

## ZMIANY WZORCÓW WYCHODZENIA Z BEZROBOCIA NA PRZYKŁADZIE SZCZECINA<sup>1</sup>

**Beata Bieszk-Stolorz**  <https://orcid.org/0000-0001-8086-9037>

Instytut Ekonomii i Finansów  
Uniwersytet Szczeciński  
e-mail: beata.bieszk-stolorz@usz.edu.pl

**Joanna Landmesser-Rusek**  <https://orcid.org/0000-0001-7286-8536>

Instytut Ekonomii i Finansów  
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie  
e-mail: joanna\_landmesser@sggw.pl

**Iwona Markowicz**  <https://orcid.org/0000-0003-1119-0789>

Instytut Ekonomii i Finansów  
Uniwersytet Szczeciński  
e-mail: iwona.markowicz@usz.edu.pl

**Streszczenie:** W artykule podjęto próbę oceny wpływu wybranych cech demograficznych oraz społeczno-ekonomicznych na czas trwania bezrobocia rejestrowanego w Szczecinie w latach 2011-2023. W badaniu wykorzystano metody analizy przeżycia, w tym model przyspieszonej porażki. Analiza danych w kolejnych latach posłużyła odpowiedzi na pytanie: czy wzorce wychodzenia z bezrobocia do pracy ulegały zmianom w czasie. Cel ten osiągnięto poprzez weryfikację hipotez badawczych. Założono, że przyjęte w badaniu zmienne mają wpływ na czas trwania bezrobocia oraz że wpływ ten nie był jednakowy w całym okresie badawczym.

**Słowa kluczowe:** bezrobocie, analiza przeżycia, model AFT

**JEL classification:** J64, C41

---

<sup>1</sup> Badania przeprowadzono w ramach stażu naukowego (2023) *Wykorzystanie metod analizy trwania w badaniach zjawisk społeczno-ekonomicznych na rynku pracy* – współpraca SGGW i US.

## WSTĘP

Bezrobocie jest istotnym problemem społecznym i ekonomicznym. Dotyczy całego społeczeństwa i poszczególnych osób. Bezrobocie ma negatywny wpływ na jakość życia osób, które chcą podjąć pracę. Badania skutków bezrobocia mają charakter interdyscyplinarny i dotyczą różnych aspektów. Z jednej strony analizowane są zmiany w sytuacji jednostki jako skutek bezrobocia, takie jak zdrowie fizyczne i psychiczne, sytuacja finansowa, relacje w rodzinie, sytuacja dzieci [Karwacki, Błędowski 2020]. Z drugiej strony badania dotyczą wpływu bezrobocia na kapitał ludzki i społeczny [Nichols, Mitchell, Lindner 2013], w szczególności w okresie pandemii Covid-19 [Petcu i Cismasu 2023]. Podjęty w artykule temat ma duże znaczenie, zarówno na poziomie indywidualnym jak i społecznym.

W literaturze przedmiotu coraz częściej można znaleźć prezentację wyników badań dotyczących bezrobocia z wykorzystaniem metod analizy trwania. Cooper, McCausland i Theodossiou [2006] wykazali negatywny wpływ bezrobocia na czas pozostawania w dobrym zdrowiu, stosując model przyspieszonej porażki (Accelerated Failure Time, AFT). Autorzy na podstawie metody Kaplana-Meiera wykazali też różnice w przebiegu funkcji przeżycia w dobrym zdrowiu osób bezrobotnych w poszczególnych krajach (13 krajów UE w latach 1994-2001). Dogan [2019] wykorzystał model przyspieszonej porażki do badania wpływu zasiłków dla bezrobotnych na czas trwania w bezrobociu we Francji i Polsce. Dla obu krajów wykazano wydłużenie czasu trwania bezrobocia jako skutek otrzymywania zasiłku. Inne badanie [Dogan 2020], także z zastosowaniem modelu AFT, wykazało że prawidłowość zaobserwowana we Francji i Polsce dotyczy także Hiszpanii i Turcji, a na rynku holenderskim obserwowana jest zależność odwrotna, otrzymywanie zasiłków skraca czas trwania bezrobocia. Innym przykładem zastosowania modelu AFT w analizie bezrobocia jest praca [Grzenda 2019] dotycząca czasu trwania w bezrobociu długotrwałym w Polsce. Autorka wykazała, że dłuższy czas wychodzenia z bezrobocia dotyczy kobiet i osób opiekujących się dziećmi. Natomiast w pracy [Grzenda 2023] wskazano na coraz większe trudności z powrotem na rynek pracy wśród osób starszych. Modele hazardu, a w tym również model AFT, stosuje Landmesser [2009] w analizie czasu trwania w bezrobociu ludności w miastach i na wsi, z uwzględnieniem innych cech osób bezrobotnych a także do oceny wpływu szkolenia zawodowego na czas trwania bezrobocia [Landmesser 2011]. Mała i Cabla [2022] w badaniach czasu trwania w bezrobociu również stosują modele hazardu. Wykazali oni, że w okresie kryzysowym na czeski rynek pracy szybciej wchodził krótkotrwale bezrobotni. Cabla [2012] przy zastosowaniu modelu AFT wykazał natomiast istnienie związku między czasem trwania bezrobocia a płcią, stanem cywilnym, wiekiem i statusem niepełnosprawności. Metody analizy przeżycia wykorzystał w swoich badania Nagler [2013] i stwierdził, że wprowadzenie ubezpieczenia na wypadek bezrobocia

w Chile doprowadziło do skrócenia czasu trwania zatrudnienia. W pracy Bieszk-Stolorz i Markowicz [2022] stwierdzono szybsze wychodzenie z bezrobocia w okresie kryzysu zdrowotnego związanego z pandemią Covid-19.

Przedmiotem badań jest weryfikacja wpływu czynników na czas trwania bezrobocia w Szczecinie. Analiza dotyczy lat 2011-2021. Wykorzystano metod analizy przeżycia, w tym model przyspieszonej porażki, aby wykazać wpływ wybranych czynników (wiek, płeć, wykształcenie i staż pracy badanych bezrobotnych) na czas trwania bezrobocia. Analiza danych w kolejnych latach posłużyła odpowiedzi na pytanie: czy wzorce wychodzenia z bezrobocia ulegały zmianom w czasie. Cel ten osiągnięto poprzez weryfikację postawionych dwóch hipotez badawczych:

H1: Płeć, wiek, wykształcenie i staż pracy osoby bezrobotnej miały wpływ na czas trwania bezrobocia

H2: Wpływ płci, wieku, wykształcenia i stażu pracy osoby bezrobotnej na czas trwania bezrobocia nie był jednakowy w całym okresie badawczym.

## METODYKA BADANIA

W badaniu wykorzystano metody analizy trwania. Metody te są coraz częściej stosowane w naukach ekonomicznych. Ich zaletą jest fakt, że dane mogą mieć charakter kompletnych bądź niekompletnych. Jedną z przyczyn niekompletności danych jest zjawisko cenzurowania jednostek, które wiąże się z ich eliminacją z pola obserwacji. Podczas śledzenia zbioru obiektów mogą pojawić się jednostki, którym będzie towarzyszyć zdarzenie kończące ich obserwację. Mogą też pojawić się jednostki, w przypadku których takie zdarzenie nie wystąpi do zakończenia procesu obserwacji. W tej drugiej sytuacji jednostki nazywamy cenzurowanymi. Cenzurowanie może się odbywać ze względu na czas lub ze względu na zakończenie badania w określonym terminie. Wśród metod analizy trwania wyróżnia się modele nieparametryczne, semiparametryczne i parametryczne [Bieszk-Stolorz, Markowicz 2019; Bieszk-Stolorz, Landmesser, Markowicz 2020].

Podstawą analizy trwania jest funkcja przeżycia zdefiniowana następująco [Kleinbaum, Klein 2012]:

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

gdzie:  $T$  – czas trwania zjawiska,  $F(t)$  – dystrybuanta rozkładu zmiennej losowej  $T$ .

Funkcja przeżycia (trwania) określa prawdopodobieństwo tego, że określone zdarzenie nie zajdzie co najmniej do czasu  $t$ . Korzystając ze wzoru (1) można wyznaczyć kwantyle czasu trwania. Są to momenty czasu, dla których funkcja trwania przyjmuje odpowiednio wartości 0,75, 0,5, 0,25. Ze względu na możliwość wystąpienia obserwacji cenzurowanych kwantyle mogą nie istnieć. Wynika to z tego, że zdarzenie może nie wystąpić w przyjętym okresie obserwacji.

Najczęściej stosowanym nieparametrycznym estymatorem funkcji trwania jest estymator Kaplana-Meiera określony wzorem [Kaplan, Meier 1958]:

$$\hat{S}(t_i) = \prod_{j=1}^i \left(1 - \frac{d_j}{n_j}\right) \text{ dla } i = 1, 2, \dots, k, \quad (2)$$

gdzie:  $t_i$  – punkt czasu, w którym wystąpiło co najmniej jedno zdarzenie,  $t_1 < t_2 < \dots < t_k$ ,  $t_0 = 0$ ,  $d_i$  – liczba zdarzeń w czasie  $t_i$ ,  $n_i$  – liczba jednostek objętych obserwacją w czasie  $t_i$ ,  $n_i = n_{i-1} - d_{i-1} - z_{i-1}$ ,  $z_i$  – liczba obserwacji cenzurowanych w czasie  $t_i$ .

Krzywą trwania można wyznaczyć dla wszystkich obserwacji ogółem, a także dla grup wyodrębnionych ze względu na cechy jednostek. Krzywe takie można porównywać. Istnieje wiele testów do badania istotności różnic pomiędzy dwiema krzywymi przeżycia i różna jest ich moc. Weryfikowana jest hipoteza zerowa postaci  $H_0: S_1(t) = S_2(t)$ . Hipoteza alternatywna może przyjąć jedną z następujących postaci:  $H_1: S_1(t) \neq S_2(t)$ ,  $H_1: S_1(t) > S_2(t)$  lub  $H_1: S_1(t) < S_2(t)$ . Nie ma kryteriów pozwalających rozstrzygnąć, który z testów ma największą moc i powinien być wykorzystany w danym badaniu. Niektóre z nich są bardziej czułe na przebieg krzywych trwania w początkowej ich fazie, a inne w końcowej. Badania przeprowadzone przez Latta [1981] wskazują, że moc testów zmienia się również w zależności od wielkości badanej próby, od sposobu cenzurowania i rozkładu funkcji hazardu. Do porównania przebiegu krzywych przeżycia dla dwóch grup wykorzystano dwa testy: test Peto i Peto [Peto, Peto 1972] oraz test log-rank [Mantel, Haenszel 1959; Mantel 1966]. Test Peto i Peto przypisuje wagi zależne od oszacowanego percentyla rozkładu czasu trwania [Karadeniz, Ercan 2017]. Zdarzenia zachodzące wcześniej, gdy szacowana funkcja przeżycia jest duża, otrzymują większe wagi, natomiast zdarzenia znajdujące się w prawym ogonie rozkładu czasu trwania – mniejsze [Prentice, Marek 1979]. Test ten stosuje się, gdy współczynnik ryzyka pomiędzy grupami nie jest stały (Stevenson, 2009). Test log-rank jest najczęściej stosowanym testem do porównywania krzywych trwania [Karadeniz, Ercan 2017]. Przywiązuje taką samą wagę do wczesnych i późnych zdarzeń [Stevenson 2009; Allison 2010]. Test log-rank zakłada, że funkcje hazardu dla obu grup są równoległe, co oznacza, że współczynniki ryzyka dwóch grup są stałe we wszystkich okresach [Dawson, Trapp 2001; Stevenson 2009].

Wiele parametrycznych modeli przeżycia to modele przyspieszonej porażki (AFT), w których czas trwania jest wyrażany jako funkcja zmiennych objaśniających. Założenie leżące u podstaw modeli AFT mówi, że dla dowolnej ustalonej wartości funkcji przeżycia  $S(t)$  stosunek czasów trwania dla dwóch jednostek jest stały, czyli  $T_1/T_2 = const$ . Oznacza to również, że  $S_1(t) = S_2(\psi t)$  dla  $t \geq 0$ . Występujące w tych modelach zmienne objaśniające powodują odpowiednie „kurczenie się” lub „wydłużenie” czasu. Przyjmując  $\psi = \exp(-\mathbf{x}'\boldsymbol{\beta})$ , założenie AFT można wyrazić jako  $S_1(t) = S_2(\exp(-\mathbf{x}'\boldsymbol{\beta}) t)$ .

Formalnie, model AFT można zapisać w następującej postaci:

$$\ln T = \mathbf{x}'\boldsymbol{\beta} + \sigma\varepsilon \quad (3)$$

gdzie:  $T$  – czas trwania,  $\mathbf{x}$  – wektor zmiennych objaśniających,  $\boldsymbol{\beta}$  – wektor parametrów modelu,  $\varepsilon$  – składnik losowy,  $\sigma$  – czynnik skalujący.

Różne rozkłady składnika losowego  $\varepsilon$  prowadzą do różnych postaci modeli AFT. Po przekształceniu równania (1) otrzymuje się  $T \exp(-\mathbf{x}'\boldsymbol{\beta}) = \exp(\sigma\varepsilon)$ , czyli  $T\psi = \exp(\sigma\varepsilon)$ . Parametr  $\psi$  nosi nazwę parametru przyspieszającego, a działanie jego polega na skalowaniu czasu. U jednostki, dla której  $\psi > 1$ , czas jest przyspieszany i zdarzenie wystąpi szybciej; z kolei  $\psi < 1$  oznacza, że czas jest spowalniany, a zdarzenie nastąpi z opóźnieniem. Podejmowana w modelach AFT parametryzacja dotyczy rozkładu składnika losowego  $\varepsilon$ . Jeśli dla  $\varepsilon$  przyjmie się rozkład wartości ekstremalnych z jednym parametrem, to czas trwania  $T$  będzie podlegał rozkładowi wykładniczemu, jeśli zaś  $\varepsilon$  będzie miał rozkład wartości ekstremalnych z dwoma parametrami, to  $T$  zyska rozkład Weibulla. Założenie dla  $\varepsilon$  rozkładu logistycznego wiąże się z rozkładem log-logistycznym dla  $T$ , zaś rozkład normalny dla  $\varepsilon$  skutkuje dla  $T$  rozkładem logarytmiczno-normalnym.

W wykładniczym modelu AFT funkcja przeżycia ma postać  $S(t) = \exp(-\exp(-\mathbf{x}'\boldsymbol{\beta})t)$ , zaś w modelu Weibulla  $S(t) = \exp(-\exp(-p\mathbf{x}'\boldsymbol{\beta})t^p)$ ,  $p > 0$ . Oba te modele spełniają założenia zarówno AFT jak i proporcjonalnego hazardu. Wiele zjawisk społecznych charakteryzuje niska początkowa intensywność procesu, rosnąca do pewnego momentu, a następnie malejąca (np. wychodzenie ze stanu bezrobocia). W takiej sytuacji zastosowanie może znaleźć rozkład logarytmiczno-normalny lub log-logistyczny. W modelu logarytmiczno-normalnym funkcja przeżycia wyraża się jako  $S(t) = 1 - \Phi\left(\frac{\ln t - \mathbf{x}'\boldsymbol{\beta}}{\sigma}\right)$ , gdzie  $\Phi(z)$  jest funkcją dystrybuanty dla standardowego rozkładu normalnego. W modelu log-logistycznym funkcja przeżycia to  $S(t) = \frac{1}{1 + (\lambda t)^{1/\gamma}}$ , gdzie  $\lambda = \exp(-\mathbf{x}'\boldsymbol{\beta})$  oraz  $\gamma > 0$ . Gdy  $\gamma \geq 1$  funkcja hazardu jest monotonicznie malejąca; gdy  $\gamma < 1$  funkcja hazardu najpierw rośnie, potem maleje. Log-logistyczny model AFT nie spełnia założenia modelu proporcjonalnych hazardów, spełnia natomiast założenie modelu proporcjonalnych „odds” [Kleinbaum, Klein 2012]. Wyboru najlepszego modelu przeżycia spośród kilku alternatywnych modeli o różnych rozkładach podstawowych można dokonać na podstawie wartości kryterium informacyjnego Akaike’a (AIC).

## DANE WYKORZYSTANE W BADANIU

W badaniu wykorzystano dane indywidualne z Powiatowego Urzędu Pracy (PUP) w Szczecinie. Badanie zatem dotyczyło bezrobocia rejestrowanego. Należy tutaj zaznaczyć, że źródłem takich danych w Polsce są jedynie powiatowe urzędy pracy. Główny Urząd Statystyczny w Polsce i Eurostat dysponują jedynie danymi zagregowanymi, które są nieprzydatne w przypadku badań z wykorzystaniem

modeli analizy trwania. Z tego powodu dane wykorzystane w badaniu są unikatowe. Pozyskane dane obejmowały 193919 osób wyrejestrowanych z urzędu pracy w latach 2011-2021. Dane te zawierały datę rejestracji oraz datę wyrejestrowania, płeć, poziom wykształcenia, wiek osoby wyrejestrowanej, staż pracy, a także powód wyrejestrowania. Powody wyrejestrowania mogą być różne. Jest ich kilkadziesiąt. Ponieważ analizowano zdarzenie polegające na podjęciu pracy, więc wszystkich bezrobotnych podzielono na dwie grupy: tych, którzy podjęli pracę oraz tych, którzy zostali wyrejestrowani z pozostałych przyczyn (np. rezygnacja z pośrednictwa urzędu, wyjazd za granicę, przejście na rentę lub emeryturę). Na podstawie daty zarejestrowania i wyrejestrowania wyznaczono czas bezrobocia rejestrowanego, czyli zmienną losową  $T$ . Zdarzeniem kończącym obserwację jest podjęcie pracy. Wyrejestrowanie z innych powodów zostało przyjęte za obserwację cenzurowaną. Analizę przeprowadzono dla 11 lat wchodzących w skład okresu badawczego. 55,33% wszystkich obserwacji stanowiły obserwacje cenzurowane. W tabeli 1 przedstawiono cechy osób bezrobotnych wraz z podziałem na kategorie oraz zanzaczono grupy referencyjne.

Tabela 1. Cechy osób bezrobotnych i ich kategorie

Cecha	Kategoria	Cecha	Kategoria
Płeć	Kobieta		Co najwyżej gimnazjalne*
	Mężczyzna*		Zasadnicze zawodowe
Grupa wieku	18-24*	Poziom wykształcenia	Średnie ogólnokształcące
	25-34		Średnie zawodowe
	35-44		Wyższe
	45-54	Staż pracy	Brak stażu*
	55-59		Posiadanie stażu
	60-64		

\* – kategoria referencyjna

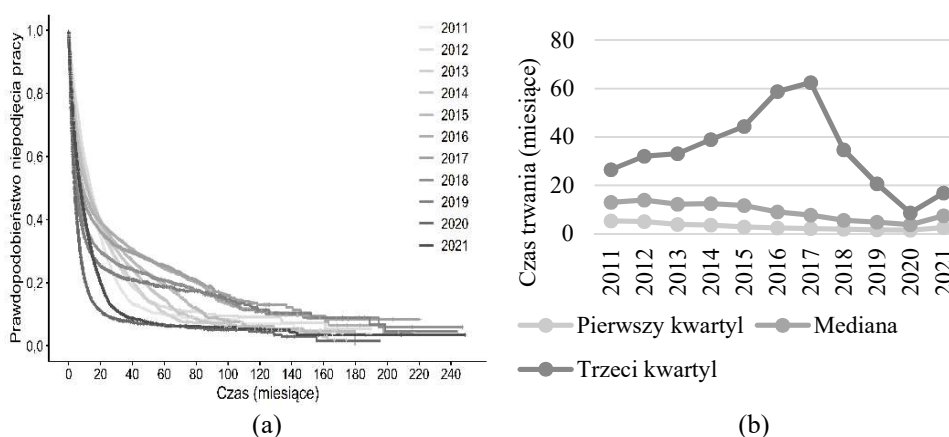
Źródło: opracowanie własne

## WYNIKI BADANIA

Badanie przeprowadzono w dwóch etapach. Pierwszy etap obejmował wstępną ocenę czasu wychodzenia z bezrobocia rejestrowanego. Na rysunku 1(a) przedstawiono oszacowania Kaplana-Meiera dla krzywych przeżycia dla 11 analizowanych lat (2011-2021). W tym przypadku krzywe przeżycia (trwania) opisują prawdopodobieństwo niewyrejestrowania z powodu podjęcia pracy, czyli inaczej mówiąc prawdopodobieństwo niepodjęcia pracy. Na rysunku 1(b) przedstawiono kwartyle czasu trwania wyznaczone na podstawie oszacowań Kaplana-Meiera. Krzywe przeżycia przedstawione na rysunku 1(a) różnią się przebiegiem. Różnice w czasach trwania są również widoczne na rysunku 1(b). Wartości pierwszego kwartyła i mediany malały do roku 2020. W 2021 roku

nastąpił wzrost ich wartości. Nieco inne były zmiany wartości trzeciego kwartyla. Jego wartości rosły do 2017 roku, następnie malały do 2020 roku. W 2021 roku nastąpił wzrost wartości trzeciego kwartyla w porównaniu do roku poprzedniego. Wynika z tego, że czas do podjęcia pracy przez osoby zarejestrowane w PUP ulegał systematycznemu skróceniu, co jest zjawiskiem pożądanym. Jednak pogarszała się sytuacja osób długotrwale bezrobotnych. Od 2017 roku widoczne jest skrócenie czasu do podjęcia pracy przez wszystkich bezrobotnych. W najlepszym, 2020 roku, 75% osób wyrejestrowanych z PUP, to osoby o czasie zarejestrowania nie dłuższym niż 8,56 miesiąca.

Rysunek 1. Oszacowania Kaplana-Meiera dla krzywych przeżycia (ocena prawdopodobieństwa niepodjęcia pracy) i kwartyle czasu trwania



Źródło: opracowanie własne w programie R (a) i programie Excel (b)

Różnice w przebiegu krzywych trwania zostały potwierdzone testem dla wielu krzywych ( $\chi^2 = 4814$  dla 10 stopni swobody,  $p < 2e-16$ ). Test ten jest rozwinięciem (lub uogólnieniem) testu Gehana, Peto i Peto oraz log-rank. Jest on oparty na statystyce chi-kwadrat. Korzystając z testu Peto i Peto oraz testu log-rank porównano parami wszystkie krzywe trwania. Wyniki przedstawiono w tabeli 2. Na ich podstawie należy stwierdzić, że w większości przypadków należy odrzucić hipotezę zerową o równości krzywych trwania. Jednak w kilku przypadkach brak jest podstaw do odrzucenia hipotezy zerowej. Wartości testu Peto i Peto wskazują na to, że przebieg krzywych trwania 2011 i 2012, 2013 i 2014, w ich początkowej fazie nie różni się istotnie od siebie. Natomiast wartości testu log-rank wskazują na to, że w całym przebiegu krzywe trwania 2014 i 2015, 2018 i 2021 nie różnią się istotnie od siebie.

Tabela 2. Porównanie krzywych przeżycia – wartości  $p$  testu Peto i Peto (log-rank)

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
2012	0,1426 (0,0010)									
2013	0,0000 (0,0001)	0,0000 (0,0000)								
2014	0,0000 (0,0061)	0,0000 (0,0000)	0,4726 (0,0240)							
2015	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,1514)	0,0000 (0,0008)						
2016	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0004)					
2017	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0004 (0,0007)				
2018	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)			
2019	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0078 (0,0002)		
2020	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	
2021	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0000 (0,0000)	0,0149 (0,0000)	0,0000 (0,4857)	0,0000 (0,0000)	0,0000 (0,0000)

Źródło: opracowanie własne

W drugim etapie badania wykorzystano modele przyspieszonej porażki (AFT). Spośród możliwych do opisu procesu podejmowania pracy przez osoby bezrobotne modeli AFT rozpatrzono cztery: wykładniczy, Weibulla, log-normalny i log-logistyczny. W pierwszym kroku oszacowano równania modeli dla całego badanego okresu czasu (2011-2021) na podstawie połączonej próby (obserwacji nie grupowano w żaden sposób) uzyskując najlepszy wynik dla modelu log-logistycznego (ten model charakteryzowała najmniejsza wartość kryterium AIC; patrz tabela 3).

Tabela 3. Współczynniki w modelach AFT szacowanych na podstawie połączonych danych (lata 2011-2021 razem)

Zmienna	Rozkłady			
	wykładniczy	Weibulla	log-normalny	log-logistyczny
plec	0,256***	0,222***	0,174***	0,205***
wiek2	0,378***	0,343***	0,251***	0,277***
wiek3	0,607***	0,549***	0,379***	0,404***
wiek4	0,683***	0,591***	0,334***	0,399***
wiek5	0,992***	0,935***	0,569***	0,710***
wiek6	1,749***	1,815***	1,363***	1,571***
wyksz2	-0,340***	-0,444***	-0,398***	-0,434***
wyksz3	-0,460***	-0,604***	-0,533***	-0,594***
wyksz4	-0,509***	-0,680***	-0,581***	-0,658***
wyksz5	-0,921***	-1,175***	-1,063***	-1,119***
staz	-0,261***	-0,347***	-0,349***	-0,324***



cons	2,743***	3,092***	2,605***	2,563***
p		0,740***		
sigma			1,879***	
gamma				1,047***
Obs	191884	191884	191884	191884
lnL	-232705,1	-224803,7	-224475,3	-223878,9
AIC	465434,2	449633,5	448976,6	<b>447783,8</b>

\*\*\* – poziom istotności 0,01

Źródło: opracowanie własne w programie R

Następnie wszystkie powyższe modele AFT szacowano w podokresach rocznych. W tabeli 4 przedstawiono liczbę obserwacji oraz wartości kryterium AIC dla oszacowanych 44 modeli. Można zauważyć, że monotoniczny hazard Weibulla był najbardziej adekwatny dla danych z lat 2011-2014, następnie w latach 2015-2017 najlepiej wypadł model logarytmiczno-normalny opisujący niemonotoniczny hazard, a dla lat 2018-2021 najlepszym modelem okazał się model log-logistyczny, który może opisywać hazard zarówno monotoniczny jak i niemonotoniczny.

Tabela 4. Liczba obserwacji i wartość AIC w szacowanych modelach AFT (wyłuszczone najniższe wartości dla kryterium AIC)

Rok	Liczba obserwacji	Obserwacje niepełne (niepodjęcie pracy)	Rozkłady			
			wykładniczy	Weibulla	log-normalny	log-logistyczny
2011	15194	60,82%	29484,05	<b>29457,36</b>	30436,40	29752,33
2012	15570	61,27%	31137,57	<b>30978,02</b>	31711,41	31159,57
2013	23762	53,87%	56404,34	<b>55539,87</b>	56303,12	55800,91
2014	24443	55,09%	56872,58	<b>55803,42</b>	56156,18	55920,95
2015	25568	56,72%	62502,88	59845,83	<b>59732,64</b>	59742,03
2016	23447	57,55%	59790,07	55632,09	<b>54828,85</b>	55038,44
2017	19697	59,71%	46912,42	44151,29	<b>43562,79</b>	43599,24
2018	14873	58,15%	35806,34	33752,23	33070,93	<b>33046,76</b>
2019	12680	55,50%	29829,79	28845,56	27980,29	<b>27975,22</b>
2020	7772	31,24%	22084,51	21842,82	21505,41	<b>21424,95</b>
2021	8878	36,89%	24753,41	24553,84	24623,19	<b>24512,29</b>

Źródło: opracowanie własne w programie R

W celu prześledzenia wpływu zmiennych objaśniających na wyjście osób z bezrobocia w PUP Szczecin w okresie 2011-2021, zdecydowano się dalszą analizę ograniczyć do jednego typu modeli AFT – modelu log-logistycznego. Model ten jest elastyczny pod względem opisu przebiegu hazardu oraz dla czterech spośród modelowanych 11 lat okazał się najlepszy. Tabela 5 prezentuje uzyskane w procesie estymacji wartości współczynników beta oraz gamma. Przykładowa interpretacja oceny parametru stojącego przy zmiennej płeć w modelu dla 2011 roku jest następująca: kobiety, w porównaniu z mężczyznami miały dłuższy czas wychodzenia z bezrobocia do pracy (ponieważ  $t_K = \exp(0,085)t_M = 1,089t_M$ ).

Dla wszystkich lat parametr ten przyjmował wartość dodatnią, wskazując tym samym na trudniejszą pozycję na rynku pracy kobiet niż mężczyzn, zaś w latach 2016-2019 sytuacja ta wydaje się najtrudniejsza.

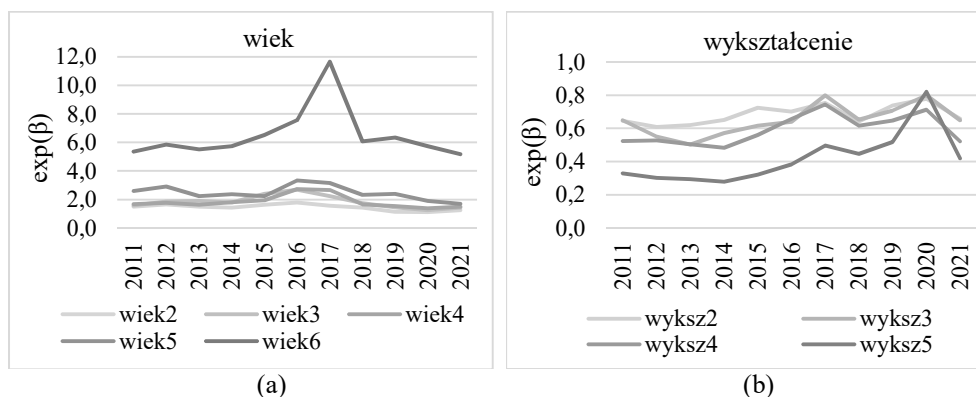
Tabela 5. Oceny parametrów w modelach log-logistycznych szacowanych dla lat 2011-2021

Zmienna	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
plec	0,085***	0,154***	0,087***	0,118***	0,216***	0,398***	0,404***	0,297***	0,339***	0,194***	0,113***
wiek2	0,408***	0,501***	0,402***	0,361***	0,490***	0,580***	0,445***	0,353***	0,127**	0,110	0,226***
wiek3	0,462***	0,623***	0,648***	0,587***	0,869***	0,989***	0,802***	0,562***	0,368***	0,228***	0,360***
wiek4	0,515***	0,568***	0,496***	0,594***	0,665***	1,004***	0,983***	0,506***	0,429***	0,321***	0,406***
wiek5	0,957***	1,072***	0,807***	0,861***	0,809***	1,207***	1,147***	0,845***	0,876***	0,645***	0,534***
wiek6	1,676***	1,765***	1,707***	1,746***	1,874***	2,023***	2,456***	1,803***	1,848***	1,746***	1,644***
wyksz2	-0,437***	-0,496***	-0,476***	-0,429***	-0,322***	-0,354***	-0,282***	-0,445***	-0,303***	-0,251***	-0,416***
wyksz3	-0,432***	-0,598***	-0,689***	-0,557***	-0,484***	-0,447***	-0,224***	-0,427***	-0,348***	-0,224***	-0,432***
wyksz4	-0,649***	-0,639***	-0,684***	-0,728***	-0,579***	-0,420***	-0,295***	-0,483***	-0,434***	-0,337***	-0,652***
wyksz5	-1,112***	-1,199***	-1,227***	-1,278***	-1,136***	-0,959***	-0,702***	-0,805***	-0,660***	-0,198***	-0,871***
staz	-0,400***	-0,462***	-0,785***	-0,666***	-0,721***	-1,012***	-0,720***	0,101**	0,190***	0,025	0,040
cons	2,870***	2,895***	3,099***	2,990***	2,661***	2,432***	1,962***	1,549***	1,371***	1,107***	1,971***
gamma	0,859***	0,918***	0,982***	0,992	1,120***	1,185***	1,124***	1,054***	0,911***	0,773***	0,843***

\*\*\* – poziom istotności 0,01, \*\* – poziom istotności 0,05

Źródło: opracowanie własne w programie R

Rysunek 2. Wpływ wieku oraz wykształcenia na czas trwania w bezrobociu (wartości  $\exp(\beta_k) > 1$  oznaczają wydłużenie czasu trwania w porównaniu z grupą referencyjną,  $\exp(\beta_k) < 1$  jego skrócenie)



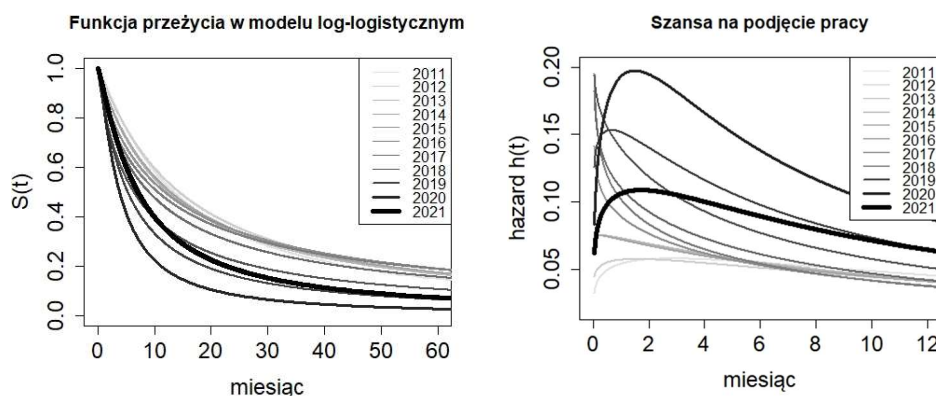
Źródło: opracowanie własne

Analiza wpływu wieku na czas trwania w bezrobociu pozwala zauważyć, że wraz ze wzrostem wieku czas bez pracy się wydłużał. W najgorszej pod tym względem sytuacji w odniesieniu do grupy referencyjnej (osoby w wieku od 18 do 24 lat) były osoby najstarsze (w wieku od 60 do 65 lat). Grupa osób najstarszych znajdowała się w wyjątkowo trudnym położeniu w latach 2016 i 2017, co dodatkowo dokumentuje rysunek 2(a). Na podstawie oszacowanych parametrów w

modelach AFT można też stwierdzić, że czas wychodzenia do pracy skracał się (świadczą o tym ujemne oceny parametrów) wraz ze wzrostem poziomu wykształcenia. Na uwagę jednak zasługuje pogorszenie się sytuacji osób z wykształceniem wyższym w roku 2020 (patrz rysunek 2(b)). Osoby z udokumentowanym stażem zawodowym miały krótszy czas wychodzenia z bezrobocia w latach 2011-2017, zaś od roku 2018 efekt stażu był przeciwny, przy czym w modelach dla lat 2020 i 2021 wystąpił brak statystycznej istotności parametru przy tej zmiennej.

Szacunki parametru gamma w modelach log-logistycznych wskazują na niemonotoniczną (najpierw rosnącą, potem malejącą w miarę trwania w bezrobociu przez kolejne miesiące) intensywność podejmowania zatrudnienia w latach 2011-2014, monotoniczną (malejącą) intensywność w latach 2015-2018 oraz ponownie niemonotoniczną w okresie 2019-2021. Przebiegi krzywych przeżycia oraz hazardu (wyrażającego intensywność wychodzenia z bezrobocia w zatrudnienie) dla poszczególnych lat wynikające z modelu AFT log-logistycznego przedstawiono na rysunku 3. Wynika z niego, że w najlepszej sytuacji bezrobotni znajdowali się w roku 2020.

Rysunek 3. Przebiegi funkcji przeżycia oraz hazardu dla zjawiska wychodzenia z bezrobocia w stan zatrudnienia w PUP Szczecin w latach 2011-2021



Źródło: opracowanie własne

## PODSUMOWANIE

Celem przeprowadzonych badań była ocena wpływu czynników (wiek, płeć, wykształcenie i staż pracy badanych bezrobotnych) na czas trwania bezrobocia w Szczecinie w latach 2011-2021. Analiza danych czasowych dała możliwość sprawdzenia, czy wzorce wychodzenia z bezrobocia ulegały zmianom. Założono, że przyjęte w badaniu zmienne mają wpływ na czas trwania bezrobocia oraz że wpływ ten nie był jednakowy w całym okresie badawczym. Przebieg

wyznaczonych krzywych przeżycia (estymator Kaplana-Meiera) wskazuje na różnice w czasie trwania bezrobocia w kolejnych badanych latach. Czas do podjęcia pracy ulegał systematycznemu skróceniu. Rozwinięciem wstępnego etapu badań było zastosowanie modelu przyspieszonej porażki z rozkładem log-logistycznym dla zmiennej  $T$ . Ogólnie uzyskane wyniki wskazują na wpływ badanych zmiennych na czas trwania bezrobocia. Okazało się, że dłuższy czas dotyczył kobiet w stosunku do mężczyzn, starszych bezrobotnych, z niższym wykształceniem, bez stażu pracy (zazwyczaj). Tym samym pozytywnie została zweryfikowana postawiona hipoteza badawcza H1. W latach 2012-2014 w Polsce zauważalny był wzrost bezrobocia, szczególnie wśród osób młodych. Był to skutek kryzysu gospodarczego z lat 2008-2009. Jednak nasze badania nie wskazują na trudniejszą sytuację młodych osób bezrobotnych w tym czasie. Ciekawym wnioskiem jest natomiast zauważona zmiana wzorców wychodzenia z bezrobocia na przestrzeni czasu, czyli pozytywnie została zweryfikowana również hipoteza badawcza H2. Wyodrębniono trzy charakterystyczne okresy: lata 2011-2014, 2015-2018 i 2019-2020. W okresach tych zmieniała się monotoniczność intensywności podejmowania zatrudnienia przez bezrobotnych. Szczególnym przypadkiem jest 2020 rok, w którym intensywność ta była duża w początkowym okresie trwania bezrobocia. Zatem kryzysowa sytuacja zdrowotna (Covid-19) nie wpłynęła na wydłużenie czasu poszukiwania pracy, a wręcz przeciwnie czas ten uległ skróceniu.

## BIBLIOGRAFIA

- Allison P. D. (2010) *Survival Analysis Using SAS: A Practical Guide*, 2nd edition, SAS Press, North Carolina.
- Bieszk-Stolorz B., Landmesser J., Markowicz I. (2020) *Analiza trwania w badaniach ekonomicznych. Modele parametryczne*. CeDeWu, Warszawa.
- Bieszk-Stolorz B., Markowicz I. (2019) *Analiza trwania w badaniach ekonomicznych. Modele nieparametryczne i semiparametryczne*. CeDeWu, Warszawa.
- Bieszk-Stolorz B., Markowicz I. (2022) The Impact of the COVID-19 Pandemic on the Situation of the Unemployed in Poland. A Study Using Survival Analysis Methods. *Sustainability*, 14(19), 12677. <https://doi.org/10.3390/su141912677>.
- Cabla A. (2012) Modeling Unemployment Duration in the Czech Republic from LFS. *Research Journal of Economics, Business and ICT*, 7, 1-5.
- Cooper D., McCausland W. D., Theodossiou I. (2006) The health hazards of unemployment and poor education: The socioeconomic determinants of health duration in the European Union. *Economics and Human Biology*, 4, 273-297. <https://doi.org/10.1016/j.ehb.2006.06.001>.
- Dawson B., Trapp R. G. (2001) *Basic&Clinical Biostatistics*. McGraw Hill, Boston.
- Dogan F. I. (2019) Unemployment Benefits and Unemployment Duration in France and Poland. *Marmara Journal of European Studies*, 27(1), 191-216. <https://doi.org/10.29228/mjes.30>.

- Dogan F. I. (2020) Social Transfers and Unemployment Duration: an Empirical Evidence of the EU and Turkey. *Marmara Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 42(2), 247-264. <https://doi.org/10.14780/muiibd.854363>.
- Grzenda W. (2019) Socioeconomic Aspects of Long-Term Unemployment in the Context of the Ageing Population of Europe: the Case of Poland. *Economic Research-Ekonomska Istraživanja*, 32(1), 1561-1582. <https://doi.org/10.1080/1331677X.2019.1638289>.
- Grzenda W. (2023) Estimating the Probability of Leaving Unemployment for Older People in Poland Using Survival Models with Censored Data. *Statistics in Transition new series*, 24(3), 241-256. <https://doi.org/10.59170/stattrans-2023-046>.
- Kaplan E. L., Meier P. (1958) Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282), 457-481. <https://doi.org/10.2307/2281868>.
- Karwacki A., Błędowski P. (2020) Bezrobocie jako współczesna kwestia społeczna – wybrane aspekty socjologiczne i ekonomiczne. *Studia Socjologiczne*, 1(236), 135-164.
- Kleinbaum D. G., Klein M. (2012) *Survival Analysis. A Self-Learning Text*. 3rd ed.; New York: Springer-Verlag. <https://doi.org/10.1007/978-1-4419-6646-9>.
- Karadeniz P. G., Ercan I. (2017) Examining Tests for Comparing Survival Curves with Right Censored Data. *Statistics in Transition*, 18(2), 311-28.
- Landmesser J. M. (2009) Econometric Analysis of Unemployment Duration Using Hazard Models. *Studia Ekonomiczne*, 1-2, 79-92. <https://www.ceeol.com/search/article-detail?id=201814>.
- Landmesser J. M. (2011) The Impact of Vocational Training on the Unemployment Duration. *International Advances in Economic Research*, 17, 89-100. <https://doi.org/10.1007/s11294-010-9291-y>.
- Latta R. B. (1981) A Monte Carlo Study of Some Two-Sample Rank Tests with Censored Data. *Journal of the American Statistical Association*, 76(375), 713-719. <https://doi.org/10.2307/2287536>.
- Mala I., Cabla A. (2022) Modelling of the Unemployment Duration in the Czech Republic Based on Aggregated Complete and Individual Censored Data. *Journal of Economics/Ekonomický časopis*, 70(2), 171-187.
- Mantel N., Haenszel W. (1959) Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease. *Journal of National Cancer Institute*, 22 (4), 719-748. <https://www.ncbi.nlm.nih.gov/pubmed/13655060>.
- Mantel N. (1966) Evaluation of Survival Data and Two New Rank Order Statistics Arising in its Consideration. *Cancer Chemotherapy Reports*, 50(3), 163-170. <https://www.ncbi.nlm.nih.gov/pubmed/5910392>.
- Nagler P. (2013) How Unemployment Insurance Savings Accounts Affect Employment Duration: Evidence from Chile. *IZA Journal of Labor & Development*, 2(9). <http://www.izajold.com/content/2/1/9>.
- Nichols A., Mitchell J., Lindner S. (2013) *Consequences of Long-Term Unemployment*. Washington: The Urban Institute Report.
- Petcu M. A., Cismasu I. D. (2023) Exploring Economic Effects of Work Out of Company Office Based on Bibliometric Analysis. *Journal of Business Economics and Management*, 24(4), 676-695. <https://doi.org/10.3846/jbem.2023.20247>.

- Peto R., Peto J. (1972) Asymptotically Efficient Rank Invariant Test Procedures. *Journal of the Royal Statistical Society*, 135(2), 185-207.  
[https://www.jstor.org/stable/2344317?seq=1#page\\_scan\\_tab\\_contents](https://www.jstor.org/stable/2344317?seq=1#page_scan_tab_contents).
- Prentice R. L., Marek P. (1979) A Qualitative Discrepancy between Censored Data Rank Tests. *Biometrics*, 35(4), 861-867.
- Stevenson M. (2009) *An Introduction to Survival Analysis*, EpiCentre, IVABS. Massey Massey University.  
[http://www.massey.ac.nz/massey/fms/Colleges/College%20of%20Sciences/Epicenter/docs/ASVCS/Stevenson\\_survival\\_analysis\\_195\\_721.pdf](http://www.massey.ac.nz/massey/fms/Colleges/College%20of%20Sciences/Epicenter/docs/ASVCS/Stevenson_survival_analysis_195_721.pdf).

### CHANGES IN PATTERNS OF LEAVING UNEMPLOYMENT: EVIDENCE FROM SZCZECIN

**Abstract:** The article attempts to assess the impact of selected factors (age, gender, education and work experience of the unemployed) on the duration of registered unemployment in Szczecin in years 2011-2023. The study used survival analysis methods, including an acceleration failure time model. The analysis of data in subsequent years was used to assess whether the patterns of leaving unemployment for work changed. This goal was achieved by verifying the research hypotheses. It was assumed that the variables used in the study had an impact on the duration of unemployment and that this impact was not the same throughout the research period.

**Keywords:** unemployment, survival analysis, AFT model

**JEL classification:** J64, C41