

## Prawdopodobieństwo wyjścia z bezrobocia rejestrowanego na przykładzie Szczecina

Beata Bieszk-Stolorz<sup>a</sup> , Krzysztof Dmytrów<sup>a</sup> 

**Streszczenie.** Celem artykułu jest ustalenie wpływu płci, wieku i wykształcenia na prawdopodobieństwo wyjścia z bezrobocia rejestrowanego w Szczecinie. Na potrzeby badania zastosowano metodę analizy trwania. W tym celu wykorzystano drzewa przeżycia zbudowane w oparciu o estymatory Kaplana-Meiera, a za kryterium podziału przyjęto statystyki testu log-rank. Analizowano dwie najczęstsze przyczyny wyrejestrowywania z urzędu pracy – podjęcie pracy oraz wykreślenie z przyczyn leżących po stronie osoby bezrobotnej. Wyodrębniono podgrupy osób podejmujących pracę oraz rezygnujących z pośrednictwa urzędu pracy w najkrótszym i najdłuższym czasie. Analizę oparto na danych indywidualnych z Powiatowego Urzędu Pracy w Szczecinie. Dotyczyły one osób zarejestrowanych w 2013 r. i obserwowanych do końca 2014 r. Obliczenia przeprowadzono w programie R, korzystając z pakietu *partykit*, funkcji *ctree*. Z badania wynika, że prawdopodobieństwo wyłączenia z ewidencji bezrobotnych z powodu podjęcia pracy zależało tylko od wieku i wykształcenia, natomiast z powodów leżących po stronie osoby bezrobotnej było warunkowane przez płeć, wiek oraz wykształcenie.

**Słowa kluczowe:** drzewa przeżycia, estymator Kaplana-Meiera, test log-rank, bezrobocie rejestrowane

### The probability of exit from registered unemployment, based on the example of Szczecin

**Abstract.** The aim of the paper is to determine the influence of sex, age and education on the probability of exit from the registered unemployment in Szczecin. For the purposes of the study, the authors employed the survival analysis method, where they used survival trees built on the basis of the Kaplan-Meier estimators and adopted the statistic of the log-rank test as the splitting criterion. The research analysed the two most frequent reasons for deregistration, namely starting a job and the unemployed person's failure to meet the conditions for being registered as unemployed. In addition, the study extracted subgroups of persons whom it took shortest and longest to start a job or deregister from a labour office. The analysis was based on the microdata from the Powiat Labour Office in Szczecin concerning persons who registered as unemployed in 2013 and were monitored until the end of 2014. The calculations were made in the R computer programme, using the *partykit* package and the *ctree* function. The research demonstrated that the probability of deregistration from the unemployment register because of finding a job depends solely on the age and education of the unemployed person, while the probability of getting removed from the unemployment register – on the two former determinants plus sex.

**Keywords:** survival trees, Kaplan-Meier estimator, log-rank test, registered unemployment

JEL: C38, C41, J64

---

<sup>a</sup> Uniwersytet Szczeciński, Wydział Ekonomii, Finansów i Zarządzania.

Z danych statystycznych gromadzonych przez powiatowe urzędy pracy wynika, że istnieje kilkadziesiąt powodów wyłączenia z ewidencji bezrobotnych. Należą do nich np.: przejście na rentę lub emeryturę, wyjazd za granicę na okres dłuższy niż 30 dni, zmiana miejsca zamieszkania, przyznanie zasiłku przedemerytalnego, śmierć. Najczęstszą przyczyną jest podjęcie pracy. W latach 2008–2018 odsetek osób w Polsce, które z tego powodu wyrejestrowały się z urzędu pracy, wzrósł z 38% do ponad 50% (tabl. 1). Wzrost ten wiąże się m.in. z poprawą sytuacji gospodarczej w kraju po kryzysie finansowym. Drugą w kolejności przyczyną jest wykreślenie z rejestru bezrobotnych z powodu braku gotowości do pracy. Dzieje się tak w przypadku odmowy podjęcia proponowanego zatrudnienia lub niestawiennictwa w urzędzie pracy w wyznaczonym terminie. W latach 2008–2018 było to od prawie 19% do 32% wszystkich wyrejestrowań w Polsce. Od 2011 r. widać wyraźny spadek tego odsetka, co wiąże się również z polepszającą się koniunkturą na rynku pracy.

**TABL. 1. WYREJESTROWANI BEZROBOTNI W POLSCE**

L a t a	Ogółem w tys.	Wyłączeni z ewidencji		
		z powodu podjęcia pracy	ze względu na brak gotowości do podjęcia pracy	pozostali
		w %		
2008 .....	2749,4	38,26	31,81	29,93
2009 .....	2664,8	37,99	31,44	30,57
2010 .....	2979,8	39,71	29,33	30,96
2011 .....	2563,5	44,11	32,29	23,60
2012 .....	2500,8	43,84	29,63	26,53
2013 .....	2688,2	46,98	27,01	26,01
2014 .....	2786,2	46,15	26,49	27,36
2015 .....	2630,6	48,81	22,27	28,92
2016 .....	2398,0	50,40	19,57	30,03
2017 .....	2168,3	49,06	18,58	32,36
2018 .....	1774,8	50,69	19,15	30,16

Ź r ó d ł o: opracowanie własne na podstawie: GUS (2011, 2014, 2016, 2018) i Banku Danych Lokalnych GUS.

Część osób bezrobotnych nie informuje urzędu o znalezieniu pracy, sądząc, że obowiązek ten należy do pracodawcy. Formalnie osoby podejmujące pracę powinny to zrobić w ciągu 7 dni. Dopóki bezrobotny nie zostanie wyłączony z ewidencji, urząd płaci za niego składki. W celu zmniejszenia skali tego zjawiska odzyskanie statusu bezrobotnego, a co za tym idzie – prawa do ubezpieczenia zdrowotnego i zasiłku, jest utrudnione dla osób wykreślonych z rejestru z własnej winy.

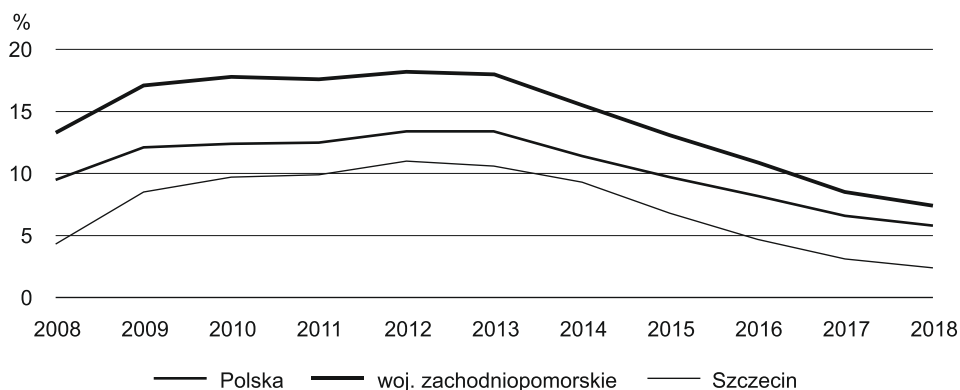
Celem badania było ustalenie wpływu płci, wieku i wykształcenia na prawdopodobieństwo wyjścia z bezrobocia rejestrowanego w Szczecinie. Analizowano dwie najczęstsze przyczyny wyrejestrowania z urzędu pracy: podjęcie pracy oraz wykreślenie z rejestru z przyczyn leżących po stronie osoby bezrobotnej. Ze względu na występowanie obserwacji cenzurowanych w badaniu wykorzystano metody analizy trwania.

## BEZROBOCIE W POLSCE I SZCZECINIE

Wybrane metody badawcze wymagały przeprowadzenia badania dla obszaru jednorodnego z dużą ilością danych indywidualnych. Obszar jednorodny charakteryzuje się tym, że osoby znajdujące się na rynku pracy podlegają działaniu takich samych czynników. Szczecin, liczący ponad 400 tys. mieszkańców, spełnia te warunki, a Powiatowy Urząd Pracy (PUP) w Szczecinie udostępnił dla celów badawczych pełne dane o osobach bezrobotnych.

Sytuacja na rynku pracy w Polsce w ostatnich latach poprawiła się, o czym świadczy malejąca stopa bezrobocia (wykr. 1). Największą stopę bezrobocia – 13,4% – odnotowano w latach 2012 i 2013. W 2018 r. spadła do 5,8%. W woj. zachodniopomorskim była w tych latach wyższa. W 2012 r. osiągnęła maksimum – 18,2%, a w 2018 r. minimum – 7,4%. Sytuacja w Szczecinie była lepsza niż w Polsce ogółem. Stopa bezrobocia spadła tam z 11% w 2012 r. do 2,4% w 2018 r., co ma związek z sytuacją na rynku światowym, a szczególnie europejskim. Lata 2008–2012 były czasem globalnego kryzysu finansowego, a co za tym idzie globalnej dekoniunktury także na rynku pracy. Odzwierciedla to wysoka stopa bezrobocia. W okresie wychodzenia z kryzysu sytuacja na rynku pracy się poprawiła.

**WYKR. 1. STOPA BEZROBOCIA REJESTROWANEGO W POLSCE, W WOJEWÓDZTWIE ZACHODNIOPOMORSKIM I W SZCZECINIE**



Jak wskazują badania, procesy zachodzące na polskim rynku pracy upodabniają go do rynków słowackiego i węgierskiego (Hadaś-Dyduch, Pietrzak i Balcerzak, 2016). Wpływ na bezrobocie ma polityka społeczna danego kraju, realizowana m.in. przez urzędy pracy. Działania aktywizujące skierowane do grup osób zagrożonych bezrobociem wpływają pozytywnie na rynek pracy, ale są one krótkotrwałe, zróżnicowane terytorialnie i nie zawsze przekładają się na spadek stopy bezrobocia (Bieszk-Stolorz i Dmytrów, 2018, 2019). Natomiast rozbudowany system zasiłków dla osób bezrobotnych może wydłużać czas poszukiwania pracy (Bieszk-Stolorz i Markowicz, 2015).

### DANE WYKORZYSTANE W BADANIU

W badaniu wykorzystano anonimowe dane indywidualne 22078 osób bezrobotnych zarejestrowanych w 2013 r. w PUP w Szczecinie i obserwowanych do końca 2014 r. Jest to początek spadku stopy bezrobocia sygnalizujący wychodzenie z globalnego kryzysu<sup>1</sup>. Zdarzeniem kończącym obserwację każdej jednostki był moment wyrejestrowania z urzędu. Analizowano czas  $T$  od momentu zarejestrowania do momentu wyrejestrowania z określonego powodu. Rozpatrywano dwa rodzaje zdarzeń kończących obserwację: podjęcie pracy i wykreślenie z przyczyn leżących po stronie osoby bezrobotnej. Wyrejestrowania z innych powodów oraz obserwacje, które nie zakończyły się zdarzeniem przed końcem 2014 r. (1856 obserwacji), uznano za obserwacje cenzurowane. W badanym okresie pracę podjęło prawie 44% zarejestrowanych bezrobotnych. Nie mniej liczną grupę (prawie 41%) stanowiły osoby, które nie zgłosiły się do urzędu w wyznaczonym terminie lub nie przyjęły propozycji pracy (wykreślone z rejestru bezrobotnych). Liczebność wszystkich grup przedstawia tabl. 2. Wstępna analiza danych wskazuje, że wraz ze wzrostem poziomu wykształcenia rósł odsetek osób podejmujących pracę, natomiast malał odsetek osób wykreślonych z rejestru bezrobotnych. Mężczyźni częściej niż kobiety byli wykreślani z rejestru, natomiast kobiety częściej podejmowały pracę. W grupie najmłodszych osób (18–24 lata) ponad 54% bezrobotnych zrezygnowało z pośrednictwa urzędu pracy, a nieco ponad 36% wyłączono z ewidencji ze względu na podjęcie pracy. Generalnie, od grupy wieku 25–34 lata im osoby były starsze, tym wyższy był odsetek osób wyrejestrowanych z urzędu z innych przyczyn. Należy zwrócić uwagę, że od 1 stycznia 2013 r. zaczęto stopniowo podwyższać wiek emerytalny kobiet i mężczyzn do 67 lat, w związku z czym wśród osób zarejestrowanych znalazły się kobiety, które ukończyły 60 lat, oraz mężczyźni, którzy ukończyli 65 lat.

<sup>1</sup> Planowana jest kontynuacja badań z wykorzystaniem drzew przeżycia dotyczących późniejszego okresu przy stabilizującej się sytuacji na rynku pracy.

**TABL. 2. BEZROBOTNI W SZCZECINIE (stan na XII 2014 r.)**

Wyszczególnienie	Ogółem		W tym			
			podejmujący pracę		wykreśleni z rejestru bezrobotnych	
		w %		w %		w %
<b>O g ó ł e m</b> .....	<b>22078</b>	<b>100,00</b>	<b>9678</b>	<b>43,84</b>	<b>8965</b>	<b>40,61</b>
<b>Płeć</b>						
Kobiety (K) .....	9770	44,25	4836	49,50	3264	33,41
Mężczyźni (M) .....	12308	55,75	4842	39,34	5701	46,32
<b>Wiek w latach</b>						
18–24 ( $W_1$ ) .....	4148	18,79	1506	36,31	2257	54,41
25–34 ( $W_2$ ) .....	7356	33,32	3614	49,13	2966	40,32
35–44 ( $W_3$ ) .....	4259	19,29	1869	43,88	1734	40,71
45–54 ( $W_4$ ) .....	3497	15,84	1642	46,95	1214	34,72
55–59 ( $W_5$ ) .....	2185	9,90	837	38,31	629	28,79
60–65 ( $W_6$ ) .....	633	2,87	210	33,18	165	26,07
<b>Wykształcenie</b>						
Co najwyżej gimnazjalne ( $S_1$ ) ....	5123	23,20	1410	27,52	2932	57,23
Zasadnicze zawodowe ( $S_2$ ) .....	5016	22,72	1968	39,23	2220	44,26
Średnie ogólnokształcące ( $S_3$ ) ...	2859	12,95	1226	42,88	1223	42,78
Średnie zawodowe ( $S_4$ ) .....	4086	18,51	1943	47,55	1415	34,63
Wyższe ( $S_5$ ) .....	4994	22,62	3131	62,70	1175	23,53

Ź r ó d ł o: opracowanie własne.

Przez podjęcie pracy (nazywane dalej pracą) rozumie się podjęcie pracy lub innego zatrudnienia, pracy subsydiowanej przez urząd (m.in. w ramach: robót publicznych, prac interwencyjnych, utworzonego przez pracodawcę dodatkowego miejsca pracy z tytułu udzielonej pożyczki, dofinansowania wynagrodzenia za zatrudnienie bezrobotnego, który ukończył 50. rok życia) lub działalności gospodarczej (podjęcie pozarolniczej działalności gospodarczej, przyznanie jednorazowo środków na podjęcie działalności gospodarczej, rozpoczęcie działalności gospodarczej ze środków PFRON).

Z wykreśleniem z rejestru bezrobotnych mamy do czynienia, gdy osoba bezrobotna nie wykazała chęci współpracy z urzędem i została z własnej winy lub na własną prośbę wyłączona z ewidencji. Może to być spowodowane m.in.:

- odmową, bez uzasadnionej przyczyny, przyjęcia propozycji odpowiedniego zatrudnienia lub innej pracy zarobkowej, wykonywania prac interwencyjnych lub robót publicznych albo udziału w szkoleniu, stażu, przygotowaniu zawodowym w miejscu pracy;
- niestawieniem się w urzędzie pracy w wyznaczonym terminie i niepowiadomieniem w terminie do 5 dni o uzasadnionej przyczynie tego niestawienia;
- nieprzedstawieniem zaświadczenia o niezdolności do pracy wskutek choroby;

- odmową, bez uzasadnionej przyczyny, udziału w działaniach w ramach Programu Aktywizacja i Integracja (PAI) lub przerwaniem z własnej winy udziału w działaniach w ramach PAI;
- brakiem gotowości do podjęcia pracy przez co najmniej 10 dni;
- przerwaniem z własnej winy szkolenia, stażu, realizacji indywidualnego planu działań;
- wnioskiem bezrobotnego o wykreślenie z ewidencji.

Pozostałe powody wyrejestrowania są mniej liczne i jak wykazały wcześniejsze badania, każdy z nich z osobna miał marginalny wpływ na prawdopodobieństwo wyłączenia z ewidencji. Bezrobotnych podzielono na podgrupy ze względu na płeć, wiek i wykształcenie, będące ważnymi determinantami wyjścia z bezrobocia (Bieszk-Stolorz, 2017).

## METODA BADANIA

Analizę trwania coraz częściej stosuje się do badania zjawisk społeczno-ekonomicznych. Za pomocą tej metody, powszechnie stosowanej w demografii i naukach medycznych do badania czasu trwania życia ludzkiego, można analizować czas trwania zjawisk ekonomicznych. Coraz częściej stosuje się ją do analizy rynku pracy. Można w ten sposób badać aktywność ekonomiczną ludności (Landmesser, 2013) lub czas trwania bezrobocia (Bieszk-Stolorz, 2013). Analizie podlega czas trwania jednostki w danym stanie (czas działania firmy, czas bezrobocia, czas spłaty kredytu), który jest zmienną losową  $T$ . Podstawą tego typu analizy jest funkcja trwania, nazywana również funkcją przeżycia, zdefiniowana następująco:

$$S(t) = P(t < T) = 1 - F(t) \quad (1)$$

gdzie:

$T$  – czas trwania zjawiska,

$F(t)$  – dystrybuanta rozkładu zmiennej losowej  $T$ .

Drugą funkcją stosowaną w analizie trwania jest funkcja hazardu. Opisuje ona intensywność zajścia zdarzenia w momencie  $t$  pod warunkiem przetrwania do czasu  $t$  i jest zdefiniowana następująco (Kleinbaum i Klein, 2005):

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad (2)$$

Do najczęściej stosowanych estymatorów funkcji trwania należy estymator Kaplana-Meiera (Kaplan i Meier, 1958):

$$\hat{S}(t) = \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{n_j}\right) \quad (3)$$

gdzie:

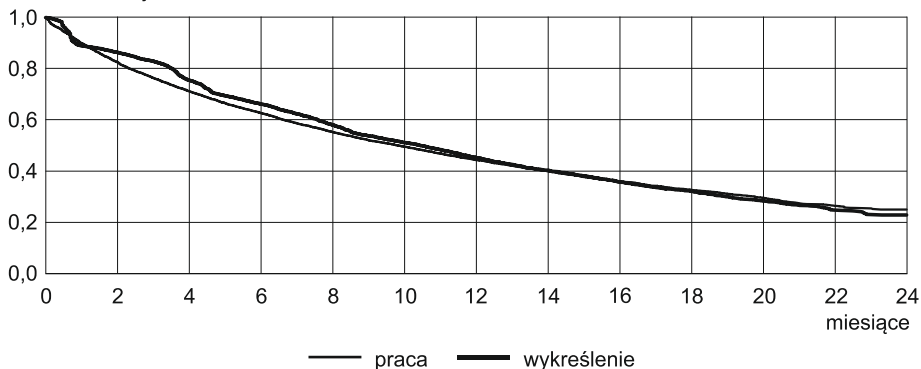
$d_j$  – liczba zdarzeń w momencie  $t_j$ ,

$n_j$  – liczba jednostek zagrożonych do momentu  $t_j$ .

Funkcja trwania informuje o tym, jakie jest prawdopodobieństwo niezajścia zdarzenia co najmniej do czasu  $t$ . Czasami uzasadnione jest analizowanie dystrybuanty  $F(t)$ , która wyraża prawdopodobieństwo, że zdarzenie zajdzie najpóźniej do czasu  $t$ . Jeżeli zdarzeniem jest wykreślenie z rejestru bezrobotnych, to estymator funkcji trwania informuje o tym, jakie jest prawdopodobieństwo pozostania w rejestrze z powodu niewykreślenia, a estymator dystrybuanty pozwala określić, jakie jest prawdopodobieństwo wykreślenia z rejestru bezrobotnych. W niniejszym przypadku  $d_j$  to liczba wyrejestrowań z urzędu pracy z danej przyczyny w momencie  $t_j$  (podjęcie pracy lub wykreślenie z powodu braku gotowości do podjęcia pracy). W przypadku wyrejestrowania ze względu na podjęcie pracy pożądane jest, aby krzywe przeżycia były położone nisko, natomiast w przypadku wykreślenia z rejestru bezrobotnych – odwrotnie. Na wyk. 2 przedstawiono przebieg estymatorów Kaplana-Meiera dla osób wyrejestrowanych z powodu podjęcia pracy i wykreślonych z przyczyn leżących po stronie osób bezrobotnych. W tym przypadku obie krzywe są do siebie zbliżone, co świadczy o tym, że prawdopodobieństwo pozostania w rejestrze dla obu rodzajów wyrejestrowań jest podobne. Na podstawie funkcji przeżycia wyznacza się kwantyle czasu trwania. Są one określone w punktach, w których krzywa przeżycia przyjmuje wartości odpowiednio 0,75, 0,50 i 0,25 (Aalen, Borgan i Gjessing, 2008).

**WYKR. 2. ESTYMATORY KAPLANA-MEIERA NIEWYREJESTROWANIA ZE WZGLĘDU NA NIEPODJĘCIE PRACY I NIEWYKREŚLENIA Z REJESTRU BEZROBOTNYCH**

prawdopodobieństwo  
pozostania w rejestrze



Źródło: opracowanie własne.

Analizowane zbiorowości można dzielić na grupy ze względu na badane cechy, oszacować funkcje trwania dla każdej z tych grup i zbadać istotność różnic między nimi. Ponieważ nie są znane rozkłady czasu trwania, stosuje się testy nieparametryczne oparte na porządku rangowym czasów trwania. Nie ma niestety powszechnie akceptowanych metod wyboru testu w danej sytuacji. Większość z nich daje rzetelne wyniki tylko przy dużych próbach, natomiast efektywność testów przy małych próbach jest mniej poznana.

Do porównania dwóch krzywych trwania często stosuje się test log-rank (Kleinbaum i Klein, 2005). Służy on do weryfikacji hipotezy  $H_0: S_1(t) = S_2(t)$  o równości krzywych trwania wyznaczonych dla obu grup. Hipoteza alternatywna  $H_1: S_1(t) \neq S_2(t)$  mówi o tym, że krzywe przeżycia nie są takie same. Statystyka testowa jest porównywana przy danym poziomie istotności z rozkładem chi-kwadrat o jednym stopniu swobody. Test ten ma największą moc, gdy różnica między funkcjami hazardu dla pojedynczych podgrup jest stała w czasie (Landmesser, 2013). Wstępna analiza z wykorzystaniem funkcji  $\ln(-\ln S(t))$  oraz pewne ograniczenia wynikające z założeń dla innych testów potwierdziły zasadność wykorzystania w badaniu testu log-rank.

Bardzo przydatnym narzędziem służącym do dzielenia badanej zbiorowości na jednorodny grupy ze względu na kształtowanie się krzywych przeżycia są drzewa przeżycia (De Rose i Pallara, 1997; Mikulec i Misztal, 2018; von Wangenheim, 2013), będące podgrupą metod zwanych drzewami wnioskowania warunkowego. Drzewa te opierają się na podziale binarnym. Zyskują one w ostatnim czasie dużą popularność w porównaniu z innymi metodami (np. analizą dyskryminacyjną), ponieważ wymagają mniej założeń i mogą poradzić sobie z różnymi strukturami danych (Al-Nachawati, Ismail i Almohisen, 2010; Bou-Hamad i in., 2009; LeBlanc i Crowley, 1993; Zhou i McArdle, 2015). Budowa każdego drzewa wiąże się z dwoma aspektami (Cappelli i Zhang, 2007):

- podziałem danych, czyli rozrostem drzewa;
- przycinaniem drzewa, aby zmniejszyć jego rozmiar i tym samym zwiększyć czytelność wyników.

Podział danych polega na wyodrębnieniu jednorodnych grup pod względem badanych współzmiennych. Kryterium podziału jest chęć otrzymania maksymalnie homogenicznych grup, czyli jak najmniejszej heterogeniczności – „zanieczyszczenia” (*impurity*). W przypadku drzew przeżycia heterogeniczność wyodrębnionych grup bada się za pomocą testu log-rank. Podział następuje do momentu, aż zostanie osiągnięte kryterium zatrzymania. Podział danych powoduje często niekontrolowany rozrost drzewa prowadzący do jego „przeuczenia”. Stąd bierze się konieczność jego przycinania. Podział przestaje następować, jeżeli empiryczny poziom istotności statystyki testu log-rank przekracza założoną wartość. Nie zawsze jest to podejście efektywne dla dużej liczby obserwacji. Innym kryterium jest ustalenie minimalnej liczebności grupy, przy której może nastąpić



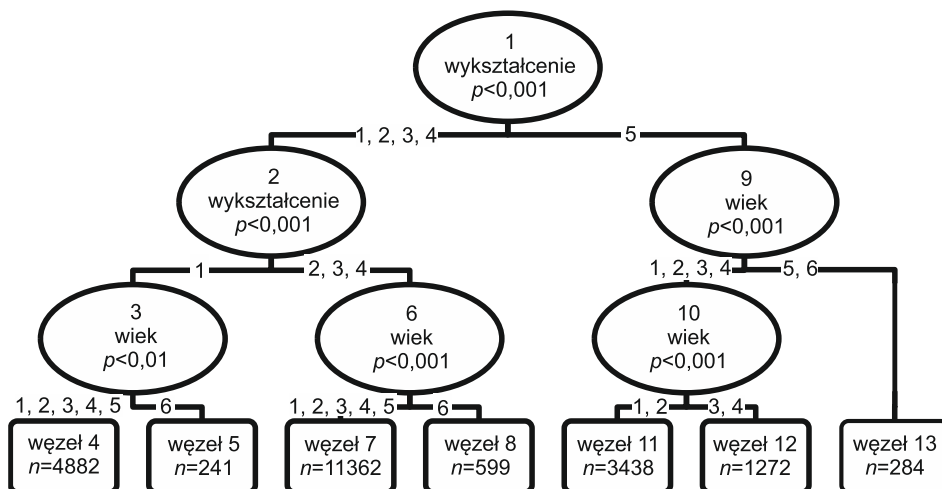
podział, oraz minimalnej liczebności grupy w węźle końcowym. Można też ustalić maksymalną głębokość drzewa (Mudunuru, 2016).

Zaprezentowane w artykule drzewa przeżycia zbudowano, korzystając z funkcji *ctree* w pakiecie *partykit* w programie R. Były to drzewa wnioskowania warunkowego. Każdego obserwowanego bezrobotnego opisano za pomocą następującego trypletu:  $\{y_i, \delta_i, \mathbf{x}_i\}$ , gdzie  $y_i$  oznacza czas zarejestrowania,  $\delta_i$  – czy obserwacja jest cenzurowana, czy nie (1 – obserwacja niecenzurowana, 0 – cenzurowana), a wektor  $\mathbf{x}_i$  zawiera trzy współzmiennie: płeć, wiek oraz wykształcenie. Czas zarejestrowania był zmienną ilościową ciągłą, cenzurowanie – zmienną ilościową skokową o dwóch wartościach (0 oraz 1), a pozostałe zmienne – zmiennymi jakościowymi. W funkcji *ctree\_control* zmieniono kilka domyślnych parametrów. Ze względu na dużą liczebność zbiorowości za pomocą parametru *mincriterion* poziom istotności ustalono na 0,01. Natomiast maksymalną głębokość drzewa ustalono za pomocą parametru *maxdepth* na trzy poziomy (kryterium zatrzymania).

### ANALIZA CZASU WYREJESTROWANIA ZE WZGLĘDU NA PODJĘCIE PRACY

Analizę przeprowadzono w dwóch etapach. Najpierw wyodrębniono jednorodne grupy bezrobotnych według prawdopodobieństwa wyrejestrowania ze względu na podjęcie pracy.

SCHEMAT 1. DRZEWO PRZEŻYCIA DLA WYREJESTROWANIA  
ZE WZGLĘDU NA PODJĘCIE PRACY



Na schemacie 1 widać, że na żadnym poziomie płeć nie stanowiła istotnego kryterium podziału ze względu na szybkość wychodzenia z bezrobocia i podjęcia pracy. W pierwszej kolejności bezrobotni zostali podzieleni według poziomu wykształcenia – na osoby z wykształceniem wyższym i pozostałe. Następnie dokonano podziału na trzy grupy: osoby z wykształceniem co najwyżej gimnazjalnym, średnim lub zasadniczym zawodowym oraz wyższym. W dalszej kolejności osoby były dzielone ze względu na grupę wieku. Ostatecznie otrzymano siedem węzłów końcowych (zestawienie 1).

**ZESTAWIENIE 1.**  
**JEDNORODNE GRUPY BEZROBOTNYCH WYREJESTROWANYCH**  
**ZE WZGLĘDU NA PODJĘCIE PRACY**

Wyszczególnienie	Numer węzła końcowego						
	4	5	7	8	11	12	13
Wykształcenie .....	$S_1$	$S_1$	$S_2-S_4$	$S_2-S_4$	$S_5$	$S_5$	$S_5$
Wiek .....	$W_1-W_5$	$W_6$	$W_1-W_5$	$W_6$	$W_1, W_2$	$W_3, W_4$	$W_5, W_6$

Źródło: jak przy wykr. 1.

Najwolniej ze względu na podjęcie pracy były wyrejestrowywane osoby w wieku co najmniej 60 lat z wykształceniem co najwyżej gimnazjalnym (węzeł 5). Dla tej grupy mediana była dłuższa niż okres obserwacji, czyli 24 miesiące (tabl. 3). Najszybciej wyrejestrowywano osoby młode (do 34. roku życia) z wykształceniem wyższym (węzeł 11). Mediana czasu wychodzenia z bezrobocia dla tych osób była najniższa i wynosiła ok. 5 miesięcy.

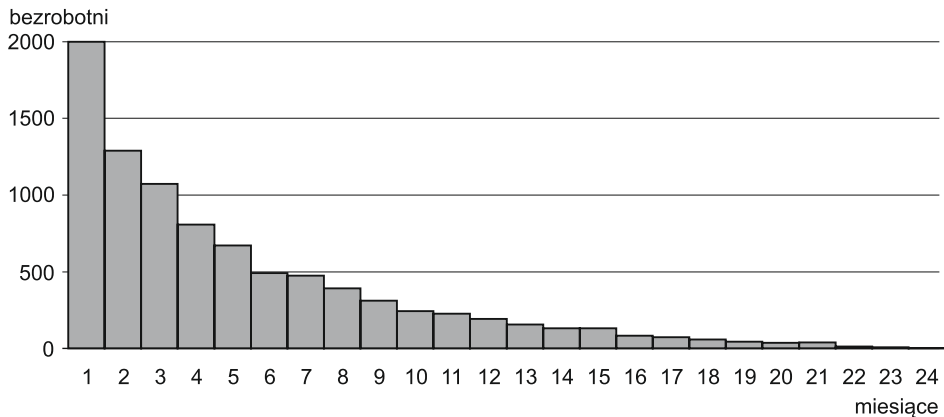
**TABL. 3. KWARTYLE CZASU DO PODJĘCIA PRACY**

Numer węzła końcowego	Kwartył pierwszy	Mediana	Kwartył trzeci
	w miesiącach		
4 .....	6,12	18,44	–
5 .....	14,53	–	–
7 .....	3,22	9,62	22,53
8 .....	10,76	22,06	–
11 .....	1,97	4,96	11,05
12 .....	2,45	6,67	15,87
13 .....	4,51	16,64	–

Źródło: jak przy wykr. 1.

Rozkład czasu wyrejestrowania ze względu na podjęcie pracy przedstawiono na wykr. 3. Rozkład czasu wychodzenia z bezrobocia rejestrowanego i podjęcia pracy był skrajnie prawostronnie asymetryczny. Najwięcej osób (1999 r. – 20,65%) podejmowało zatrudnienie w ciągu pierwszego miesiąca od zarejestrowania.

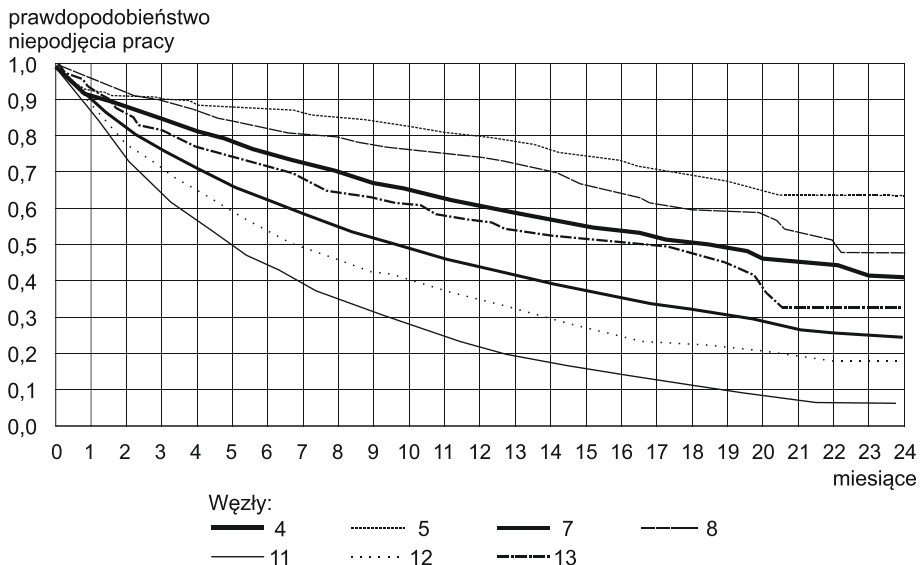
**WYKR. 3. ROZKŁAD CZASU POZOSTAWANIA W EWIDENCJI BEZROBOTNYCH DO MOMENTU WYREJESTROWANIA ZE WZGLĘDU NA PODJĘCIE PRACY**



Źródło: opracowanie własne.

Na wyk. 4 przedstawiono estymatory Kaplana-Meiera dla grup wyznaczonych przez węzły końcowe.

**WYKR. 4. ESTYMATORY KAPLANA-MEIERA NIETYREJESTROWANIA ZE WZGLĘDU NA NIEPODJĘCIE PRACY**



Źródło: opracowanie własne.

Istotność różnic między krzywymi przeżycia niepodjęcia zatrudnienia zbadano za pomocą testu log-rank (tabl. 4). Nie ma podstaw do odrzucenia hipotezy o równości krzywych przeżycia dla następujących par węzłów końcowych: 5 i 8 oraz 4 i 13. Pierwsza para to osoby mające co najmniej 60 lat z wykształceniem co najwyżej gimnazjalnym oraz z wykształceniem zasadniczym zawodowym, średnim ogólnokształcącym lub średnim zawodowym. Są to grupy osób o najmniejszym prawdopodobieństwie podjęcia pracy. Wynika z tego, że w przypadku osób najstarszych o szybkim znalezieniu pracy decydowało posiadanie wykształcenia wyższego. Druga para to osoby w wieku do 59 lat z wykształceniem nie wyższym oraz osoby w wieku co najmniej 55 lat z wykształceniem średnim zawodowym lub wyższym.

**TABL. 4. TEST LOG-RANK DLA PODJĘCIA PRACY**

Numer węzła końcowego	5	7	8	11	12	13
4 .....	-4,0809 (0,0000)	13,4463 (0,0000)	4,5747 (0,0000)	24,8693 (0,0000)	13,9294 (0,0000)	1,1290 <sup>a</sup> (0,2589)
5 .....	x	7,8012 (0,0000)	1,4971 <sup>a</sup> (0,1344)	11,9572 (0,0000)	9,5367 (0,0000)	4,2782 (0,0000)
7 .....	x	x	-10,7772 (0,0000)	16,7728 (0,0000)	5,3692 (0,0000)	-3,2969 (0,0010)
8 .....	x	x	x	17,2561 (0,0000)	12,6752 (0,0000)	4,0028 (0,0000)
11 .....	x	x	x	x	-5,5753 (0,0000)	-8,2552 (0,0000)
12 .....	x	x	x	x	x	-5,3790 (0,0000)

a Brak istotności na przyjętym poziomie 0,01.

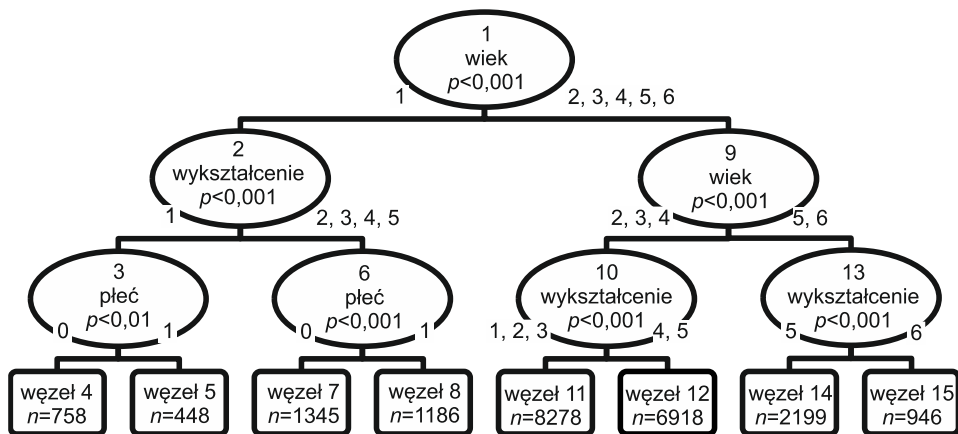
U w a g a. W nawiasie podano wartości *p*.

Ź r ó d ł o: jak przy wykr. 1.

## ANALIZA CZASU DO WYKREŚLENIA Z REJESTRU BEZROBOTNYCH

Drugi etap badania polegał na wyodrębnieniu jednorodnych grup bezrobotnych ze względu na prawdopodobieństwo wykreślenia z rejestru osób bezrobotnych. Jak widać na schemacie 2, osoby wykreślane z rejestru w pierwszej kolejności były dzielone według wieku – na osoby do 24 lat i starsze. Osoby najmłodsze były w dalszej kolejności dzielone według poziomu wykształcenia, a następnie według płci. Osób najstarszych nie dzielono ani według płci, ani według poziomu wykształcenia.

**SCHEMAT 2. DRZEWO PRZEŻYCIA DLA WYKREŚLENIA Z REJESTRU BEZROBOTNYCH**



Źródło: opracowanie własne.

W każdej grupie więcej osób było wykreślanych z rejestru bezrobotnych w ciągu pierwszego miesiąca (wartości krzywej przeżycia mały najszybciej). Drugą najliczniejszą grupę stanowiły osoby wykreślane w ciągu czwartego miesiąca. Ostatecznie otrzymano osiem węzłów końcowych (zestawienie 2). Najwolniej były wykreślane osoby w wieku 60 lat i więcej (węzeł 15). Dla tej grupy mediana była dłuższa niż okres obserwacji, czyli 24 miesiące (tabl. 5). Największe prawdopodobieństwo wykreślenia z rejestru osób bezrobotnych występowało w przypadku najmłodszych mężczyzn z wykształceniem co najmniej gimnazjalnym (węzeł 4). Mediana czasu do ich wykreślenia była niższa i wynosiła ok. 3,5 miesiąca.

**ZESTAWIENIE 2.  
JEDNORODNE GRUPY BEZROBOTNYCH WYKREŚLONYCH  
Z REJESTRU BEZROBOTNYCH**

Wyszczególnienie	Numer węzła końcowego							
	4	5	7	8	11	12	14	15
Wykształcenie .....	S <sub>1</sub>	S <sub>1</sub>	S <sub>2</sub> -S <sub>5</sub>	S <sub>2</sub> -S <sub>5</sub>	S <sub>1</sub> -S <sub>3</sub>	S <sub>4</sub> , S <sub>5</sub>	S <sub>1</sub> -S <sub>5</sub>	S <sub>1</sub> -S <sub>5</sub>
Wiek .....	W <sub>1</sub>	W <sub>1</sub>	W <sub>1</sub>	W <sub>1</sub>	W <sub>2</sub> -W <sub>4</sub>	W <sub>2</sub> -W <sub>4</sub>	W <sub>5</sub>	W <sub>6</sub>
Płeć .....	M	K	M	K	K, M	K, M	K, M	K, M

Źródło: opracowanie własne.

**TABL. 5. KWARTYLE CZASU POZOSTAWANIA W EWIDENCJI  
BEZROBOTNYCH DO MOMENTU WYKREŚLENIA  
Z REJESTRU BEZROBOTNYCH**

Numer węzła końcowego	Kwartył pierwszy	Mediana	Kwartył trzeci
	w miesiącach		
4 .....	0,76	3,55	6,86
5 .....	2,00	5,56	12,29

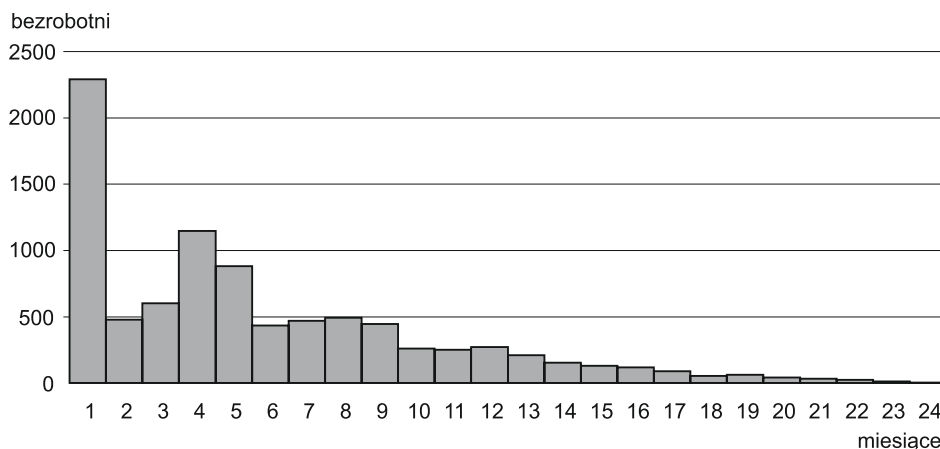
**TABL. 5. KWARTYLE CZASU POZOSTAWANIA W EWIDENCJI BEZROBOTNYCH DO MOMENTU WYKREŚLENIA Z REJESTRU BEZROBOTNYCH (dok.)**

Numer węzła końcowego	Kwartył pierwszy	Mediana	Kwartył trzeci
	w miesiącach		
7 .....	2,47	4,69	9,87
8 .....	3,39	7,53	16,68
11 .....	3,72	8,52	18,15
12 .....	6,28	15,02	–
14 .....	8,04	17,13	–
15 .....	11,79	–	–

Ź r ó d ł o: opracowanie własne.

Rozkład czasu pozostawania w ewidencji bezrobotnych do momentu wykreślenia z rejestru bezrobotnych przedstawiono na wyk. 5.

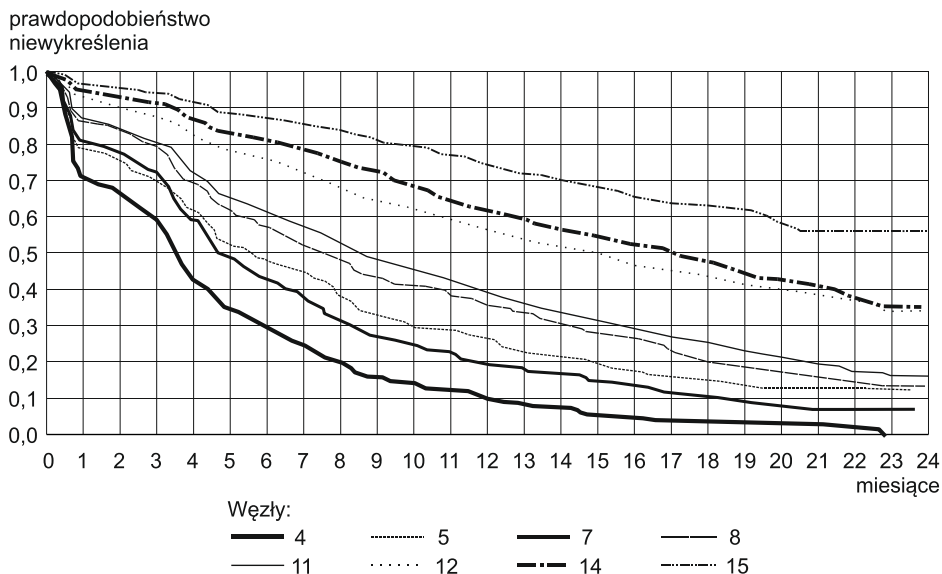
**WYKR. 5. ROZKŁAD CZASU DO WYKREŚLENIA Z REJESTRU BEZROBOTNYCH**



Ź r ó d ł o: opracowanie własne.

Warto zauważyć, że rozkład czasu do wykreślenia różni się znacznie od rozkładu wyrejestrowania ze względu na podjęcie pracy. Jest to rozkład dwumodalny – największa liczba wykreśleń z rejestru bezrobotnych (2290 osób – 25,5%) następuje w ciągu pierwszego miesiąca od zarejestrowania, następnie w czwartym miesiącu (1148 osób – 12,8%). Widać to na krzywych przeżycia na wyk. 6 przedstawiającym estymatory Kaplana-Meiera dla grup wyznaczonych przez węzły końcowe – ich przebieg małał najbardziej w pierwszym miesiącu, drugi największy spadek występował w czwartym miesiącu.

**WYKR. 6. ESTYMATORY KAPLANA-MEIERA DLA NIWYKREŚLENIA Z REJESTRU BEZROBOTNYCH**



Źródło: opracowanie własne.

Wyniki testu log-rank dla krzywych przeżycia dla wykreślenia z rejestru bezrobotnych przedstawia tabl. 6.

**TABL. 6. TEST LOG-RANK DLA WYKREŚLENIA Z REJESTRU BEZROBOTNYCH**

Numer węzła końcowego	5	7	8	11	12	14	15
4 .....	-5,7340 (0,0000)	-6,6633 (0,0000)	-11,4510 (0,0000)	-18,3279 (0,0000)	-27,5179 (0,0000)	-26,3308 (0,0000)	-24,3055 (0,0000)
5 .....	x	0,7085 <sup>a</sup> (0,4787)	-3,9738 (0,0001)	-6,7149 (0,0000)	-14,2782 (0,0000)	-15,9294 (0,0000)	-17,0176 (0,0000)
7 .....	x	x	-5,7012 (0,0000)	-11,2926 (0,0000)	-21,0416 (0,0000)	-21,1883 (0,0000)	-20,3912 (0,0000)
8 .....	x	x	x	-2,3739 <sup>a</sup> (0,0176)	-11,0911 (0,0000)	-13,3089 (0,0000)	-14,7212 (0,0000)
11 .....	x	x	x	x	-16,6979 (0,0000)	-16,5267 (0,0000)	-17,0624 (0,0000)
12 .....	x	x	x	x	x	-5,1496 (0,0000)	-9,1224 (0,0000)
14 .....	x	x	x	x	x	x	-5,5185 (0,0000)

a Brak istotności na przyjętym poziomie 0,01.

U w a g a. Jak przy tabl. 4.

Źródło: opracowanie własne.

Nie ma podstaw do odrzucenia hipotezy o równości krzywych przeżycia dla następujących par węzłów końcowych 5 i 7 oraz 8 i 11. Pierwsza para to kobiety w wieku do 24 lat z wykształceniem co najwyżej gimnazjalnym oraz mężczyźni w wieku do 24 lat z wykształceniem powyżej gimnazjalnego. Druga para to kobiety w wieku do 24 lat z wykształceniem powyżej gimnazjalnego oraz osoby (kobiety i mężczyźni) w wieku od 25 lat do 54 lat z wykształceniem co najwyżej średnim.

## PODSUMOWANIE

Na podstawie przeprowadzonego badania określono wpływ płci, poziomu wykształcenia oraz wieku na prawdopodobieństwo podjęcia pracy oraz wykreślenia z rejestru osób bezrobotnych z powodu braku gotowości do podjęcia pracy. Potwierdziło ono wyniki otrzymane za pomocą innych metod analizy przeżycia w odniesieniu do osób bezrobotnych zarejestrowanych w PUP w Szczecinie w latach 2007–2011.

Wykazano, że płeć osoby bezrobotnej nie była determinantą prawdopodobieństwa wyrejestrowania ze względu na podjęcie pracy. W badanym okresie mężczyźni i kobiety tak samo często korzystali z ofert urzędu pracy. Silnymi determinantami tego prawdopodobieństwa były z kolei wiek i poziom wykształcenia. Osoby młode z wykształceniem wyższym charakteryzowały się najwyższym prawdopodobieństwem podjęcia pracy. Połowa z nich podejmowała pracę w czasie krótszym niż 5 miesięcy. W najgorszej sytuacji byli bezrobotni w wieku 60 lat i więcej, najsłabiej wykształceni. Jednym z celów analizy rynku pracy jest wyodrębnienie grup szczególnie zagrożonych bezrobociem i określenie dla nich programów aktywizujących. Wiele z tych programów jest skierowanych do ludzi młodych. Z drugiej strony osoby młode i słabo wykształcone były najszybciej wykreślane z rejestru osób bezrobotnych z przyczyn będących po ich stronie. Wskazuje to na brak zainteresowania tych osób ofertami urzędów pracy, takimi jak: szkolenia, staże czy zapewnienie doposażenia stanowisk pracy.

Należy podkreślić, że płeć stanowiła ważną determinantę wykreślenia z rejestru bezrobotnych. Mężczyźni charakteryzowali się znacznie większym prawdopodobieństwem wykreślenia z rejestru osób bezrobotnych niż kobiety. Może to być spowodowane większą skłonnością mężczyzn do poszukiwania pracy bez pośrednictwa urzędu lub podejmowania pracy nierejestrowanej. Bez względu na płeć i wykształcenie najmniej skłonne do rezygnacji z pośrednictwa urzędu były osoby najstarsze. W powiązaniu z wnioskami dotyczącymi podejmowania pracy może to świadczyć o oczekiwaniu tych osób na świadczenia przedemerytalne, rentę lub emeryturę.

W badaniu wykorzystano metody analizy trwania. Obserwacji podlegała kohorta osób zarejestrowanych w 2013 r. i obserwowanych do końca 2014 r. Są to dane obejmujące początek spadku stopy bezrobocia w Polsce i Szczecinie,



który świadczy o wychodzeniu z globalnego kryzysu finansowego. Zastosowane metody badawcze należą do metod nieparametrycznych (nie są wymagane żadne założenia dotyczące czasu trwania w bezrobociu obserwowanych osób), w związku z czym możliwe będzie porównanie otrzymanych wyników z wynikami dla planowanych badań dotyczących późniejszych okresów (z pełną stabilizacją gospodarczą i dobrą sytuacją na rynku pracy).

#### BIBLIOGRAFIA

- Aalen, O., Borgan, O., Gjessing, H. (2008). *Survival and Event History Analysis. A Process Point of View*. New York: Springer-Verlag.
- Al-Nachawati, H., Ismail, M., Almohisen, A. (2010). Tree-structured analysis of survival data and its application using SAS software. *Journal of King Saud University (Science)*, 22, 251–255. DOI: 10.1016/j.jksus.2010.05.006.
- Bieszk-Stolorz, B. (2013). *Analiza historii zdarzeń w badaniu bezrobocia*. Szczecin: Volumina.pl Daniel Krzanowski.
- Bieszk-Stolorz, B. (2017). Cumulative Incidence Function in Studies on the Duration of the Unemployment Exit Process. *Folia Oeconomica Stetinensia*, 17(1), 138–150. DOI: 10.1515/fole-2017-0011.
- Bieszk-Stolorz, B., Dmytrów, K. (2018). Efektywność form aktywizacji zawodowej w przekroju wojewódzkim. *Wiadomości Statystyczne*, 63(12), 57–74.
- Bieszk-Stolorz, B., Dmytrów, K. (2019). Spatial diversity of effectiveness of forms of professional activation in Poland in years 2008–2014 by poviats. *Oeconomia Copernicana*, 10(1), 113–130. DOI: 10.24136/oc.2019.006.
- Bieszk-Stolorz, B., Markowicz, I. (2015). Influence of Unemployment Benefit on Duration of Registered Unemployment Spells. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 10(3), 167–183. DOI: 10.12775/EQUIL.2015.031.
- Bou-Hamad, I., Larocque, D., Ben-Ameur, H., Mâsse, L. C., Vitaro, F., Tremblay, R. E. (2009). Discrete-time survival trees. *Canadian Journal of Statistics / La revue canadienne de statistique*, 37(1), 17–32. DOI: 10.1002/cjs.10007.
- Cappelli, C., Zhang, H. (2007). Survival Trees. W: W. Härdle, Y. Mori, P. Vieu (red.), *Statistical Methods for Biostatistics and Related Fields* (s. 167–179). Berlin: Springer-Verlag. DOI: 10.1007/978-3-540-32691-5.
- De Rose, A., Pallara, A. (1997). Survival Trees: An Alternative Non-Parametric Multivariate Technique for Life History Analysis. *European Journal of Population / Revue Européenne de Démographie*, 13(3), 223–241. DOI: 10.1023/A:1005844818027.
- GUS. (2011). *Rocznik Statystyczny Pracy 2010*. Warszawa: Główny Urząd Statystyczny.
- GUS. (2014). *Rocznik Statystyczny Pracy 2012*. Warszawa: Główny Urząd Statystyczny.
- GUS. (2016). *Rocznik Statystyczny Pracy 2015*. Warszawa: Główny Urząd Statystyczny.
- GUS. (2018). *Rocznik Statystyczny Pracy 2017*. Warszawa: Główny Urząd Statystyczny.
- Hadaś-Dyduch, M., Pietrzak, M. B., Balcerzak, A. P. (2016). *Wavelet Analysis of Unemployment Rate in Visegrad Countries* (Institute of Economic Research Working Papers No. 37).
- Kaplan, E. L., Meier, P. (1958). Non-parametric estimation from incomplete observations. *Journal of American Statistical Association*, 53(282), 457–481. DOI: 10.2307/2281868.
- Kleinbaum, D., Klein, M. (2005). *Survival Analysis. A Self-Learning Text*. New York: Springer-Verlag. DOI: 10.1007/0-387-29150-4.

- Landmesser, J. (2013). *Wykorzystanie metod analizy czasu trwania do badania aktywności ekonomicznej ludności w Polsce*. Warszawa: Wydawnictwo SGGW.
- LeBlanc, M., Crowley, J. (1993). Survival trees by goodness of split. *Journal of the American Statistical Association*, 88(422), 457–467. DOI: 10.2307/2290325.
- Mikulec, A., Misztal, M. (2018). Zastosowanie metody rekurencyjnego podziału w analizie trwania przedsiębiorstw województwa łódzkiego. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, (507), 179–190. DOI: 10.15611/pn.2018.507.18.
- Mudunuru, V. R. (2016). *Modeling and Survival Analysis of Breast Cancer: A Statistical, Artificial Neural Network, and Decision Tree Approach*. Graduate Theses and Dissertations. Pobrane z: <http://scholarcommons.usf.edu/etd/6120>.
- von Wangenheim, B. F. (2013). *Survival trees – a new method in innovation theory: A successful introduction of a method commonly used in survival analysis into the field of innovation diffusion theory*. Hamburg: Anchor Academic Publishing.
- Zhou, Y., McArdle, J. J. (2015). Rationale and Applications of Survival Tree and Survival Ensemble Methods. *Psychometrika*, 80(3), 811–833. DOI: 10.1007/s11336-014-9413-1.