0,27425	0,72693	0,96077	1,01453	
(60, 65)	0,1920	0,70027	0,96332	1,00266	0,98829

Źródło: Obliczenia własne – *Statistica*.

Jak wynika z tabeli 5, wraz ze wzrostem wykształcenia rośnie szansa na znalezienie pracy, przy czym dla osób z wykształceniem wyższym jest ona ponad trzykrotnie większa niż dla osób z wykształceniem co najwyżej gimnazjalnym. Wartości zawarte w tabeli 6 wskazują na większe zainteresowanie pracodawców osobami młodymi. Wraz z wiekiem prawdopodobieństwo szybkiego znalezienia pracy maleje, przy czym jednocześnie maleją różnice między dwoma kolejnymi przedziałami wieku. Zatem wraz z wiekiem szansa na znalezienie pracy maleje coraz wolniej.

V. ANALIZA WIELOCZYNNIKOWA Z UWZGLĘDNIENIEM INTERAKCJI ZMIENNYCH NIEZALEŻNYCH

Wieloczynnikowy model proporcjonalnego hazardu Coxa umożliwia ocenę jednoczesnego wpływu wielu zmiennych na czas do wystąpienia określonego zdarzenia. Uwzględnienie w modelu również iloczynu zmiennych umożliwia porównanie stopnia ryzyka między poszczególnymi kategoriami jednej zmiennej przy ustalonym poziomie drugiej zmiennej². W kolejnym, trzecim badaniem

² Więcej na temat interakcji w Kleinbaum, Klein, 2005, s. 100–103.

modelu, uwzględniono zmienną dychotomiczną – płeć, zmienną ciągłą – wiek oraz interakcję między nimi. Najpierw za pomocą modelu jednoczynnikowego zbadano zasadność włączenia badanych zmiennych do konstruowanego modelu. Wyniki prezentuje tabela 7.

Tabela 7. Jednoczynnikowe modele proporcjonalnego hazardu Coxa

Zmienna	Parametr <i>beta</i>	Błąd	Wykładnik <i>beta</i>	Statystyka Walda	<i>p</i>
Płeć	0,170428	0,043492	1,185812	15,35526	0,000089
Wiek	-0,039938	0,002073	0,960850	371,0848	0,000000

Źródło: Obliczenia własne – *Statistica*.

Dla obydwu zmiennych parametry są istotne, a zatem można je włączyć do modelu wieloczynnikowego z interakcją (tabela 8).

Tabela 8. Wieloczynnikowy model Coxa z interakcją między zmiennymi niezależnymi ($\chi^2 = 408,441$, $p = 0,0000$)

Zmienna	Parametr <i>beta</i>	Błąd	Wykładnik <i>beta</i>	Statystyka Walda	<i>p</i>
Płeć	0,159304	0,143703	1,172694	1,2289	0,267628
Wiek	-0,038378	0,002876	0,962349	178,0960	0,000000
Płeć × Wiek	-0,002454	0,004000	0,997549	0,3764	0,539544

Źródło: Obliczenia własne – *Statistica*.

Celem budowy modelu z interakcją było otrzymanie ilorazu szans na znalezienie pracy bezrobotnych kobiet i mężczyzn w danym wieku. Brak istotności parametru przy zmiennej Płeć × Wiek świadczy o braku interakcji między płcią i wiekiem. Rozpatrując wpływ tych dwóch cech na czas poszukiwania pracy można stwierdzić, że determinantą jest wiek, a płeć osób bezrobotnych w określonym wieku nie wpływa na intensywność znajdowania pracy.

VI. PODSUMOWANIE

Na rynku pracy w Szczecinie w ostatnim kwartale 2006 roku zgłaszany był popyt na pracowników młodych i dobrze wykształconych. To właśnie takie osoby najszybciej znajdowały pracę. Im wyższe było wykształcenie bezrobotnego tym większe było prawdopodobieństwo znalezienia przez niego pracy, natomiast im bezrobotny był starszy tym – tym to prawdopodobieństwo mniejsze. Interesu-

jące jest zjawisko szybszego znajdowania pracy przez kobiety w porównaniu do mężczyzn. Krzywa przeżycia kobiet w populacji (czyli wśród osób zarejestrowanych w PUP) leży poniżej krzywej przeżycia mężczyzn, co oznacza, że to właśnie one w krótszym czasie podejmowały zatrudnienie. Brak podstaw do odrzucenia hipotezy o statystycznej równości funkcji przeżycia potwierdził przeprowadzony test Gehana dla dwóch prób. Może to być spowodowane większą desperacją w poszukiwaniu pracy oraz mniejszymi wymaganiami płacowymi. Jednak przy jednoczesnym uwzględnieniu wieku, płeć nie miała istotnego wpływu na czas poszukiwania pracy.

BIBLIOGRAFIA

- Bednarski T., 2005, *Ocena przydatności danych BAEL dla charakterystyki rozkładu czasu poszukiwania pracy na przykładzie danych z lat 2001-2002*, „Studia Ekonomiczne” nr 4.
- Cox D. R., Oakes D., 1984, *Analysis of Survival Data*, Chapman and Hall, London.
- Domański C., Pruska K., 2000, *Nieklasyczne metody statystyczne*, PWE, Warszawa.
- Frątczak E., Gach-Ciepiela U., Babiker H., 2005, *Analiza historii zdarzeń. Elementy teorii, wybrane przykłady zastosowań*, SGH, Warszawa.
- Gehan, E. A., 1965a, A generalized Wilcoxon test for comparing arbitrary single-censored samples, *Biometrika*, nr 52.
- Gehan, E. A., 1965b, A generalized two-sample Wilcoxon test for double-censored data, *Biometrika* nr 52.
- Hosmer D.W., Lemeshow S., 1999, *Applied Survival Analysis. Regression Modeling of Time to Event Data*, John Wiley & Sons, INC, New York.
- Kleinbaum D. G., Klein M., 2005, *Survival Analysis, Second Edition*, Springer, New York.
- Lawless J. F., 1982, *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, New York.
- Markowicz I., Stolorz B., 2006, *Analysis of the Survival Function of Firms*. Baltic Business Development. SME Management International Degree Programmes. Faculty of Economics and Management, Szczecin University, Szczecin.
- Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, 2004, red. E. Gattnar, M. Walesiak, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław.

Iwona Markowicz, Beata Bieszk-Stolorz

DETERMINANTS OF JOB AWAITING TIME AND THEIR INTERACTION

Abstract

In the paper results of survey concerning the influence of gender, age and education of unemployed on job awaiting time will be presented. For the sake of specificity of survey, especially of occurrence of censored data, the proportional hazard model of Cox will be used for identification of determinants and research their influence for time needed to take up a job. To the model except of independent variables, their interactions that enable calculation of interplay of researched variables are attached. 5074 persons registered as unemployed in the Local Labour Office in Szczecin and due to various reasons unregistered in the fourth quarter of 2006 have been included in the research. If unregistering occurred due to different reasons than finding a job, such a observation is considered as censored.