

STATYSTYKA W PRAKTYCE

Beata BIESZK-STOLORZ

Ocena wpływu płci na formę wyjścia z bezrobocia¹

Streszczenie. Liczne badania wskazują na zróżnicowanie sytuacji kobiet i mężczyzn na rynku pracy. W celu wyrównania szans urzędy pracy oferują bezrobotnym różne formy aktywizacji. Zaobserwowano jednak, że obok wyrejestr wywania z urzędów osób podejmujących pracę częstym przypadkiem było wykreślenie z powodu odmowy przyjęcia propozycji zatrudnienia.

Celem badania jest ocena wpływu płci osób bezrobotnych na intensywność ich wyrejestr wywania z powodu podjęcia pracy, wykreślenia lub innych przyczyn. Powody te stanowią różnego rodzaju zdarzenia konkurujące. Do oceny prawdopodobieństwa wyrejestr wywania wykorzystano estymator Kaplana-Meiera, a do określenia intensywności wyrejestr wywania z konkretnego powodu — model Lunna-McNeila. W badaniu wykorzystano dane indywidualne osób zarejestr owanych w Powiatowym Urzędzie Pracy w Szczecinie w latach 2013 i 2014.

Wyniki badania pokazały, że kobiety częściej niż mężczyźni podejmowały pracę, natomiast mężczyźni byli intensywniej wykreślani z rejestru. W przypadku pozostałych powodów intensywność wyrejestr wywania ze względu na płeć nie różniła się w sposób istotny statystycznie.

Słowa kluczowe: estymator Kaplana-Meiera, model Lunna-McNeila, ryzyko konkurujące, bezrobocie, płeć.

JEL: C41, J64

¹ Artykuł opracowano na podstawie referatu wygłoszonego podczas Ogólnopolskiej Konferencji Naukowej *Procesy integracji różnych źródeł danych w badaniach statystycznych. Źródła administracyjne, pozaadministracyjne i Big Data w statystyce publicznej*, która odbyła się 31 maja i 1 czerwca 2017 r. w Szczecinie.

Przemiany zachodzące na współczesnym rynku pracy powodują, że sytuacja kobiet w ostatnich latach ulega stopniowej zmianie. Upowszechnienie elastycznych form zatrudnienia, wzrost poziomu wykształcenia, jak również przemiany społeczne, takie jak przechodzenie od tradycyjnego modelu rodziny do partnerskiego, mogą wpływać na polepszenie sytuacji kobiet. Nadal jednak jest ona trudniejsza niż sytuacja mężczyzn, co wynika z podwójnej roli — matki i opiekunki oraz osoby aktywnej zawodowo (Kotowska, Sztanderska i Wóycicka, 2007). Zróznicowaną sytuację kobiet i mężczyzn na rynku pracy potwierdzają dane statystyczne (tabl. 1).

TABL. 1. WYBRANE CHARAKTERYSTYKI RYNKU PRACY W 2016 R.

Wskaźniki	Ogółem	Kobiety	Mężczyźni
Współczynnik aktywności zawodowej	56,3	48,2	65,1
Wskaźnik zatrudnienia	53,2	45,5	61,6
Stopa bezrobocia	5,5	5,7	5,4
Przeciętny czas poszukiwania pracy	10,7	10,4	11,0

Źródło: opracowanie własne na podstawie GUS (2017).

Według danych BAEL w IV kwartale 2016 r. sytuację kobiet w porównaniu z mężczyznami cechował niższy współczynnik aktywności zawodowej osób w wieku 15 lat i więcej, niższy wskaźnik zatrudnienia oraz wyższa stopa bezrobocia. Kobiety przeważały wśród biernych zawodowo (61,8%) i stanowiły 53,3% zarejestrowanych w urzędach pracy. Charakterystyczne jest to, że przeciętny czas poszukiwania pracy kobiet i mężczyzn był zbliżony (nieco niższy w przypadku kobiet). Taka sytuacja występowała również na lokalnym rynku pracy w Szczecinie (Batóg i Batóg, 2016; Bieszk-Stolorz, 2013). Kobiety częściej niż mężczyźni korzystały z subsydiowanych form aktywizacji, uczestniczyły w programach zwiększających aktywność zawodową i decydowały się na podjęcie działalności gospodarczej (Wiśniewski i Zawadzki, 2010). Liczne badania wskazywały również na nierówności w wynagrodzeniach kobiet i mężczyzn oraz niską obecność kobiet na wyższych stanowiskach i w zarządach firm (Kiełkowska, 2013).

Podjęcie pracy nie jest jedynym powodem wyrejestrowania z urzędu pracy. Zarejestrowani bezrobotni korzystają z możliwości przejścia na rentę lub emeryturę, czy zasiłków przedemerytalnych, podejmują naukę w systemie dziennym, a ponadto (choć są to rzadsze przypadki) wyjeżdżają za granicę na okres dłuższy niż 30 dni, zmieniają miejsce zamieszkania lub umierają. Pomimo wielu działań mających na celu przeciwdziałanie bezrobociu ciągle bardzo często wyrejestrowanie następuje z powodu odmowy, bez uzasadnionej przyczyny, przyjęcia propozycji zatrudnienia lub innej formy pomocy, a także niezgłoszenia się w urzędzie w wyznaczonym terminie. W takich sytuacjach bezrobotni są wykreśleni z rejestru i nakłada się na nich okres karencji, którego długość zależy od liczby wykreśleń (ustawa z dnia 20 kwietnia 2004 r. o promocji zatrudnienia

i instytucjach rynku pracy). Utrudnione odzyskanie statusu bezrobotnego, a co za tym idzie prawa do ubezpieczenia zdrowotnego i zasiłku, nie wpływa znacząco na zmniejszenie się odsetka osób wykreślanych z rejestru (tabl. 2). W 2014 r. było to 26,49% osób wyrejestrowanych. Również w tym przypadku widać wyraźne różnice ze względu na płeć osoby bezrobotnej. Mężczyźni byli wykreślani z rejestru częściej (32,54%) niż kobiety (19,95%).

TABL. 2. LICZBA WYREJESTROWAŃ Z URZĘDÓW PRACY W POLSCE

Osoby wyłączone z ewidencji	Jednostka miary	2008	2009	2010	2011	2012	2013	2014
O g ó ł e m								
O g ó ł e m	w tys.	2749,4	2664,8	2979,8	2563,5	2500,8	2688,2	2786,2
Które podjęły pracę	w tys.	1052,0	1012,4	1183,3	1130,7	1096,4	1262,9	1285,8
	w %	38,26	37,99	39,71	44,11	43,84	46,98	46,15
Które nie potwierdziły gotowości do pracy	w tys.	874,9	837,8	874,1	827,8	741,0	726,2	738,2
	w %	31,82	31,44	29,33	32,29	29,63	27,01	26,49
Kobiety								
O g ó ł e m	w tys.	1456,8	1330,5	1419,4	1224,5	1217,1	1285,6	1337,8
Które podjęły pracę	w tys.	543,7	500,1	556,1	564,7	552,2	628,0	644,8
	w %	37,32	37,59	39,18	46,12	45,37	48,85	48,20
Które nie potwierdziły gotowości do pracy	w tys.	394,9	345,6	331,7	316,2	282,3	265,1	266,9
	w %	27,11	25,98	23,37	25,82	23,19	20,62	19,95
Mężczyźni								
O g ó ł e m	w tys.	1292,6	1334,3	1560,4	1339,0	1283,7	1402,5	1448,4
Które podjęły pracę	w tys.	508,3	512,3	627,2	566,0	544,2	634,8	641,0
	w %	39,32	38,39	40,19	42,27	42,39	45,26	44,26
Które nie potwierdziły gotowości do pracy	w tys.	480,0	492,2	542,4	511,6	458,7	461,1	471,3
	w %	37,13	36,89	34,76	38,21	35,73	32,88	32,54

Źródło: opracowanie własne na podstawie GUS (2011, 2013, 2016).

Różne powody wyrejestrowania z urzędu pracy są zdarzeniami konkurującymi, czyli takimi, których wystąpienie wyklucza wystąpienie innego zdarzenia lub fundamentalnie zmienia prawdopodobieństwo wystąpienia tego innego zdarzenia (Gooley, Leisenring, Crowley i Storer, 1999). Spełnione jest założenie o ich niezależności, gdyż wystąpienie zdarzenia określonego typu nie ma żadnego wpływu na prawdopodobieństwo wystąpienia jakichkolwiek innych zdarzeń (Crowder, 1994, 1996, 1997). Ten rodzaj ryzyka w literaturze określa się jako bezwarunkowy model ryzyka konkurującego (*unconditional competing risks*). Jeżeli prawdopodobieństwo przejścia między stanami zależy od zmiennych objaśniających oraz od czasu i typu pobytu we wcześniejszym stanie, zjawisko to można opisać za pomocą warunkowych modeli ryzyka konkurującego (*conditional competing risks models*) (Landmesser, 2008).

Celem badania jest analiza wpływu płci osób bezrobotnych na intensywność wyrejestrowywania z urzędu pracy z powodu podjęcia pracy, wykreślenia lub pozostałych przyczyn, z zastosowaniem metod analizy trwania uwzględniających obserwacje cenzurowane. Do oceny prawdopodobieństwa wyrejestrowania

z urzędu wykorzystano estymator Kaplana-Meiera, a intensywność zdarzeń oceniono za pomocą bezwarunkowego modelu ryzyka konkurującego — modelu Lunna-McNeila. Badanie przeprowadzono na podstawie danych indywidualnych osób zarejestrowanych w Powiatowym Urzędzie Pracy (PUP) w Szczecinie wygenerowanych z systemu Syriusz.

DANE WYKORZYSTANE W BADANIU

Zastosowanie w badaniu metod analizy trwania wymaga dostępu do danych indywidualnych, które muszą zawierać informację o czasie trwania jednostki w określonym stanie. W przypadku badania bezrobocia rejestrowanego w Polsce najlepszym źródłem takich danych są rejestry powiatowych urzędów pracy. Dzięki wdrożeniu systemu informatycznego Syriusz gromadzą one obszerne informacje o zarejestrowanych bezrobotnych. W badaniu wykorzystano anonimowe dane indywidualne o 22078 osobach bezrobotnych uzyskane z PUP w Szczecinie. Analizowano czas trwania w bezrobociu od zarejestrowania w 2013 r. do wyrejestrowania z określonego powodu do końca 2014 r. Część obserwacji nie zakończyła się zdarzeniem, czyli wyrejestrowaniem w analizowanym okresie; są to obserwacje cenzurowane prawostronnie. Podano kilkadziesiąt powodów wyrejestrowania, m.in. związanych z: podjęciem zatrudnienia, przejściem na rentę lub emeryturę, kontynuacją nauki, wyjazdem za granicę czy odmową podjęcia zatrudnienia. Zostały one zebrane w trzy grupy zdarzeń konkurujących (główne powody wyrejestrowania): praca, wykreślenie i pozostałe powody. Liczebność grup bezrobotnych wyróżnionych ze względu na główne powody wyrejestrowania oraz bezrobotnych niewyrejestrowanych podano w tabl. 3.

TABL. 3. LICZEBNOŚĆ BEZROBOTNYCH ZE WZGLĘDU NA PŁEĆ I POWÓD WYREJESTROWANIA W LATACH 2013 I 2014

Bezrobotni	Ogółem	Wyrejestrowani z powodu						Niewyrejestrowani
		podjęcia pracy		wykreślenia		pozostałych przyczyn		
	w liczbach bezwzględnych	w %	w liczbach bezwzględnych	w %	w liczbach bezwzględnych	w %		
O g ó ł e m	22078	9633	43,63	8965	40,61	1624	7,36	1856
Kobiety	9770	4809	49,22	3264	33,41	784	8,02	913
Mężczyźni	12308	4824	39,19	5701	46,32	840	6,82	943

Źródło: opracowanie własne.

Szczegółowe powody wyrejestrowania składające się na grupy głównych powodów przedstawiono w zestawieniu. Należy zauważyć, że:

- grupę praca tworzą trzy podgrupy: podjęcie pracy lub innego zatrudnienia, podjęcie pracy subsydiowanej przez urząd oraz działalność gospodarcza;

- wykreślenie grupuje powody wykreślenia z rejestru z własnej winy;
- pozostałe powody wyrejestrowania są mniej liczne i — jak wykazały wcześniejsze badania — każdy z osobna miał marginalny wpływ na prawdopodobieństwo wyrejestrowania z urzędu, dlatego postanowiono rozpatrywać je razem.

W analizowanym okresie z PUP w Szczecinie wykreślono prawie 41% osób, więcej mężczyzn (46%) niż kobiet (33%). Pracę podjęło 49% kobiet i 39% mężczyzn.

ZESTAWIENIE POWODÓW WYREJESTROWANIA

Powody główne	Powody szczegółowe
Praca	podjęcie pracy lub innego zatrudnienia podjęcie 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) działalność gospodarcza (podjęcie pozarolniczej działalności gospodarczej, przyznanie jednorazowo środków na podjęcie działalności gospodarczej, rozpoczęcie działalności gospodarczej ze środków PFRON)
Wykreślenie	odmowa, bez uzasadnionej przyczyny, przyjęcia propozycji odpowiedniego zatrudnienia lub innej pracy zarobkowej, wykonywania prac interwencyjnych lub robót publicznych albo udziału w szkoleniu, stażu czy przygotowaniu zawodowym w miejscu pracy niestawienie się w PUP w wyznaczonym terminie i niepowiadomienie w terminie do 5 dni o uzasadnionej przyczynie tego niestawienia nieprzedstawienie zaświadczenia o niezdolności do pracy wskutek choroby odmowa, bez uzasadnionej przyczyny, udziału w działaniach w ramach Programu Aktywizacja i Integracja (PAI) lub przerwaniu z własnej winy udziału w PAI brak gotowości do podjęcia pracy przez co najmniej 10 dni przerwanie, z własnej winy, szkolenia, stażu lub realizacji indywidualnego planu działań wniosek bezrobotnego o wykreślenie z ewidencji
Pozostałe	wyjazd za granicę na co najmniej 30 dni zmiana miejsca zameldowania lub pobytu poza obszar działania PUP niezdolność do pracy wskutek choroby lub przebywania w zamkniętym ośrodku odwykowym przez nieprzerwany okres 90 dni rozpoczęcie szkolenia organizowanego przez inny podmiot niż PUP powołanie do zasadniczej służby wojskowej podjęcie nauki w szkole w systemie dziennym nabycie prawa do emerytury, renty lub świadczenia rehabilitacyjnego pobieranie zasiłku stałego podleganie ubezpieczeniu emerytalno-rentowemu z tytułu stałej pracy jako domownik w gospodarstwie rolnym pobieranie świadczenia lub dodatku do zasiłku rodzinnego z tytułu samotnego wychowywania dziecka pobieranie zasiłku dla opiekuna przyznanie prawa do pobierania świadczenia/zasiłku przedemerytalnego zgon

METODYKA BADANIA

W analizie trwania bada się czas trwania jednostki w danym stanie aż do momentu wystąpienia określonego zdarzenia kończącego obserwację (zmienna losowa T). Metody analizy trwania stosuje się powszechnie w demografii i naukach medycznych; ich zaletą jest uwzględnianie obserwacji cenzurowanych. Metody analizy trwania stosuje się również w analizie zjawisk społeczno-ekonomicznych, np. do badania wpływu czynników na czas działania firmy, czas trwania bezrobocia lub czas do zaprzestania spłaty kredytu. Zdarzeniami kończącymi obserwację są w wymienionych przypadkach odpowiednio: zakończenie działalności firmy, podjęcie zatrudnienia, zaprzestanie spłaty kredytu. Metodami analizy trwania można badać żywotność firm (Markowicz, 2012), aktywność ekonomiczną ludności (Landmesser, 2013), dynamikę ubóstwa w miejskich i wiejskich gospodarstwach domowych (Sączewska-Piotrowska, 2016), ryzyko kredytowe (Matuszyk, 2015; Wycinka, 2015) czy też czas trwania bezrobocia (Bieszk-Stolorz i Markowicz, 2012; Bieszk-Stolorz, 2013).

Podstawowym pojęciem analizy trwania jest funkcja trwania (funkcja przeżycia), określająca prawdopodobieństwo braku zdarzenia co najmniej do czasu t , a zdefiniowana następująco:

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

gdzie:

T — czas trwania zjawiska,

$F(T)$ — dystrybuanta rozkładu zmiennej losowej T .

Jeżeli zdarzeniem kończącym obserwację jest zgon, to funkcja trwania opisuje prawdopodobieństwo przeżycia jednostki. W przypadku gdy zdarzenie kończące obserwację ma charakter zjawiska pożądanego (np. wyjście z bezrobocia), wygodniej jest analizować dystrybuantę $F(T) = 1 - S(t)$, która wyraża prawdopodobieństwo, że zdarzenie zajdzie najpóźniej do czasu t . Jeżeli zdarzeniem jest podjęcie pracy przez osobę bezrobotną zarejestrowaną w urzędzie, to estymator funkcji trwania wyraża prawdopodobieństwo pozostania w rejestrze, a estymator dystrybuanty pozwala określić, jakie jest prawdopodobieństwo podjęcia pracy.

W analizie trwania jednostki są obserwowane do wystąpienia zdarzenia kończącego badany proces lub stan. Ze względu na ograniczony czas obserwacji część jednostek może jednak nie doświadczyć zdarzenia końcowego. W takich przypadkach czas trwania dla badanej jednostki jest znany jedynie częściowo (obserwacje cenzurowane prawostronnie). Jeżeli rozkład czasu trwania analizo-

wanego zjawiska nie jest znany, do estymacji funkcji trwania najczęściej stosuje się estymator Kaplana-Meiera (Kaplan i Meier, 1958):

$$\hat{S}(t) = \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{n_j}\right) \quad (2)$$

gdzie:

$t_1 < t_2 < \dots < t_i < \dots < t_n$ — momenty zachodzenia zdarzeń,

d_j — liczba zdarzeń w momencie t_j ,

n_j — liczba jednostek zagrożonych do momentu t_j .

Funkcje trwania wyznaczone dla podgrup wyodrębnionych ze względu na warianty analizowanej cechy mogą być porównywane. Ze względu na nieznaną rozkładów czasu trwania stosuje się testy nieparametryczne oparte na porządku rangowym czasu trwania. Nie ma powszechnie akceptowanych metod wyboru testu w danej sytuacji. Często zastosowanie znajduje test log-rank (Kleinbaum i Klein, 2005, s. 57—61), służący do weryfikacji hipotezy $H_0: S_1(t) = S_2(t)$ o równości krzywych trwania wyznaczonych dla dwóch podgrup. Statystyka testowa jest porównywana — przy danym poziomie istotności — z rozkładem chi kwadrat o jednym stopniu swobody. Test log-rank można również wykorzystywać do porównania dopełnienia do jedności estymatora Kaplana-Meiera (1 — KM).

Drugą funkcją stosowaną w analizie trwania jest funkcja hazardu. Opisuje ona chwilowy potencjał zdarzenia pojawiającego się w momencie t pod warunkiem przetrwania do czasu t . Definiuje się ją następująco (Kleinbaum i Klein, 2005):

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \quad (3)$$

W praktyce często stosuje się model proporcjonalnych hazardów Coxa, który można zapisać:

$$h(t, X) = h_0(t) \exp \left(\sum_{j=1}^p \beta_j X_j \right) \quad (4)$$

gdzie:

$h_0(t)$ — hazard bazowy,

X_1, X_2, \dots, X_p — zmienne objaśniające.

Parametry β_i modelu pozwalają ocenić intensywność względną wystąpienia danego zdarzenia. W tym celu wyznacza się ilorazy hazardu $HR = \exp(\beta_i)$.

Oprócz obserwacji niezakończonych za cenzurowane uznaje się również takie, w których badana jednostka znika z pola widzenia lub występuje zdarzenie kończące obserwację i jednocześnie wykluczające zajście zdarzenia właściwego (Pepe, 1991), czyli zdarzenie konkurujące. W analizie przyczyn śmierci zgon jednostki może nastąpić bezpośrednio z powodu określonej choroby (np. rak) lub być spowodowany skutkami leczenia (np. powikłania po chemioterapii). Rozróżnienie tych przyczyn ma istotne znaczenie dla procesu leczenia. W analizie bezawaryjności urządzeń ryzyko konkurujące wiąże się z wpływem elementów składowych na cały system (awaria systemu może nastąpić w wyniku awarii któregośkolwiek elementu). W praktyce badawczej obserwacje zakończone zdarzeniem konkurującym można potraktować jako cenzurowane prawostronnie.

Ciekawych wyników dostarcza wykorzystanie modeli ryzyka konkurującego (Klein i Moeschberger, 1984; Klein i Bajorunaite, 2004). Niech T będzie ciągłą zmienną losową opisującą czas do zajścia pewnego zdarzenia, a C — ciągłą zmienną losową wyrażającą czas do cenzurowania. Dla K konkurujących rodzajów ryzyka obserwacji podlegają pary (X, δ) , gdzie $X = \min(T, C)$ i $\delta = 0, 1, \dots, K$. Jeżeli dana obserwacja jest cenzurowana, to $\delta = 0$, a dla obserwacji kończących się zdarzeniem (jednym z K konkurujących) $\delta = 1, \dots, K$. W przypadku różnych rodzajów ryzyka konkurującego ($K \geq 1$) K funkcji hazardu jest opisanych wzorem (Klein i Moeschberger, 2003, s. 50):

$$h_k(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t, \delta = k | T \geq t)}{\Delta t} \quad \text{dla } k = 1, 2, 3, \dots, K \quad (5)$$

Intensywność względną zajścia zdarzenia określonego typu można ocenić, korzystając z modelu proporcjonalnego hazardu Coxa, który dla K rodzajów ryzyka konkurującego określony jest wzorem (Kleinbaum i Klein, 2005, s. 400):

$$h_k(t, \mathbf{x}) = h_{0k}(t) \exp \left(\sum_{j=1}^p \beta_{jk} X_j \right) \quad \text{dla } k = 1, 2, 3, \dots, K \quad (6)$$

gdzie $\mathbf{x} = [X_1, X_2, \dots, X_p]$ — wektor zmiennych objaśniających.

Każdy rodzaj ryzyka jest opisany osobnym modelem proporcjonalnego hazardu. W każdym z nich obserwacje, które nie zakończyły się zdarzeniem typu k , traktuje się jako obserwacje cenzurowane. W przypadku występowania zdarzeń konkurujących do oceny intensywności względnej wystąpienia danego zdarzenia do czasu t można wykorzystać jeden model — Lunna-McNeila. Dane w tym modelu muszą zostać pogrupowane w specjalny sposób (tabl. 4). Jeśli występu-

je K rodzajów ryzyka, to wyjściowe dane należy zduplikować K razy. Obiekt i o czasie trwania t_i występuje K razy. Każdej obserwacji (również tej zduplikowanej) nadaje się status e_i , który informuje o przynależności obiektu i do określonej grupy ryzyka: e_i jest równe 1, jeżeli obserwacja obiektu i zakończyła się zajściem zdarzenia typu k ($k = 1, 2, \dots, K$), a w przeciwnym wypadku ma wartość 0. W przypadku obserwacji cenzurowanej e_i jest równe 0 dla każdego $k = 1, 2, \dots, K$. W modelu wprowadza się sztuczne zmienne dychotomiczne D_1, D_2, \dots, D_k , reprezentujące K rodzajów ryzyka: D_k jest równe 1 dla ryzyka typu k i 0 dla pozostałych rodzajów ryzyka.

TABL. 4. SPOSÓB DEFINIOWANIA ZMIENNYCH W MODELU LUNNA-MCNEILA

Obiekt	Czas	Status	D_1	D_2	D_3	...	D_k	X_1	...	X_p
i	t_i	e_1	1	0	0	...	0	x_{i1}	...	x_{ip}
i	t_i	e_2	0	1	0	...	0	x_{i1}	...	x_{ip}
i	t_i	e_3	0	0	1	...	0	x_{i1}	...	x_{ip}
...
i	t_i	e_K	0	0	0	...	1	x_{i1}	...	x_{ip}

Źródło: jak przy tabl. 3.

Jeżeli ryzyko $k = 1$ przyjmie się za referencyjne, to w modelu uwzględnia się $k - 1$ zmiennych sztucznych D_2, D_3, \dots, D_K . Natomiast X_1, X_2, \dots, X_p są zmiennymi objaśniającymi.

Jeżeli przez $g = 1, 2, \dots, K$ oznaczy się warstwy będące rodzajami ryzyka odpowiadające K zdarzeń konkurujących, a zdarzenie typu $k = 1$ przyjmie się za referencyjne, to model Lunna-McNeila można zdefiniować jako warstwowy model regresji Coxa z interakcjami. Ma on postać (Kleinbaum i Klein, 2005, s. 423):

$$h_g^*(t, X) = h_{0g}^*(t) \exp \left(\sum_{j=1}^p \beta_j X_j + \sum_{k=2}^K \sum_{j=1}^p \delta_{kj} D_k X_j \right) \quad (7)$$

gdzie:

X_1, X_2, \dots, X_p — zmienne objaśniające,

D_2, D_3, \dots, D_K — $K - 1$ zmiennych sztucznych.

Przy założeniu, że jedna ze zmiennych objaśniających (X_j) jest dychotomiczna, można wyznaczyć ilorazy hazardu.

W warstwie $g = 1$, czyli dla zdarzenia konkurującego $k = 1$, zachodzi $D_2 = D_3 = \dots = D_K = 0$; iloraz hazardu jest określony wzorem:

$$HR_{g=1}(X_j = 1/X_j = 0) = \exp(\beta_j) \quad (8)$$

W pozostałych warstwach ($g > 1$), czyli dla $k > 1$, $D_k = 1$ i $D_k = 0$, gdy $k' \neq k$, ilorazy hazardu mają postać:

$$HR_{g=k>1}(X_j = 1/X_j = 0) = \exp(\beta_j + \delta_{kj}) = \exp(\beta_j)\exp(\delta_{kj}) \quad (9)$$

Obok podstawowej wersji modelu Lunna-McNeila (7) stosuje się również jego wersję alternatywną o postaci (Kleinbaum i Klein, 2005, s. 427):

$$h_g^{**}(t, X) = h_{0g}^{**}(t) \exp\left(\sum_{k=1}^K \sum_{j=1}^p \delta'_{kj} D_k X_j\right) \quad (10)$$

gdzie:

X_1, X_2, \dots, X_p — zmienne objaśniające,

D_1, D_2, \dots, D_k — K zmiennych sztucznych.

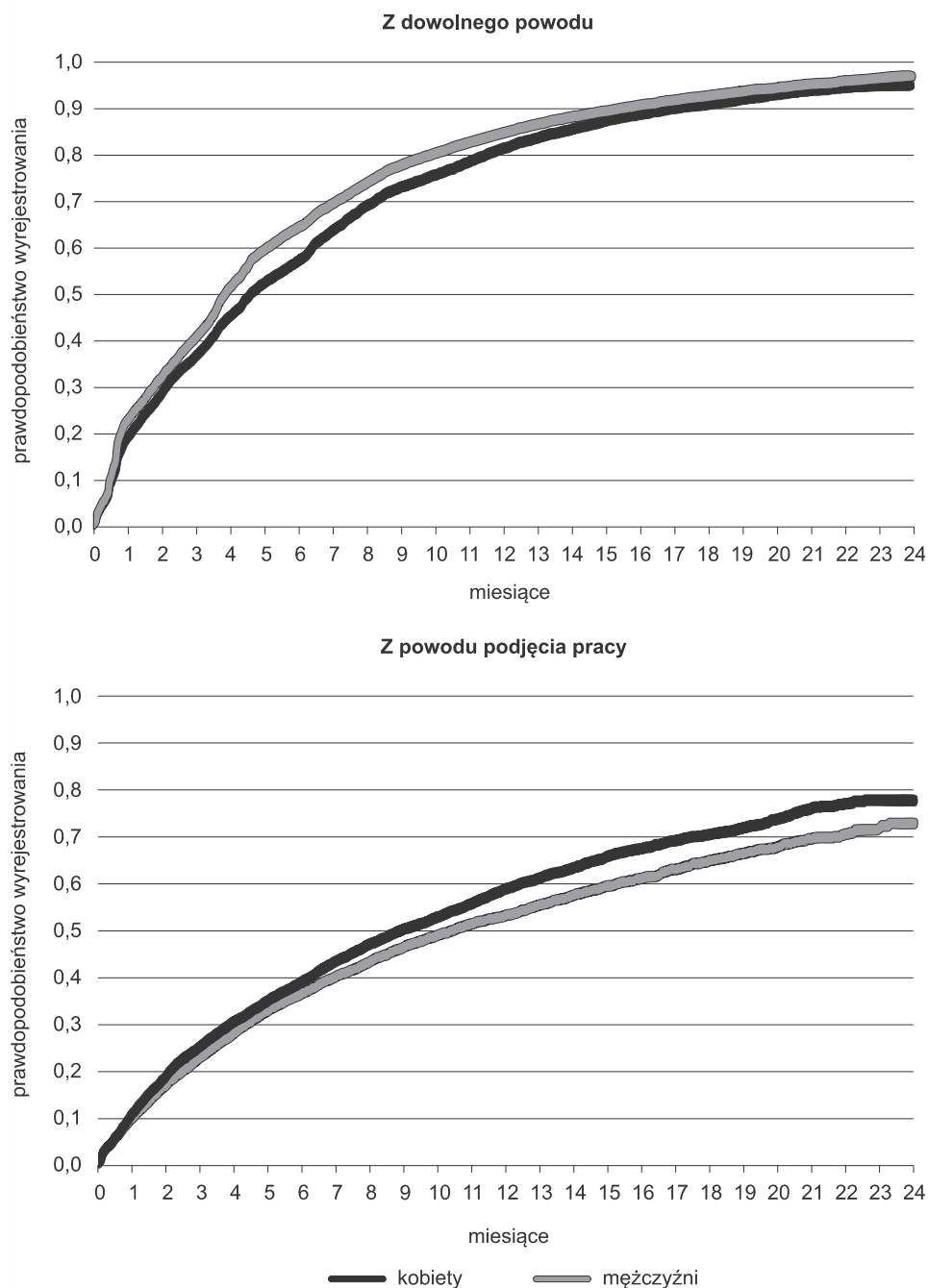
Dla określonego $g = k \geq 1$ zachodzi $D_k = 1$ i $D_{k'} = 0$, gdy $k' \neq k$, a ilorazy hazardu są określone wzorem:

$$HR_{g=k \geq 1}(X_j = 1/X_j = 0) = \exp(\delta'_{kj}) \quad (11)$$

Ilorazy hazardu w modelach interpretuje się jako hazard względny (intensywność względną) zajścia zdarzenia typu k .

ANALIZA WPŁYWU PŁCI NA PRZYCZYNY WYREJESTROWANIA Z URZĘDU PRACY

Analizę przeprowadzono w dwóch etapach. Pierwszy polegał na wykorzystaniu dopełnienia do jedności estymatora Kaplana-Meiera (estymator dystrybuan-ty) w celu oceny prawdopodobieństwa wyrejestrowania z określonego powodu bezrobotnych kobiet i mężczyzn. Najpierw analizowano wyrejestrowanie z urzędu pracy z dowolnego powodu. Za dane cenzurowane przyjęto obserwacje, które nie zakończyły się przed końcem 2014 r. Na wykresie widać, że prawdopodobieństwo wyrejestrowania bezrobotnych mężczyzn z dowolnego powodu było nieco większe niż kobiet. Istotność różnic między krzywymi została potwierdzona testem log-rank ($\chi^2 = 77,6440$; $p = 0,0000$ — tabl. 5). Prawdopodobieństwo pozostania w rejestrze urzędu po 12 miesiącach od momentu zarejestrowania (i jednocześnie prawdopodobieństwo przejścia w stan długotrwałego bezrobocia) było równe 0,16 w przypadku mężczyzn i 0,19 w przypadku kobiet.

**ESTYMATORY 1 – KM PRAWDOPODOBIENSTWA WYREJESTROWANIA
DLA KOBIECI I MĘŻCZYŹN**

Źródło: jak w tabl. 3.

Następnie przeprowadzono analizę prawdopodobieństwa wyrejestrowania z urzędu z powodu podjęcia pracy. Za obserwacje cenzurowane przyjęto obserwacje, które nie zakończyły się wyrejestrowaniem przed końcem 2014 r., oraz obserwacje, które zakończyły się wyrejestrowaniem z powodu innego niż podjęcie pracy. Prawdopodobieństwo podjęcia pracy przez kobiety było większe niż mężczyzn. Istotność różnic w przebiegu obu krzywych została potwierdzona testem log-rank ($\chi^2 = 40,3170$; $p = 0,0001$ — tabl. 5). Prawdopodobieństwo podjęcia pracy po 12 miesiącach od momentu zarejestrowania wyniosło 0,53 w przypadku mężczyzn i 0,59 w przypadku kobiet.

**TABL. 5. WYNIK TESTU LOG-RANK DLA BEZROBOTNYCH
WEDŁUG POWODÓW WYREJESTROWANIA**

Powody wyrejestrowania	Test log-rank (chi-kwadrat)	p
Dowolny powód	77,6440	0,0000
Praca	40,3170	0,0001

Źródło: jak przy tabl. 3.

Analizując wykres, można zauważyć, że krzywe obrazujące sytuację wyrejestrowania z powodu podjęcia pracy przyjmują wyraźnie niższe wartości niż te obrazujące sytuację wyrejestrowania z dowolnego powodu. Po 24 miesiącach od zarejestrowania prawdopodobieństwo wyrejestrowania z dowolnego powodu w przypadku kobiet było równe 0,94, a z powodu podjęcia pracy — 0,78 (w przypadku mężczyzn odpowiednio: 0,96 i 0,73). Oznacza to, że znaczący wpływ na ogólne prawdopodobieństwo wyrejestrowania mogły mieć również inne przyczyny niż podjęcie pracy. Bardziej szczegółowych wniosków na ten temat dostarcza analiza ryzyka konkurującego uwzględniająca różne rodzaje powodów wyrejestrowania (zdarzeń konkurujących), przeprowadzona w drugim etapie omawianego badania. Ponieważ wcześniejsze badania wykazały, że wiele powodów wyrejestrowania miało znaczenie marginalne (Bieszk-Stolorz, 2017), określono trzy grupy zdarzeń kończących obserwację: podjęcie pracy, wykreślenie i pozostałe, które stanowią zdarzenia konkurujące. Danymi cenzurowanymi były obserwacje, które nie zakończyły się przed końcem 2014 r.

W badaniu wykorzystano trzy nieliniowe modele regresji pozwalające ocenić wpływ płci na intensywność różnych form wychodzenia z bezrobocia. Zmienną objaśniającą X występującą w tych modelach jest płeć osoby bezrobotnej; to zmienna dychotomiczna przyjmująca wartość 1 dla kobiet i 0 dla mężczyzn (grupa referencyjna).

Modele proporcjonalnego hazardu Coxa (6) dla trzech zdarzeń konkurujących przyjmują postać:

$$h_k(t, X) = h_{0k}(t) \exp(\beta_k X) \quad \text{dla } k = 1, 2, 3 \quad (12)$$

Wyniki estymacji parametrów i ilorazy hazardu $\exp(\beta_k)$ przedstawiono w tabl. 6.

Kolejnym zastosowanym modelem był model Lunna-McNeila (7) o postaci:

$$h_g^*(t, X) = h_{0g}^*(t) \exp(\beta X + \delta_2 D_2 X + \delta_3 D_3 X) \quad (13)$$

Zmienne sztuczne dychotomiczne (D_k) określają formę ryzyka konkurującego i przyjmują wartość 1 dla ryzyka $g = k$ i 0 w przeciwnym wypadku. W badaniu przyjęto następujące oznaczenia ryzyka: $k = 1$ w przypadku podjęcia pracy, $k = 2$ — wykreślenia i $k = 3$ — pozostałych powodów wyrejestrowania; ustalono poziom istotności parametrów 0,05. Ilorazy hazardu dla zdarzeń konkurujących wyznaczone na podstawie wartości estymatorów parametrów modelu (tabl. 7) i wzorów (8) i (9) mają następujące wartości:

$$\begin{aligned} HR_{g=1} &= \exp(\beta) = 1,1354, \\ HR_{g=2} &= \exp(\beta + \delta_2) = 1,1354 \cdot 0,5679 = 0,6448, \\ HR_{g=3} &= \exp(\beta + \delta_3) = 1,1354 \cdot 0,9158 = 1,0398. \end{aligned}$$

Alternatywny model Lunna-McNeila (10) dla zmiennej dychotomicznej *Płeć* i trzech rodzajów ryzyka konkurującego przyjmuje postać:

$$h_g^{**}(t, X) = h_{0g}^{**}(t) \exp(\delta'_1 D_1 X + \delta'_2 D_2 X + \delta'_3 D_3 X) \quad (14)$$

Wyniki estymacji parametrów modelu przedstawiono w tabl. 7. Ilorazy hazardu wyznaczono korzystając ze wzoru (11):

$$\begin{aligned} HR_{g=1} &= \exp(\delta'_1) = 1,1354, \\ HR_{g=2} &= \exp(\delta'_2) = 0,6448, \\ HR_{g=3} &= \exp(\delta'_3) = 1,0398. \end{aligned}$$

TABL. 6. OCENY PARAMETRÓW MODELU PROPORCJONALNEGO HAZARDU COXA DLA ZDARZEŃ KONKURUJĄCYCH

Parametr	Ocena parametru	Błąd standardowy	Statystyka Walda	p	Hazard względny
Praca					
β_1	0,1270	0,0204	38,8026	0,0000	1,1354
$\chi^2 = 38,7788 \quad p = 0,0000$					
Wykreślenie					
β_2	-0,4389	0,0220	399,2858	0,0000	0,6448
$\chi^2 = 412,4404 \quad p = 0,0000$					
Pozostałe					
β_3	0,0390	0,0500	0,6168	0,4322	1,0398
$\chi^2 = 0,6166 \quad p = 0,4323$					

Źródło: jak przy tabl. 3.

TABL. 7. OCENY PARAMETRÓW MODELU LUNNA-MCNEILA

Parametry	Ocena parametru	Błąd standardowy	Statystyka Walda	p	Hazard względny
Model podstawowy					
β	0,1270	0,0204	38,8026	0,0000	1,1354
δ_2	-0,5659	0,0230	356,5345	0,0000	0,5679
δ_3	-0,0880	0,0537	2,6836	0,1014	0,9158
$\chi^2 = 451,8358 \quad p = 0,0000$					
Model alternatywny					
δ'_1	0,1270	0,0204	38,8026	0,0000	1,1354
δ'_2	-0,4389	0,0220	399,2858	0,0000	0,6448
δ'_3	0,0390	0,0500	0,6168	0,4322	1,0398
$\chi^2 = 451,8358 \quad p = 0,0000$					

Źródło: jak przy tabl. 3.

Jak widać, wszystkie zastosowane modele dają takie same wyniki. W analizowanym okresie kobiety o prawie 14% intensywniej niż mężczyźni podejmowały pracę i o nieco ponad 35% mniej intensywnie były wykreślane z rejestru. Intensywność wyrejestrowywania z pozostałych powodów w przypadku kobiet i mężczyzn była zbliżona (brak istotności parametru δ'_3).

Podsumowanie

Metody zastosowane w badaniu pozwoliły na dokonanie oceny wpływu płci na intensywność względną wyrejestrowywania z urzędu pracy z określonego powodu. Analiza z wykorzystaniem modeli Lunna-McNeila potwierdziła wyniki analizy wskaźnikowej, wskazując, że podjęcie zatrudnienia nie było jedyną przyczyną wyrejestrowania z urzędu. Pomimo stosowania przez urzędy kar wobec osób niezgłaszających się w wyznaczonym terminie lub odmawiających podjęcia pracy albo innej formy pomocy, wykreślenia nadal stanowią duży odsetek wyrejestrowań. Świadczy to o tym, że wielu bezrobotnych rezygnuje z pośrednictwa urzędu w poszukiwaniu pracy, a sankcje stosowane wobec takich osób nie są skuteczne. Średni czas do wykreślenia w analizowanym okresie był równy 5,3 miesiąca (5,7 miesiąca dla kobiet, 5 miesięcy dla mężczyzn). Wyniki analizy z zastosowaniem modelu Lunna-McNeila wskazały, że kobiety intensywniej niż mężczyźni podejmowały pracę, mężczyźni zaś intensywniej byli wykreśleni z urzędu. Płeć osoby bezrobotnej nie wpływała natomiast na intensywność wyrejestrowania z pozostałych powodów.

Z przedstawionego badania wynikają pewne uwagi metodyczne. Po pierwsze, jeżeli analizowany proces może zakończyć się zdarzeniami różnego rodzaju, to warto wyodrębnić zdarzenia konkurujące i oszacować ryzyko ich wystąpienia. W przypadku analiz dotyczących rynku pracy ważne jest nie tylko podjęcie zatrudnienia przez osobę bezrobotną, lecz także zbadanie innych form wyjścia

z bezrobocia niż podjęcie pracy. Po drugie, wyniki badania pokazują, że przedstawione w artykule modele regresji dają taką samą ocenę ryzyka względnego dla wszystkich wyróżnionych form wyrejestrowania.

dr hab. Beata Bieszk-Stolorz — Uniwersytet Szczeciński

LITERATURA

- Batóg, J., Batóg, B. (2016). Application of correspondence analysis to the identification of the influence of features of unemployed persons on the unemployment duration. *Economics and Business Review*, 16(4), 25—44. DOI: 10.18559/ebr.2016.4.2.
- 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/fofi-2017-0011.
- Bieszk-Stolorz, B., Markowicz, I. (2012). *Modele regresji Coxa w analizie bezrobocia*. Warszawa: CeDeWu.
- Crowder, M. (1994). Identifiability Crises in Competing Risks. *International Statistical Review*, 62(3), 379—391. DOI: 10.2307/1403768.
- Crowder, M. (1996). On assessing independence of competing risks when failure times are discrete. *Lifetime Data Analysis*, 2(2), 195—209. DOI: 10.1007/BF00128575.
- Crowder, M. (1997). A test for independence of competing risks with discrete failure Times. *Lifetime Data Analysis*, 3(3), 215—223. DOI: 10.1023/A:1009696830515.
- Gooley, T. A., Leisenring, W., Crowley, J., Storer, B. E. (1999). Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Statistics in Medicine*, 18(6), 695—706. DOI: 10.1002/(SICI)1097-0258(19990330)18:6<695::AID-SIM60>3.0.CO;2-O.
- GUS. (2011). *Rocznik Statystyczny Pracy 2010*. Warszawa: GUS.
- GUS. (2013). *Rocznik Statystyczny Pracy 2012*. Warszawa: GUS.
- GUS. (2016). *Rocznik Statystyczny Pracy 2015*. Warszawa: GUS.
- GUS. (2017). *Kwartalna informacja o rynku pracy*. Pobrane z: [http://stat.gov.pl/obszary-tematyczne/rynek-pracy/pracujacy-bezrobotni-bierni-zawodowo-wg-bael/monitoring-rynku-pracy-kwartalna-informacja-o-rynku-pracy-czwarty-kwartal-2016-roku,12,31.html](http://stat.gov.pl/obszary-tematyczne/rynek-pracy/pracujacy-bezrobotni-bierni-zawodowo-wg-bael/monitoring-ryнку-pracy-kwartalna-informacja-o-rynku-pracy-czwarty-kwartal-2016-roku,12,31.html).
- Kaplan, E. L., Meier, P. (1958). Non-parametric estimation from incomplete observations. *Journal of American Statistical Association*, 53, 457—481. DOI: 10.2307/2281868.
- Kielkowska, M. (red.). (2013). *Rynek pracy wobec zmian demograficznych*. Warszawa: Instytut Obywatelski.
- Klein, J. P., Bajorunaite, R. (2004). Inference for Competing Risks. W: N. Balakrishnan, C. R. Rao (red.), *Handbook of Statistics: Advances in Survival Analysis* (s. 291—311). New York: Elsevier.
- Klein, J. P., Moeschberger, M. L. (1984). Asymptotic bias of the product limit estimator under dependent competing risks. *Indian Journal of Productivity, Reliability and Quality Control*, 9, 1—7.
- Klein, J. P., Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data. Second Edition*. New York: Springer-Verlag.
- Kleinbaum, D., Klein, M. (2005). *Survival Analysis. A Self-Learning Text*. New York: Springer.

- Kotowska, I. E., Sztanderska, U., Wóycicka, I. (red.). (2007). *Aktywność zawodowa i edukacyjna a obowiązki rodzinne w Polsce w świetle badań empirycznych*. Warszawa: Wydawnictwo Naukowe SCHOLAR.
- Landmesser, J. M. (2008). Modele ryzyka konkurencyjnego dla czasu trwania czynności. W: T. Trzaskalik (red.), *Modelowanie Preferencji a Ryzyko '08* (s. 371—385). Katowice: Wydawnictwo Akademii Ekonomicznej im. Karola Adamieckiego w Katowicach.
- Landmesser, J. (2013). *Wykorzystanie metod analizy czasu trwania do badania aktywności ekonomicznej ludności w Polsce*. Warszawa: Wydawnictwo SGGW.
- Markowicz, I. (2012). *Statystyczna analiza żywotności firm*. Szczecin: Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.
- Matuszyk, A. (2015). *Zastosowanie analizy przetrwania w ocenie ryzyka kredytowego klientów indywidualnych*. Warszawa: CeDeWu.
- Pepe, M. S. (1991). Inference for Events With Dependent Risks in Multiple Endpoint Studies. *Journal of the American Statistical Association*, 86(415), 770—778. DOI: 10.1080/01621459.1991.10475108.
- Sączewska-Piotrowska, A. (2016). Dynamika ubóstwa w miejskich i wiejskich gospodarstwach domowych. *Wiadomości Statystyczne*, (7), 39—59.
- Wiśniewski, Z., Zawadzki, K. (red.). (2010). *Aktywna polityka rynku pracy w Polsce w kontekście europejskim*. Toruń: UMK-WUP.
- Wycinka, E. (2015). Modelowanie czasu do zaprzestania spłat rat kredytu lub wcześniejszej spłaty kredytu jako zdarzeń konkurujących. *Problemy Zarządzania*, 13(3), 146—157.

Summary. Numerous researches indicate that the situation of men and women on the labour market varies. In order to equalise opportunities, labour offices offer various forms of activation to the unemployed. However, it was observed that besides deregistering people taking up employment, removal of persons, who did not accept offered position, from the register was a common reason.

The aim of the research is to assess the influence of gender of the unemployed persons on the intensity of deregistration due to taking up employment, cancelling or other causes. These reasons constitute various types of competing events. In order to assess the probability of deregistration, the Kaplan-Meier estimator was used and to determine the intensity of deregistration for a specific reason — the Lunn-McNeil model. In the research individual data of persons registered in the Powiat Labour Office in Szczecin in the years 2013 and 2014 were used.

The results of the research show that intensity of finding a job was higher for women than for men. Women were also more rarely cancelled from the registry. In the case of other reasons, differences in the intensity of deregistration in terms of sex were not statistically significant.

Keywords: Kaplan-Meier estimator, Lunn-McNeil model, competing risks, unemployment, sex.