

## METODY ANALIZY TRWANIA W IDENTYFIKACJI DETERMINANT RYZYKA WYKREŚLENIA Z REJESTRU BEZROBOTNYCH

**Iwona Markowicz, Beata Bieszk-Stolorz**

Wydział Nauk Ekonomicznych i Zarządzania

Uniwersytet Szczeciński

e-mail: iwona.markowicz@wneiz.pl, beatus@wneiz.pl

**Streszczenie:** Rejestr bezrobotnych może stanowić bazę danych do badania sytuacji na regionalnym rynku pracy. Ważne są w tym przypadku zarówno przyczyna bezrobocia, jak i powód wyrejestrowania. Osoby, które nie zgłosiły się w urzędzie w wyznaczonym terminie lub nie przyjęły propozycji pracy stanowią dużą grupę wśród zarejestrowanych. Stąd podjęto próbę identyfikacji cech wpływających na ryzyko wykreślenia z rejestru. Celem artykułu jest ocena ryzyka wykreślenia z rejestru i identyfikacja jego determinant. Analizie poddano czas pozostawiania w rejestrze bezrobotnych, stąd wykorzystano metody analizy trwania.

**Słowa kluczowe:** model regresji Coxa, tablice trwania, funkcja hazardu, bezrobocie

### WSTĘP

Wśród osób rejestrujących się w Powiatowych Urzędach Pracy są osoby rzeczywiście poszukujące tej pracy i potrzebujące pomocy w jej znalezieniu, ale są i takie osoby, które z tej pomocy korzystać nie chcą. Zarejestrowani, którzy nie wyrażą zgody na określenie profilu pomocy, odmówią bez uzasadnionej przyczyny przyjęcia propozycji odpowiedniej pracy lub innej formy pomocy ze strony PUP, poddania się badaniom lekarskim lub psychologicznym, mającym na celu ustalenie zdolności do pracy, odmówią bez uzasadnionej przyczyny udziału w działaniach w ramach Programu Aktywizacja i Integracja, realizowanych przez PUP, nie stawiają się w urzędzie w wyznaczonym terminie, z własnej winy przerwą staż, szkolenie, wykonywanie prac społecznie użytecznych, realizację indywidualnego planu działania lub inną formę pomocy, z własnej winy przerwą program

przygotowania zawodowego dorosłych są z rejestru wykreślane [Ustawa z dnia 20 kwietnia 2004] bez podania przyczyny przez osobę zainteresowaną (za pierwszym razem jest to wykreślane na 120 dni). Wykreśleni stanowią dużą grupę wśród wyrejestrowanych bezrobotnych. Stąd podjęto próbę identyfikacji cech wpływających na ryzyko wykreślenia z rejestru.

Celem artykułu jest ocena ryzyka wykreślenia i identyfikacja jego determinant. Źródłem danych był rejestr bezrobotnych Powiatowego Urzędu Pracy w Szczecinie. Wśród analizowanych cech wzięto pod uwagę płeć, wiek, wykształcenie, staż pracy i liczbę kolejnych zarejestrowań osoby bezrobotnej. Przeprowadzone badanie pozwoliło na określenie grup osób bezrobotnych bardziej podatnych na rezygnację z pośrednictwa urzędu bez podania przyczyny.

## WYKORZYSTANE DANE

Dane wykorzystane w badaniu, którego wyniki są prezentowane w niniejszym artykule, pochodzą z rejestru Powiatowego Urzędu Pracy w Szczecinie. Obejmują one dane indywidualne 22078 osób rejestrujących się jako bezrobotne w 2013 roku. Obserwacja trwała do końca 2014 roku. Informacje o danych liczbowych zawarto w tabeli 1.

Tabela 1. Liczba zarejestrowanych bezrobotnych w PUP w Szczecinie w 2013 roku i obserwowanych do końca 2014 roku według cech (zmiennych w modelach)

Zmienna	Wariant	Symbol	Liczba wykreślonych (1)	Liczba pozostałych (0)	Suma
Ogółem			8965	13113	22078
Płeć	mężczyzna	M	5701	6607	12308
	kobieta	K	3264	6506	9770
Wiek	18-24 lata	W1	2257	1891	4148
	25-34 lata	W2	2966	4390	7356
	35-44 lata	W3	1734	2525	4259
	45-54 lata	W4	1214	2283	3497
	55-60 lat	W5	629	1556	2185
	60-65 lat	W6	165	468	633
Wykształcenie	co najwyżej gimnazjalne	S1	2932	2191	5123
	zasadnicze zawodowe	S2	2220	2796	5016
	średnie ogólnokształcące	S3	1223	1636	2859
	średnie zawodowe	S4	1415	2671	4086
	wyższe	S5	1175	3819	4994
Zarejestrowanie	pierwszy raz	Z0	2045	3373	5418
	kolejny raz	Z1	6920	9740	16660
Staż pracy	brak stażu	D0	4195	3362	7557
	posiadany staż	D1	4770	9751	14521

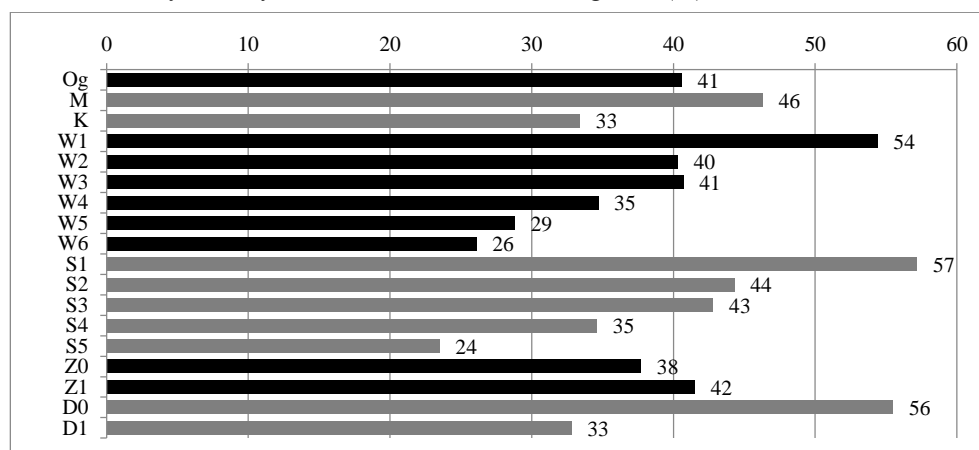
Źródło: opracowanie własne

Poza ogólną liczbą badanych, podane zostały informacje dotyczące cech osób bezrobotnych, takich jak: płeć, wiek, wykształcenie, kolejność zarejestrowania i staż pracy. Cechy te stanowią zmienne w szacowanych modelach, dlatego też wskazano zarówno warianty cech jak i ich symbole. Ostatnia kolumna

wskazuje liczbę osób bezrobotnych rejestrujących się w 2013 roku. Na tą sumę składają się liczba wykreślonych do końca 2014 roku oraz liczba pozostałych osób. Ponieważ przedmiotem zainteresowania w niniejszym badaniu jest wykreślenie bezrobotnego z rejestru, to stanowi ono obserwację pełną i zostało oznaczone wartością 1. Pozostałe przypadki stanowią tu obserwacje cenzurowane i przyjęto dla nich wartość 0 [Hosmer, Lemeshow 1999]. Wśród badanych bezrobotnych przeważają mężczyźni (56%), osoby w wieku 25-34 lata (33%), osoby z najniższym wykształceniem, czyli co najwyżej gimnazjalnym (23%), rejestrujące się po raz kolejny (75%) oraz posiadające już doświadczenia zawodowe (66%).

Udział bezrobotnych wykreślonych z rejestru jest stosunkowo wysoki, wynosi 41% wszystkich rejestrujących się w 2013 roku w PUP w Szczecinie. Udział ten według poszczególnych wariantów analizowanych zmiennych zaprezentowano na rysunku 1. Wykreślenie, jako sposób wyjścia z rejestru przeważał (ponad 50%) w następujących wariantach cech: S1 – wykształcenie co najwyżej gimnazjalne, D0 – brak stażu pracy, W1 – wiek 18-24 lata.

Rysunek 1. Udział zarejestrowanych bezrobotnych w PUP w Szczecinie w 2013 roku i wykreślonych do końca 2014 roku według cech (%)



Źródło: opracowanie własne

## METODA BADAWCZA

Analizie poddano czas pozostawania w rejestrze osób bezrobotnych, stąd wykorzystano wybrane metody analizy trwania: model regresji (hazardu) Coxa i funkcję hazardu jako jeden z elementów tablicy trwania. Metody te można stosować w przypadku występowania danych cenzurowanych. Tak właśnie jest w omawianym badaniu. Wszystkie osoby bezrobotne zarejestrowane w 2013 roku i niewykreślone do końca 2014 roku stanowią właśnie dane cenzurowane (zakodowane jako 0).

Rozkład czasu pozostawania w rejestrze bezrobotnych może być wyrażony za pomocą funkcji hazardu  $h(t)$ . Jest ona stosunkiem funkcji gęstości prawdopodobieństwa  $f(t)$  i funkcji przetrwania  $S(t)$  [Balicki 2006]. Funkcja hazardu określa ryzyko wystąpienia określonego zdarzenia w krótkim przedziale czasu  $(t; t + \Delta t)$ , pod warunkiem, że nie wystąpiło ono do czasu  $t$  [Bieszk-Stolorz, Markowicz 2012].

Omawiane metody służą właśnie do wyznaczenia funkcji hazardu. Analizy intensywności wykreśleń bezrobotnych według ich cech dokonano przy zastosowaniu modelu Coxa [Cox, Oakes 1984], nazywanego modelem proporcjonalnego hazardu [szerzej na temat zastosowania modelu czasu trwania w naukach ekonomicznych: Bieszk-Stolorz 2013, Markowicz 2012, Landmesser 2013, Frątczak, Gach-Ciepiela, Babiker 2005, Gruszczyński 2012, Ostasiewicz 2000]. Model ten można zapisać następująco:

$$h(t; x_1, x_2, \dots, x_n) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n) \quad (1)$$

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.

Parametry tego modelu są wyznaczone przez maksymalizację funkcji częściowej wiarygodności [Cox 1972, Cox 1975, Cox, Oakes 1984]. Jest to model semiparametryczny, ponieważ funkcja hazardu jest iloczynem niewyspecyfikowanej parametrycznie bazowej funkcji hazardu i funkcji zmiennych objaśniających, dla których szacowane są parametry.

Zdarzeniem początkowym była rejestracja osoby bezrobotnej w 2013 roku, a zdarzenie końcowe to wykreślenie tej osoby z rejestru. Jeżeli wykreślenie nie nastąpiło, to taką obserwację uznano za cenzurowaną.

Zastosowany sposób definiowania (kodowania) 0-1 zmiennych (dla przekształconych zmiennych dychotomicznych), pozwala na wyznaczenie  $n - 1$  parametrów dla  $n$  zmiennych (uniknięcie współliniowości) [Bieszk-Stolorz, Markowicz 2012]. Jedną, wybraną zmienną staje się punktem odniesienia. W przeprowadzonej analizie jako grupę referencyjną w przypadku poszczególnych cech przyjęto: mężczyzn, bezrobotnych w wieku od 18 do 24 lata, bezrobotnych z wykształceniem co najwyżej gimnazjalnym, zarejestrowanych po raz pierwszy i nie posiadających doświadczenia zawodowego. Wartość  $\exp(\beta)$  jest zatem interpretowana jako ryzyko względne (*hazard ratio* – *HR*), czyli ryzyko zajścia zdarzenia dla danych dwóch jednostek. Jeśli jedna z nich należy do określonej grupy ( $j$ ), a druga do grupy referencyjnej (0), to iloraz hazardu można zapisać wzorem:

$$HR_{j/0} = \frac{h_0 \exp\left(\sum_{i=1}^n \beta_i x_{ij}\right)}{h_0 \exp\left(\sum_{i=1}^n \beta_i x_{i0}\right)} = \exp(\beta_j) \quad (2)$$

Funkcję hazardu jako jeden z elementów tablicy trwania, można zapisać wzorem [Markowicz 2012]:

$$\hat{h}_t = \frac{w_t}{\left(n_t^* - \frac{w_t}{2}\right)a_t} \quad (3)$$

gdzie:  $t$  – początek przedziału czasu trwania  $\langle t, t+1 \rangle$ ,  $n_t^*$  – liczba bezrobotnych narażonych w przedziale,  $w_t$  – liczba bezrobotnych wykreślonych w przedziale,  $a_t$  – długość przedziału czasu trwania (przyjęto miesięczne przedziały).

## WYNIKI BADAŃ

W celu identyfikacji determinant ryzyka wykreślenia z rejestru bezrobotnych zastosowano funkcje hazardu: w modelu Coxa (wskazanie ryzyka względnego) oraz w tablicach trwania (miesięczna stopa hazardu). W tabeli 2 przedstawiono wyniki estymacji modelu Coxa dla badanych bezrobotnych ogółem ( $\chi^2 = 2241,19$ ;  $p = 0,0000$ ;  $p(\beta_i) < 0,05$ ). Elementami wykorzystywanymi w niniejszej analizie są hazard względny oraz jego przedział ufności. Zostały one przedstawione na rysunku 2. Dla wszystkich zmiennych pierwszy wariant stanowi grupę referencyjną (tabela 1). Zatem należy stwierdzić, że ryzyko wykreślenia z rejestru dla kobiet jest mniejsze o 34% niż dla mężczyzn (wartość hazardu poniżej 1). Uwzględniając wiek i wykształcenie, grupą o największym ryzyku są bezrobotni najmłodsi oraz najslabiej wykształceni. Częściej wykreślani są bezrobotni rejestrujący się po raz kolejny oraz mający doświadczenie zawodowe (posiadający staż pracy).

Model regresji Coxa wyznaczono także dla kobiet ( $\chi^2 = 604,80$ ;  $p = 0,0000$ ;  $p(\beta_i) < 0,05$ ;  $p(S2) > 0,05$ ) i mężczyzn ( $\chi^2 = 1246,46$ ;  $p = 0,0000$ ;  $p(\beta_i) < 0,05$ ). Wyniki przedstawiono na rysunkach 3 i 4. Wnioski są w zasadzie podobne jak w analizie ogółem, z małymi wyjątkami. Parametr przy zmiennej S2 dla kobiet jest nieistotny, co oznacza, że ryzyko wykreślenia dla kobiet z wykształceniem zasadniczym zawodowym nie różni się istotnie od ryzyka dla kobiet z wykształceniem co najwyżej gimnazjalnym. Mężczyźni w wieku 60-64 lata (brak kobiet w tej grupie wieku) stanowią drugą grupę ryzyka wykreślenia zaraz po najmłodszych mężczyznach.

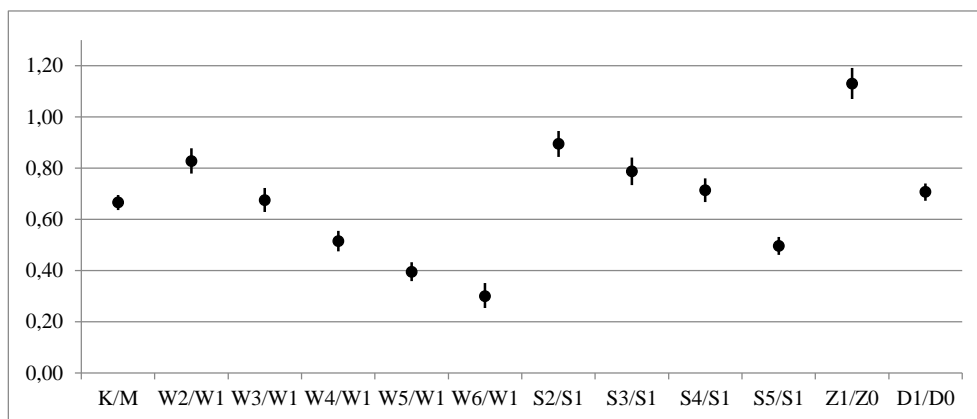
Tabela 2. Wyniki estymacji modelu (1) dla bezrobotnych ogółem

Zmienna	Ocena parametru	Błąd standardowy	Statystyka Walda	Poziom $p$	Hazard względny	Hazard względny przedział ufności 95%	
K/M	-0,4083	0,0226	326,7747	0,0000	0,6648	0,6360	0,6949
Z1/Z0	0,1216	0,0273	19,8263	0,0000	1,1294	1,0705	1,1915
W2/W1	-0,1903	0,0306	38,5943	0,0000	0,8267	0,7785	0,8779
W3/W1	-0,3945	0,0357	122,0670	0,0000	0,6740	0,6284	0,7229
W4/W1	-0,6665	0,0393	288,1717	0,0000	0,5135	0,4755	0,5546
W5/W1	-0,9315	0,0481	375,1093	0,0000	0,3940	0,3585	0,4329
W6/W1	-1,2084	0,0827	213,5862	0,0000	0,2987	0,2540	0,3512

Zmienna	Ocena parametru	Błąd standardowy	Statystyka Walda	Poziom $p$	Hazard względny	Hazard względny przedział ufności 95%	
S2/S1	-0,1126	0,0288	15,2948	0,0001	0,8935	0,8444	0,9454
S3/S1	-0,2408	0,0347	48,1913	0,0000	0,7860	0,7343	0,8413
S4/S1	-0,3394	0,0330	105,6640	0,0000	0,7122	0,6675	0,7598
S5/S1	-0,7029	0,0366	368,1418	0,0000	0,4952	0,4609	0,5320
D1/D0	-0,3481	0,0242	206,9950	0,0000	0,7060	0,6733	0,7403

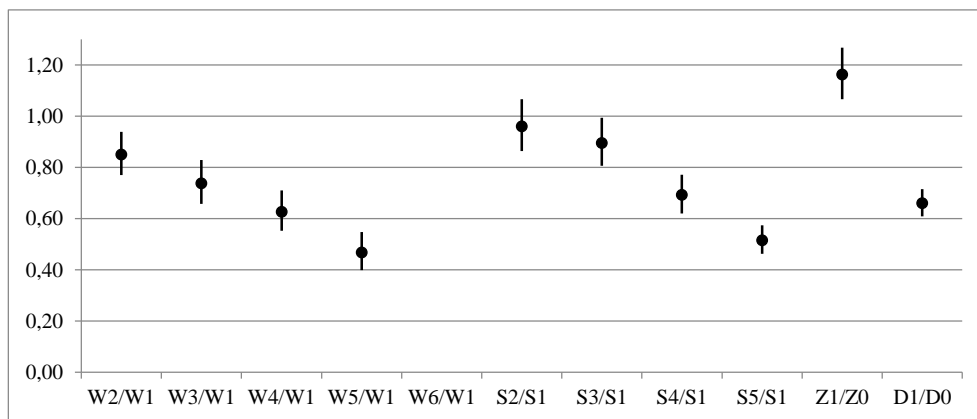
Źródło: opracowanie własne (oznaczenia jak w tabeli 1)

Rysunek 2. Hazard względny wykreślenia z rejestru dla bezrobotnych ogółem



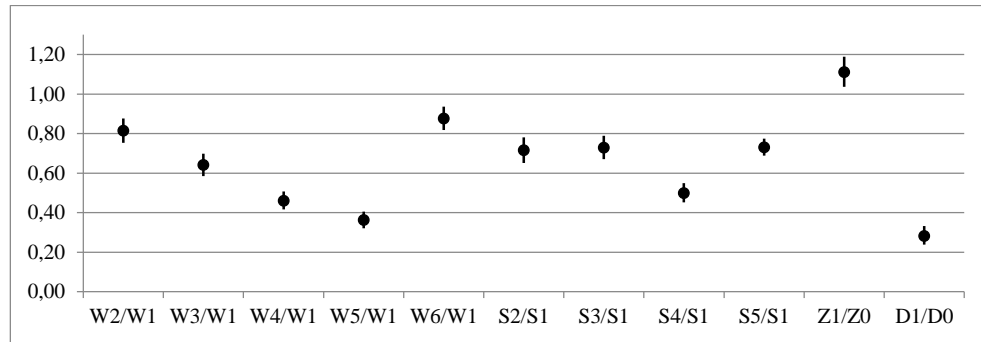
Źródło: opracowanie własne

Rysunek 3. Hazard względny wykreślenia z rejestru dla bezrobotnych kobiet



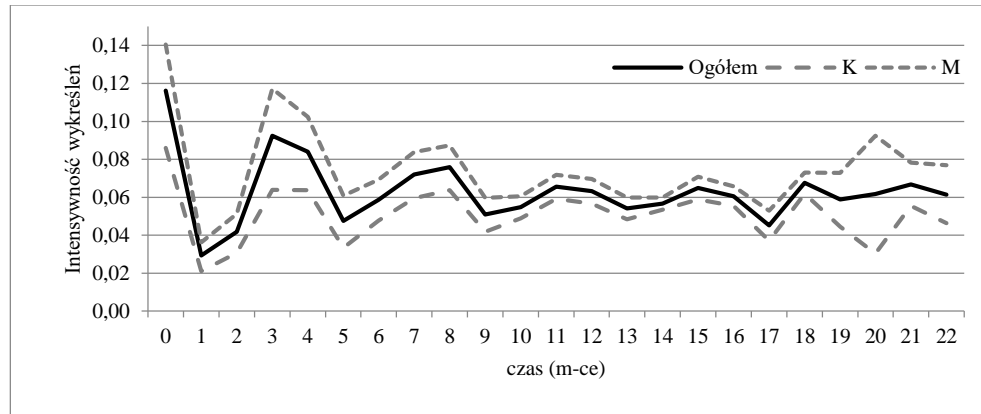
Źródło: opracowanie własne

Rysunek 4. Hazard względny wykreślenia z rejestru dla bezrobotnych mężczyzn



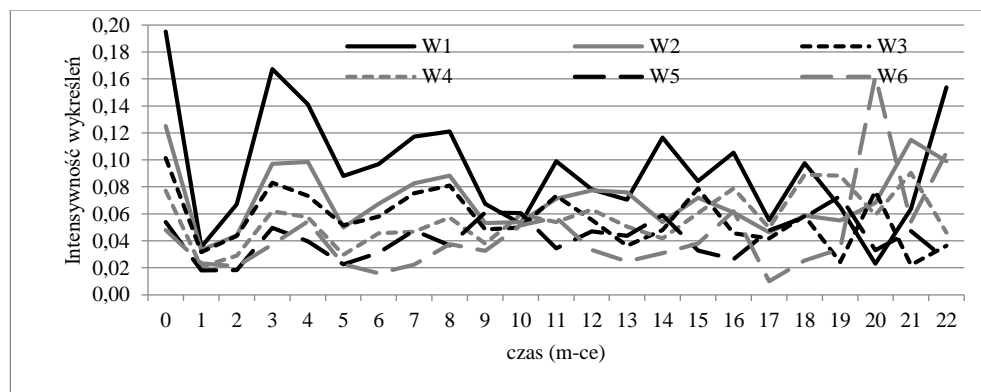
Źródło: opracowanie własne

Rysunek 5. Intensywność wykreśleń z rejestru bezrobotnych ogółem oraz według płci



Źródło: opracowanie własne

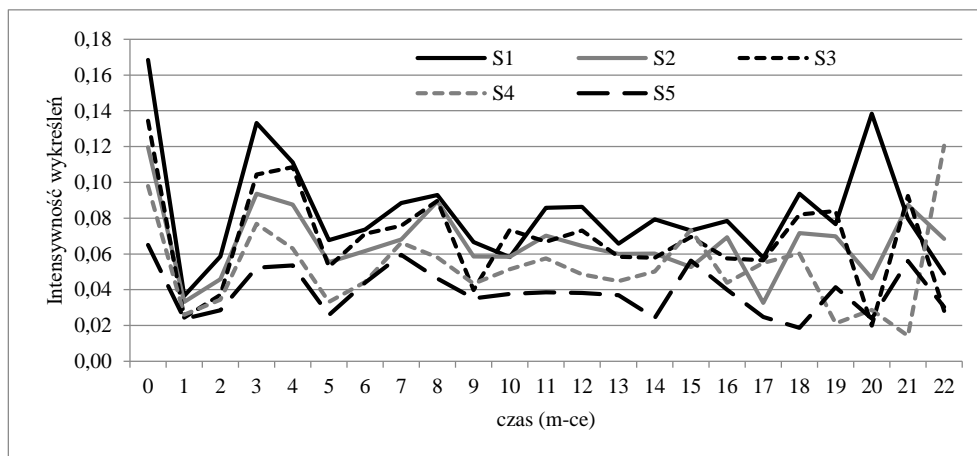
Rysunek 6. Intensywność wykreśleń z rejestru bezrobotnych według wieku



Źródło: opracowanie własne

Drugim etapem badania była budowa tablic trwania w bezrobociu dla bezrobotnych ogółem oraz bezrobotnych według płci, wieku, wykształcenia, kolejności rejestracji w urzędzie pracy oraz posiadania stażu pracy. Utworzono tablice dla miesięcznych przedziałów czasu. Wyznaczona funkcja hazardu pozwala na porównanie ryzyka wykreślenia danej grupy bezrobotnych z inną grupą oraz na obserwację zmian ryzyka wykreślenia wraz z upływem czasu od rejestracji. Wyniki przedstawiono na rysunkach 5-7 (na osi czasu podano początek przedziału czasu trwania bezrobocia). Szczególnie wysoka intensywność wykreśleń miała miejsce w pierwszym (przedział 0-1) i czwartym miesiącu trwania bezrobocia. Jest to początek okresu posiadania statusu bezrobotnego i okres wezwań do urzędu pracy. Niestawianie się na wezwania już na początku trwania bezrobocia sugeruje brak gotowości do podjęcia pracy. Większe ryzyko wykreśleń w całym badanym okresie dotyczyło mężczyzn niż kobiet. Biorąc pod uwagę wiek i wykształcenie bezrobotnych, można stwierdzić, że ryzyko wykreśleń maleje wraz ze wzrostem wieku i wykształcenia. Zatem to osoby młode i słabo wykształcone rejestrują się w urzędzie, ale nie oczekują pomocy w znalezieniu zatrudnienia czy uczestniczenia w szkoleniu, stażu. W przypadku grup wieku szczególnie wysoka intensywność wykreśleń dotyczy osób młodych (18-24 lata). W przypadku wykształcenia należy podkreślić stosunkowo wysoką intensywność osób z wykształceniem ogólnokształcącym (często wyższą niż osób z niższym wykształceniem – zasadniczym zawodowym).

Rysunek 7. Intensywność wykreśleń z rejestru bezrobotnych według wykształcenia



Źródło: opracowanie własne

Funkcje hazardu wyznaczono również dla grup bezrobotnych według kolejności zgłoszenia oraz posiadanego doświadczenia zawodowego. Intensywność wykreśleń dla osób rejestrujących się pierwszy i kolejny raz były dość podobne.



Natomiast intensywność wykreśleń dla osób bez stażu pracy była w całym okresie wyższa niż dla osób posiadających już staż.

## PODSUMOWANIE

Przeprowadzona analiza dotyczyła osób zarejestrowanych jako bezrobotne w 2013 roku w PUP w Szczecinie, ale które nie zgłosiły się w urzędzie w wyznaczonym terminie lub nie przyjęły propozycji pracy i zostały pozbawione statusu osoby bezrobotnej. Stąd podjęto próbę identyfikacji cech wpływających na ryzyko wykreślenia z rejestru (płeć, wiek, wykształcenie, kolejność zarejestrowania i staż pracy). Wyniki badań wskazują na następujące wnioski:

1. bezrobotni wykreśleni z rejestru stanowili 41% wszystkich rejestrujących się,
2. ponad 50% udział wykreśleń dotyczył grup: z wykształceniem co najwyżej gimnazjalnym (57%), rejestrujących się pierwszy raz (56%) i w wieku 18-24 lata (54%),
3. ryzyko wykreślenia z rejestru dla kobiet jest mniejsze o 34% niż dla mężczyzn,
4. grupą według wieku o największym ryzyku są najmłodsi bezrobotni (18-24 lata)
5. grupą według wykształcenia o największym ryzyku są osoby najslabiej wykształcone (co najwyżej gimnazjalne),
6. częściej wykreślanymi są bezrobotni rejestrujący się po raz kolejny oraz mający doświadczenie zawodowe (posiadający staż pracy),
7. z analizy funkcji hazardu dla miesięcznych przedziałów czasu trwania bezrobocia (tablice trwania) wynika, że szczególnie wysoka intensywność wykreśleń miała miejsce w pierwszym i trzecim miesiącu trwania bezrobocia; jest to początek okresu posiadania statusu bezrobotnego i okres wezwań do urzędu pracy (brak gotowości do podjęcia pracy),
8. większe ryzyko wykreśleń w całym badanym okresie dotyczyło mężczyzn niż kobiet,
9. ryzyko wykreśleń maleje wraz ze wzrostem wieku i wykształcenia,
10. intensywność wykreśleń dla osób rejestrujących się pierwszy i kolejny raz były podobne
11. intensywność wykreśleń dla osób bez stażu pracy była w całym okresie wyższa niż dla osób posiadających już staż.

Zatem to przede wszystkim mężczyźni, osoby młode, słabo wykształcone i bez doświadczeń zawodowych rejestrują się w urzędzie pracy, ale nie oczekują pomocy w znalezieniu zatrudnienia czy udziału w szkoleniu, stażu.

**BIBLIOGRAFIA**

- Balicki A. (2006) Analiza przeżycia i tablice wymieralności. PWE, Warszawa.
- Bieszk-Stolorz B. (2013) Analiza historii zdarzeń w badaniu bezrobocia. Volumina.pl Daniel Krzanowski, Szczecin.
- Bieszk-Stolorz B., Markowicz I. (2012) Modele regresji Coxa w analizie bezrobocia. CeDeWu, Warszawa.
- Blanchard O., Diamond P. (1994) Ranking, Unemployment Duration and Wages. *Review of Economic Studies*, 61, 417-434.
- Cox D. R., Oakes D. (1984) *Analysis of Survival Data*. Chapman and Hall, London.
- Fratczak E., Gach-Ciepiela U., Babiker H. (2005) Analiza historii zdarzeń. Elementy teorii, wybrane przykłady zastosowań. SHG, Warszawa
- Gruszczyński M. (red.) (2012) *Mikroekonometria. Modele i metody analizy danych indywidualnych*. Oficyna Wolters Kluwer S. A., Warszawa.
- Hosmer D. W., Lemeshow S. (1999) *Applied Survival Analysis. Regression Modeling of Time to Event Data*. John Wiley & Sons, Inc., New York.
- Landmesser J. (2013) Wykorzystanie metod analizy czasu trwania do badania aktywności ekonomicznej ludności w Polsce. Wydawnictwo SGGW, Warszawa.
- Markowicz I. (2012) Statystyczna analiza żywotności firm. Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.
- Markowicz I., Stolorz B. (2009) Model proporcjonalnego hazardu Coxa przy różnych sposobach kodowania zmiennych. *Przegląd Statystyczny*, (56) 2, 106-115.
- Ostasiewicz S. red. (2000) *Metody oceny i porządkowania ryzyka w ubezpieczeniach życiowych*. Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu.
- Ustawa z dnia 20 kwietnia 2004 r. o promocji zatrudnienia i instytucjach rynku pracy (Dz. U. z 2017 r. poz. 1065) – art. 33 ust. 4.

**METHODS OF DURATION ANALYSIS IN IDENTIFICATION  
OF DETERMINANTS OF RISK OF ERASURE FROM THE  
UNEMPLOYMENT REGISTRY**

**Abstract:** The unemployment registry may be the database for analysis of the situation on the regional labour market. In this case there are important both the cause of unemployment and the deregistration reason. The lack of clearly defined cause of erasure from the labour office registry makes analyses difficult. Therefore the authors tried to identify the features of these persons influencing the risk of erasure from the unemployment registry. The aim of the paper is assessment of the risk of erasure from the registry and identification of its determinants. Because the time of staying in the unemployment registry was analysed, selected methods of duration analysis were applied.

**Keywords:** the Cox hazard model, duration tables, hazard function, unemployment