

## Badanie dynamiki ubóstwa gospodarstw domowych z wykorzystaniem wybranych modeli analizy historii zdarzeń

### Streszczenie

Celem artykułu jest identyfikacja czynników zwiększających i zmniejszających szanse wejścia do sfery ubóstwa i wyjścia z niej. W analizie zastosowano modele analizy historii zdarzeń – semiparametryczny model Coxa oraz modele o czasie dyskretnym (logitowy oraz komplementarny logistyczno-logistyczny). W modelach uwzględniono następujące zmienne objaśniające: płeć, wiek i wykształcenie głowy gospodarstwa domowego, klasę miejscowości zamieszkania, grupę społeczno-ekonomiczną oraz status gospodarstwa domowego na rynku pracy. Analiza została poszerzona o opisową analizę czasu przeżycia w sferze ubóstwa oraz czasu przeżycia poza sferą ubóstwa z wykorzystaniem estymatora Kaplana–Meiera. Prawdopodobieństwo przeżycia w sferze ubóstwa przez długi czas jest mniejsze niż prawdopodobieństwo przeżycia poza tą sferą. Gospodarstwa domowe utrzymujące się z niezarobkowych źródeł są w najgorszej sytuacji ekonomicznej, ponieważ mają małe szanse na wyjście ze sfery ubóstwa oraz duże szanse na wejście do niej.

**Słowa kluczowe:** dynamika ubóstwa, model Coxa, model logitowy, model komplementarny logistyczno-logistyczny, estymator Kaplana–Meiera

### 1. Wstęp

Ubóstwo jest zjawiskiem niekorzystnym, mogącym dotykać gospodarstwa domowe przez krótszy lub dłuższy czas. Dłuższy pobyt w sferze ubóstwa przyczynia się często do nawarstwienia problemów i pojawienia się kolejnych niekorzystnych zjawisk, np. wykluczenia społecznego. Im szybciej gospodarstwo domowe opuści tę sferę, tym większe ma szanse na szybką inkluzję społeczną. Z punktu widzenia prowadzonej przez państwo polityki społecznej ważne jest poznanie czynników zwiększających i zmniejszających szanse gospodarstwa

---

<sup>1</sup> Uniwersytet Ekonomiczny w Katowicach, Wydział Ekonomii.

domowego na wejście do sfery ubóstwa i wyjście z niej. Identyfikacja tych czynników jest celem niniejszego opracowania, a cel ten osiągnięto, stosując modele analizy historii zdarzeń – semiparametryczny model Coxa oraz modele o czasie dyskretnym (logitowy i komplementarny log-log). W modelach uwzględniono następujące zmienne objaśniające: płeć, wiek i wykształcenie głowy gospodarstwa domowego, klasę miejscowości zamieszkania, grupę społeczno-ekonomiczną i status gospodarstwa domowego na rynku pracy. Analiza historii zdarzeń jest ogólnym pojęciem odnoszącym się do grupy statystycznych metod pozwalających analizować czas oczekiwania na wystąpienie zdarzenia<sup>2</sup>, czyli tzw. czas przeżycia, czas trwania lub epizod. W przeprowadzonym badaniu wykorzystano dane panelowe, co oznacza, że czasy zdarzeń były mierzone w dyskretnych przedziałach, tym samym powinny być stosowane modele o czasie dyskretnym. Czas oczekiwania na wystąpienie zdarzenia potraktowano również jako zmienną o czasie ciągłym, co pozwoliło porównać wyniki badań z wynikami uzyskanymi na podstawie modeli o czasie dyskretnym będących pewnym przybliżeniem modeli o czasie ciągłym.

Analizę determinant poprzedzono opisową analizą (dla ogółu gospodarstw domowych) czasu przeżycia w sferze ubóstwa oraz w sferze poza ubóstwem z wykorzystaniem estymatora Kaplana–Meiera. Badanie dynamiki ubóstwa przeprowadzono na przykładzie Polski, wykorzystując dane z ośmiu fal panelu przeprowadzonego w ramach projektu „Diagnoza społeczna”.

W artykule przyjęto kilka hipotez badawczych. Można się spodziewać, że czas przeżycia poza sferą ubóstwa jest dużo dłuższy niż czas przeżycia w tej sferze. Ubóstwo jest zazwyczaj stanem, który trwa krótko (np. z powodu chwilowej utraty pracy), a tylko mały odsetek gospodarstw domowych nie ma szans (np. z powodu własnej niechęci do podwyższenia kwalifikacji w celu znalezienia pracy) zmiany przynależności do sfery ubóstwa. Można również przypuszczać, że na czas przeżycia w sferze ubóstwa i poza nią istotnie wpływają cechy gospodarstwa domowego i jego głowy. Sformułowano hipotezę głoszącą, że w najgorszej sytuacji ekonomicznej są gospodarstwa utrzymujące się z niezarobkowych źródeł – mają z jednej strony bardzo małe szanse na wyjście z biedy, z drugiej – bardzo duże szanse na wejście do tej sfery. Jest to grupa gospodarstw uzyskujących niskie dochody (głównie są to zasiłki) i jednocześnie często nawet niestarájących się poprawić swojej sytuacji ekonomicznej.

---

<sup>2</sup> M. Mills, *Survival and event history analysis*, Sage Publications, Los Angeles–London–New Dehli–Singapore–Washington 2011, s. 257.

Należy zaznaczyć, że w opracowaniu celowo nie skupiono się na problemach metodologicznych dotyczących pomiaru ubóstwa, ponieważ szczegółowy opis można znaleźć w literaturze przedmiotu. W Polsce problemy związane z pomiarem ubóstwa omawia m.in. T. Panek<sup>3</sup>.

## 2. Przegląd literatury

Prekursorami badań ubóstwa opartych na czasie przeżycia (ang. *spell-based approach*) byli M.J. Bane i D.T. Ellwood<sup>4</sup>. W swoich badaniach, przeprowadzonych na podstawie amerykańskiego badania panelowego dynamiki dochodów (*Panel Study on Income Dynamics – PSID*) z lat 1970–1982, skupili się na długości okresu ubóstwa oraz na przejściach ze sfery ubóstwa do sfery poza ubóstwem i odwrotnie. Wyniki badań wskazały, że wraz ze wzrostem czasu trwania ubóstwa malały szanse na wyjście z tej sfery.

A.H. Stevens<sup>5</sup> zwróciła uwagę na to, że wadą badań Ellwooda było skupienie się na pojedynczych epizodach. Przeprowadziła swoje badania na podstawie PSID za lata 1970–1988, biorąc pod uwagę wielokrotne czasy przeżycia w sferze ubóstwa oraz poza nią. Wyniki estymacji modeli hazardu pozwoliły jej na sformułowanie wniosku, że osoby, które wyszły ze sfery ubóstwa po długim okresie, są bardziej zagrożone ponownym wejściem do tej sfery, i potwierdzić wysuniętą przez siebie hipotezę, że analiza pojedynczych epizodów może nie doszacowywać czasu trwania ubóstwa.

Badanie ubóstwa na podstawie *British Household Panel Survey* (BHPS) z lat 1991–1997 z wykorzystaniem metod analizy przeżycia przeprowadził F. Devicienti<sup>6</sup>. Na podstawie uzyskanych wyników stwierdził on m.in., że są pewne grupy osób w większym stopniu zagrożonych nie tylko wejściem do sfery ubóstwa, ale również dłuższym pozostaniem w tej sferze.

---

<sup>3</sup> T. Panek, *Ubóstwo, wykluczenie społeczne i nierówności. Teoria i praktyka pomiaru*, Oficyna Wydawnicza SGH, Warszawa 2011.

<sup>4</sup> M.J. Bane, D.T. Ellwood, *Slipping into and out of poverty: the dynamics of spell*, „The Journal of Human Resources” 1986, vol. 21, no. 1, s. 1–23.

<sup>5</sup> A.H. Stevens, *Climbing out of poverty, falling back in: measuring the persistence of poverty over multiple spells*, „Journal of Human Resources” 1999, vol. 34, no. 3, s. 557–588.

<sup>6</sup> F. Devicienti, *Poverty persistence in Britain: a multivariate analysis using the BHPS, 1991–1997*, „Journal of Economics” 2002, vol. 9, no. 1, s. 307–340.

Na podstawie danych z niemieckiego panelu *German Socio-Economic Panel* (GSOEP) za lata 1984–2000 M. Biewen<sup>7</sup> stwierdził, że czas przeżycia w sferze ubóstwa w negatywny sposób wpływa na możliwość wejścia do sfery ubóstwa oraz na możliwość wyjścia z niej.

Analizę porównawczą determinant zmian przynależności do sfery ubóstwa w 14 państwach członkowskich Unii Europejskiej (w badaniu nie uwzględniono Polski) przeprowadzili E. Andriopoulou i P. Tsakloglou<sup>8</sup>. W analizie wykorzystali dane z badania *European Community Household Panel* za lata 1994–2000. Na podstawie uzyskanych wyników stwierdzili, że w większości badanych państw istnieje ujemna zależność pomiędzy czasem trwania w ubóstwie (poza ubóstwem) a prawdopodobieństwem wyjścia (ponownego wejścia) z biedy. Dłuższym czasem spędzonym w ubóstwie charakteryzowały się gospodarstwa domowe, których głowa była młodą lub była starszą osobą, oraz gospodarstwa domowe z dziećmi na utrzymaniu.

Przy użyciu danych panelowych z bazy CHER (*Consortium of Household Panels for European Socio-Economic Research*) za lata 1997–2000 N. Nehrebecka<sup>9</sup> analizowała czas trwania ubóstwa w Polsce. Wyniki badań potwierdziły negatywną zależność pomiędzy stopą wyjścia a trwaniem ubóstwa. Z przeprowadzonej analizy wynika również, że istnieją grupy populacji częściej zagrożone wejściem do sfery ubóstwa i z większym prawdopodobieństwem pozostające w ubóstwie przez dłuższy okres.

Metody analizy historii zdarzeń w badaniu ubóstwa w Polsce stosowała również A. Sączewska-Piotrowska<sup>10</sup>. Uzyskane wyniki potwierdzają to, że prawdopodobieństwo wejścia do sfery ubóstwa i wyjścia z niej zmniejsza się z biegiem czasu. Gospodarstwa domowe osób pracujących na własny rachunek trwają dłużej poza sferą ubóstwa i krócej w sferze ubóstwa niż pozostałe grupy społeczno-ekonomiczne. Przeprowadzona analiza ubóstwa z wykorzystaniem modeli

---

<sup>7</sup> M. Biewen, *Who are the chronic poor? Evidence on the extent and the composition of chronic poverty in Germany*, IZA Discussion Paper no. 779, 2003; M. Biewen, *Who are the chronic poor? An econometric analysis of chronic poverty in Germany*, „Research on Economic Inequality” 2006, vol. 13, no. 1, s. 31–62.

<sup>8</sup> E. Andriopoulou, P. Tsakloglou, *The determinants of poverty transitions in Europe and the role of duration dependence*, IZA Discussion Paper no. 5692, 2011.

<sup>9</sup> N. Nehrebecka, *Analiza ubóstwa w Polsce w latach 1997–2000 z wykorzystaniem modeli hazardu*, „Ekonomista” 2010, nr 1, s. 95–116.

<sup>10</sup> A. Sączewska-Piotrowska, *Analysis of poverty transitions in Poland using multilevel discrete-time event history models*, w: *Applications of Mathematics and Statistic in Economics*, red. Z. Rusnak, B. Zmyślona, Proceedings of the 17th AMSE, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2014, s. 219–228; A. Sączewska-Piotrowska, *Poverty duration of households of the self-employed*, „Ekonometria” 2015, nr 1(47), s. 44–55.

analizy historii zdarzeń dla zdarzeń powtarzających się pozwoliła stwierdzić, że cechy gospodarstwa domowego i jego głowy istotnie wpływają na szanse wyjścia (i wejścia) ze sfery ubóstwa (do sfery ubóstwa), przy czym wszystkie zmienne objaśniające w modelach miały charakter dychotomiczny (np. wiek głowy gospodarstwa do 59 lat oraz wiek 60 lat i więcej).

### 3. Dane i metoda

Zjawisko ubóstwa może dotyczyć gospodarstwa domowe wielokrotnie – mogą one bowiem wchodzić do sfery ubóstwa (wychodzić z niej) wiele razy, co oznacza, że zdarzenia mogą się powtarzać. Ze względu na małą średnią liczbę epizodów<sup>11</sup> przypadającą na gospodarstwo domowe analizę ograniczono do pierwszego epizodu, czyli czasu oczekiwania na pierwsze wystąpienie zdarzenia (pierwsze wejście do sfery ubóstwa i pierwsze wyjście z niej). W analizie uwzględniono tylko epizody rozpoczynające się w trakcie okresu obserwacji, co oznacza, że nie brano pod uwagę epizodów lewostronnie cenzurowanych.

Analizę czasu przeżycia rozpoczęto od oszacowania funkcji przeżycia za pomocą estymatora Kaplana–Meiera<sup>12</sup>. Dysponując próbą  $n$  gospodarstw domowych, oznaczmy uporządkowane czasy trwania epizodów jako  $t_1 < t_2 < \dots < t_l$ . Wtedy estymator Kaplana–Meiera jest określony wzorem:

$$\hat{S}(t) = \prod_{l: t_l \leq t} \left( 1 - \frac{d_l}{n_l} \right), \quad (1)$$

gdzie  $d_l$  jest liczbą zdarzeń występujących w czasie  $t_l$  oraz  $n_l$  jest liczbą jednostek narażonych na ryzyko wystąpienia zdarzenia bezpośrednio przed czasem  $t_l$  (uwzględniając cenzurowane czasy przeżycia w czasie  $t_l$ ).

Funkcję hazardu w przedziale  $t_l \leq t \leq t_{l+1}$  można szacować w następujący sposób<sup>13</sup>:

<sup>11</sup> P.D. Allison, *Survival analysis*, w: *The reviewer's guide to quantitative methods in the social sciences*, red. G.R. Hancock, R.O. Mueller, Routledge, New York 2010, s. 413–425.

<sup>12</sup> E.L. Kaplan, P. Meier, *Nonparametric estimation from incomplete observations*, „Journal of the American Statistical Association” 1958, vol. 53, s. 457–481.

<sup>13</sup> D. Collett, *Modelling survival data in medical research*, Chapman and Hall/CRC, Boca-Raton 2004, s. 30.

$$\hat{h}(t) = \frac{d_i}{n_i(t_{i+1} - t_i)}. \quad (2)$$

Jest to wskaźnik występowania zdarzenia na jednostkę czasu w przedziale  $[t_i, t_{i+1})$ .

Dla danych dotyczących czasu przeżycia opracowano wiele modeli regresji. Jednymi z najbardziej popularnych są modele Coxa – proporcjonalnego i nieproporcjonalnego hazardu. W przypadku pierwszego rodzaju modelu zakłada się, że wpływ zmiennych objaśniających jest stały w czasie, więc ilorazy hazardu nie zmieniają się w czasie. W przypadku modelu nieproporcjonalnego hazardu iloraz hazardu zmienia się wraz z upływem czasu. W niniejszym opracowaniu zastosowano model proporcjonalnego hazardu.

Model wprowadzony przez D. Coxa modeluje wskaźnik hazardu w następujący sposób<sup>14</sup>:

$$\lambda_i(t) = \lambda_0(t) \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}), \quad (3)$$

gdzie hazard dla jednostki  $i$  w czasie  $t$  jest iloczynem dwóch czynników: niewyspecyfikowanej parametrycznie funkcji hazardu bazowego  $\lambda_0(t)$ , zapisywanej również jako  $h_0(t)$ , oraz wyspecyfikowanej funkcji wykładniczej zmiennych objaśniających, przy czym  $x_i$  jest wektorem zmiennych niezależnych od czasu. Funkcja hazardu bazowego jest interpretowana jako funkcja hazardu bazowego jednostki, dla której wszystkie zmienne przyjmują wartość 0.

Podstawową metodą estymacji modelu Coxa, który jest modelem semiparametrycznym, jest metoda częściowej wiarygodności (ang. *partial likelihood method*), zaproponowana przez Coxa<sup>15</sup>.

Po oszacowaniu parametrów  $\beta_j$  ( $j = 1, 2, \dots, k$ ) dla każdej zmiennej jest obliczany iloraz hazardu (ang. *hazard ratio*) określony wzorem:

$$HR_j = \exp(\beta_j). \quad (4)$$

Iloraz hazardu wyraża zmianę ryzyka zdarzenia, gdy dana zmienna objaśniająca rośnie o jedną jednostkę przy niezmiennych wartościach pozostałych zmiennych objaśniających.

<sup>14</sup> E. Frątczak, U. Gach-Ciepiela, H. Babiker, *Analiza historii zdarzeń*, Oficyna Wydawnicza SGH, Warszawa 2005, s. 111; M. Mills, op.cit., s. 87; B. Bieszk-Stolorz, I. Markowicz, *Modele regresji Coxa w analizie bezrobocia*, CeDeWu, Warszawa 2012, s. 31.

<sup>15</sup> D.R. Cox, *Regression models and life tables*, „Journal of the Royal Statistical Society” 1972, vol. 34, s. 187–220; D.R. Cox, *Partial likelihood*, „Biometrika” 1975, vol. 62, s. 269–276.

Istotność statystyczna wszystkich zmiennych objaśniających w modelu łącznie jest oceniana za pomocą testu ilorazu wiarygodności, natomiast istotność poszczególnych parametrów opiera się na statystyce Walda<sup>16</sup>.

W modelu proporcjonalnego hazardu Coxa jest przyjmowane założenie o niezależności zmiennych w czasie. W opracowaniu zostanie wykorzystany test bazujący na resztach Schoenfelda, szczegółowo opisany w książce D.W. Hosmera, S. Lemeshowa i S. Maya<sup>17</sup>.

W modelach o czasie dyskretnym zmienna zależna modeluje ryzyko, czyli prawdopodobieństwo tego, że zdarzenie wystąpi pod warunkiem przeżycia do tego czasu i w zależności od zmiennych objaśniających. Będzie stosowany zapis  $P(y_{ii} = 1) = \lambda_i$ , oznaczający prawdopodobieństwo wystąpienia zdarzenia, natomiast zapis  $P(y_{ii} = 0) = 1 - \lambda_i$  oznacza prawdopodobieństwo niewystąpienia zdarzenia. Równocześnie zakłada się, że prawdopodobieństwo jest funkcją zmiennych objaśniających.

Jednym z bardziej popularnych modeli o czasie dyskretnym jest model logitowy postaci<sup>18</sup>:

$$\text{logit} \lambda_i(t) = \log \frac{\lambda_i(t)}{1 - \lambda_i(t)} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}, \quad (5)$$

gdzie  $\text{logit} \lambda_i$  jest logarytmem ilorazu prawdopodobieństwa wystąpienia zdarzenia i prawdopodobieństwa niewystąpienia zdarzenia. Jest to logarytm szansy (ang. *log-odds*), nazywany popularnie logarytmem ilorazu szans. Dla oszacowanych parametrów  $\beta_j$  ( $j = 1, 2, \dots, k$ ) oblicza się iloraz szans (ang. *odds ratio*) określony wzorem:

$$OR_j = \exp(\beta_j). \quad (6)$$

Wyraża on zmianę szans na wystąpienie zdarzenia, gdy dana zmienna objaśniająca rośnie o jedną jednostkę przy niezmiennych wartościach pozostałych zmiennych objaśniających.

<sup>16</sup> Test ilorazu wiarygodności opisano m.in. w: B. Więckowska, *Podręcznik użytkownika – PQStat*, PQStat Software, 2015, s. 365, natomiast test Walda w: B. Bieszk-Stolorz, I. Markowicz, op.cit., s. 34.

<sup>17</sup> D.W. Hosmer, S. Lemeshow, S. May, *Applied survival analysis. Regression modeling of time-to-event data*, John Wiley & Sons Inc., Hoboken 2008, s. 170–184.

<sup>18</sup> M. Mills, op.cit., s. 182.

Model komplementarny log-log (cloglog) jest rzadziej stosowanym modelem o czasie dyskretnym. Przyjmuje postać:

$$\log[-\log(1 - \lambda_i)] = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}. \quad (7)$$

Model komplementarny log-log jest przybliżeniem modelu proporcjonalnego hazardu Coxa, stąd współczynniki tych modeli są bezpośrednio porównywalne. Oszacowania modelu Coxa i modelu logitowego zbliżają się do siebie wraz ze zmniejszeniem funkcji hazardu. Hazard o czasie dyskretnym jest mniejszy, gdy szerokości przedziałów czasowych stają się coraz mniejsze<sup>19</sup>.

Oprócz modeli logitowego i komplementarnego log-log w analizach jest często stosowany model probitowy, którego zastosowanie w analizie historii zdarzeń opisano m.in. w pracy M. Mills<sup>20</sup>.

Parametry modelu logitowego oraz modelu komplementarnego log-log można estymować metodą największej wiarygodności. W przypadku obydwu modeli istotność statystyczną zmiennych objaśniających łącznie można oceniać za pomocą testu ilorazu wiarygodności, natomiast istotność poszczególnych parametrów za pomocą testu Walda.

W badaniu wykorzystano dane pochodzące z ośmiu fal panelu zrealizowanego w latach 2000–2015 w ramach projektu „Diagnoza społeczna”<sup>21</sup>. Jako wskaźnik zamożności przyjęto dochody netto gospodarstw domowych w Polsce w lutym, marcu i czerwcu w latach: 2000, 2003, 2005, 2007, 2009, 2011, 2013 i 2015. W celu uwzględnienia różnic występujących w wielkości i składzie demograficznym gospodarstw domowych obliczono dochody ekwiwalentne, stosując w tym celu zmodyfikowaną skalę OECD. Skala ta przypisuje pierwszej dorosłej osobie w gospodarstwie wartość 1, każdej następnej – wartość 0,5, natomiast dziecku (każda osoba poniżej 14 lat) – wartość 0,3. Gospodarstwo domowe zostało uznane za ubogie, gdy jego dochód był mniejszy niż 60% mediany rozkładów dochodów ekwiwalentnych w danym roku.

Przeprowadzona analiza dynamiki ubóstwa z wykorzystaniem metod analizy przeżycia bazowała na niezbilansowanym panelu liczącym 26 841 gospodarstw domowych. Przeprowadzenie badania na podstawie panelu niezbilansowanego

<sup>19</sup> F. Steele, E. Washbrook, *Discrete-time event history analysis. Lectures*, Centre for Multilevel Modelling, University of Bristol, 2013, <http://www.bristol.ac.uk/media-library/sites/cmm/migrated/documents/discrete-time-eha-july2013-combined.pdf> (odczyt: 15.06.2015).

<sup>20</sup> M. Mills, op.cit., s. 183–189.

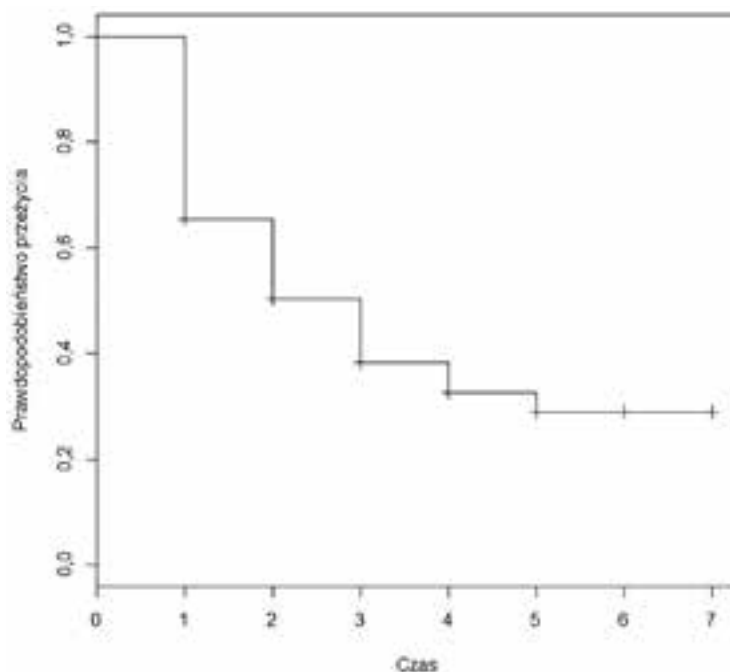
<sup>21</sup> *Diagnoza społeczna: zintegrowana baza danych*, Rada Monitoringu Społecznego, 2015, <http://www.diagnoza.com> (odczyt: 28.12.2015).



oznacza, że brane są pod uwagę gospodarstwa domowe uczestniczące w różnej liczbie fal panelu, np. część gospodarstw mogła uczestniczyć tylko w dwóch falach panelu, część w trzech falach panelu itd. W panelu 2169 gospodarstw domowych oczekiwało na pierwsze wejście do sfery ubóstwa (znany początek okresu oczekiwania na pierwsze wejście), 2015 – na pierwsze wyjście z niej (znany początek okresu oczekiwania na pierwsze wyjście).

#### 4. Analiza przeżycia w sferze ubóstwa i poza nią z wykorzystaniem estymatora Kaplana–Meiera

Analizę dynamiki ubóstwa rozpoczęto od oszacowania funkcji hazardu i przeżycia w sferze ubóstwa dla ogółu gospodarstw domowych (rysunek 1 i tabela 1).



**Rysunek 1. Funkcja przeżycia Kaplan–Meiera w sferze ubóstwa**

Źródło: opracowanie własne.

Prawdopodobieństwo wyjścia ze sfery ubóstwa pod warunkiem, że zdarzenie to wcześniej nie wystąpiło, ogólnie spada wraz z czasem spędzonym

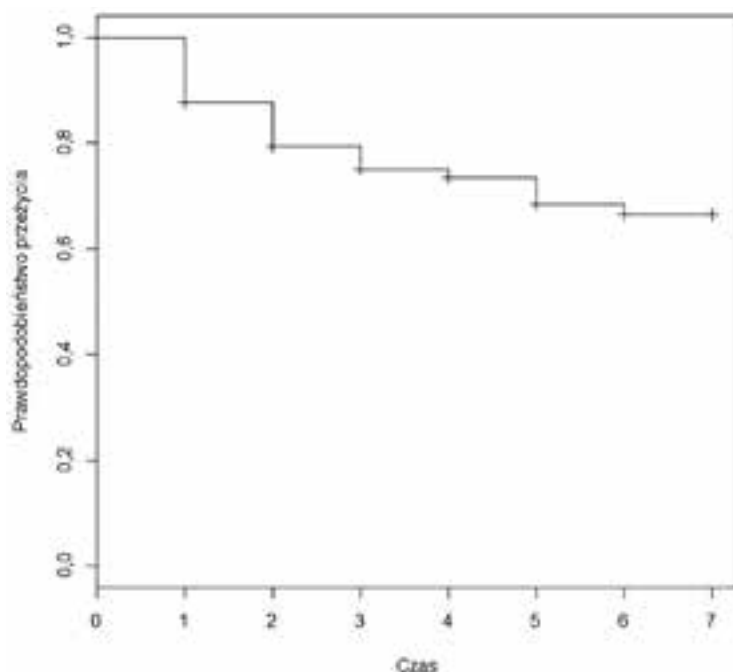
w sferze ubóstwa. Wyjątek stanowiło prawdopodobieństwo wyjścia ze sfery ubóstwa po trzech okresach, czyli po 6 latach (przerwy pomiędzy kolejnymi falami panelu wynosiły 2 lata), było ono nieco wyższe niż po dwóch okresach. 2 lata i więcej w sferze ubóstwa pozostaje ponad 65% gospodarstw domowych, natomiast 10 lat i więcej – blisko 30% gospodarstw domowych.

**Tabela 1. Oszacowane funkcje hazardu i przeżycia w sferze ubóstwa**

| Czas              | 1     | 2     | 3     | 4     | 5     | 6     | 7     |
|-------------------|-------|-------|-------|-------|-------|-------|-------|
| Funkcja przeżycia | 0,653 | 0,503 | 0,384 | 0,326 | 0,290 | 0,290 | 0,290 |
| Funkcja hazardu   | 0,347 | 0,230 | 0,237 | 0,151 | 0,110 | 0,000 | 0,000 |

Źródło: opracowanie własne.

Funkcje hazardu i przeżycia poza sferą ubóstwa przedstawiono na rysunku 2 i w tabeli 2.



**Rysunek 2. Funkcja przeżycia Kaplan-Meiera poza sferą ubóstwa**

Źródło: opracowanie własne.

Można zauważyć, że prawdopodobieństwo wejścia do sfery ubóstwa pod warunkiem, że zdarzenie to wcześniej nie wystąpiło, generalnie spada wraz z czasem spędzonym poza sferą ubóstwa. Wyjątkiem jest wzrost prawdopodobieństwa po pięciu okresach spędzonych poza sferą ubóstwa. Gospodarstwa domowe przebywają poza sferą ubóstwa długi czas: 2 lata i więcej poza sferą ubóstwa przeżywa prawie 88% gospodarstw, natomiast 14 lat i więcej – 66,5% gospodarstw.

**Tabela 2. Oszacowane funkcje hazardu i przeżycia poza sferą ubóstwa**

| Czas              | 1     | 2     | 3     | 4     | 5     | 6     | 7     |
|-------------------|-------|-------|-------|-------|-------|-------|-------|
| Funkcja przeżycia | 0,877 | 0,794 | 0,751 | 0,734 | 0,685 | 0,665 | 0,665 |
| Funkcja hazardu   | 0,123 | 0,095 | 0,054 | 0,023 | 0,067 | 0,029 | 0,000 |

Źródło: opracowanie własne.

## 5. Determinanty wejścia do sfery ubóstwa

Analizując determinanty wejść do sfery ubóstwa, w szacowanych modelach wykorzystano zmienne objaśniające odnoszące się do cech gospodarstwa domowego i jego głowy. Wszystkie zmienne uwzględnione w modelach zostały przedstawione w postaci układów zmiennych zero-jedynkowych. Szacując modele, w celu uniknięcia współliniowości w przypadku każdej ze zmiennych pominięto jedną kategorię (stała się ona punktem odniesienia). W szacowanych modelach uwzględniono następujące zmienne:

- $P$  – płeć głowy gospodarstwa domowego (1 – mężczyzna, 0 – kobieta),
- $W_i$  – przedział wieku głowy gospodarstwa domowego, gdzie:  $W_1$  – 34 lata i mniej (punkt odniesienia),  $W_2$  – 35–44 lata,  $W_3$  – 45–59 lat,  $W_4$  – 60 lat i więcej,
- $E_i$  – poziom wykształcenia głowy gospodarstwa domowego, gdzie:  $E_1$  – wyższe (punkt odniesienia),  $E_2$  – średnie,  $E_3$  – zawodowe,  $E_4$  – gimnazjalne i niższe,
- $K$  – klasa miejscowości zamieszkania (1 – wieś, 0 – miasto),
- $G_i$  – grupa społeczno-ekonomiczna, gdzie:  $G_1$  – pracownicy (punkt odniesienia),  $G_2$  – rolnicy,  $G_3$  – pracujący na własny rachunek,  $G_4$  – emeryci i renciści,  $G_5$  – utrzymujący się z niezarobkowych źródeł,
- $B$  – status gospodarstwa na rynku pracy (1 – gospodarstwa domowe z przynajmniej jedną osobą bezrobotną, 0 – gospodarstwa domowe bez osób bezrobotnych).

W przeprowadzonej analizie jako punkty odniesienia przyjęto: kobiety, wiek głowy gospodarstwa domowego 34 lata i mniej, wyższe wykształcenie głowy gospodarstwa, miasto, grupę pracowników oraz gospodarstwa domowe bez osób bezrobotnych.

Ważną kwestią przy szacowaniu parametrów modeli analizy historii zdarzeń jest struktura danych. Dane mogą być zapisane w „szerokim formacie” (typowe zbiory danych mają taką postać), w przypadku którego na każdą jednostkę przeznaczony jest jeden wiersz zawierający informacje o numerze jednostki, czasie oczekiwania na wystąpienie zdarzenia oraz cenzurowaniu. W przypadku „długiego formatu” informacja o danej jednostce zajmuje tyle wierszy, ile jest badanych okresów. Do oszacowania modelu proporcjonalnego hazardu Coxa wykorzystano „szeroki format”, natomiast do oszacowania modeli o czasie dyskretnym – „długi format” danych. Wyniki estymacji modeli przedstawiono w tabeli 3.

W każdym z oszacowanych modeli łącznie wszystkie zmienne były statystycznie istotne (na poziomie istotności 0,05). We wszystkich modelach te same zmienne okazały się statystycznie istotne, lecz w niektórych przypadkach na różnym poziomie istotności. Odnosząc się do modelu logitowego, można powiedzieć, że szanse wejścia do sfery ubóstwa są:

- o ok. 54% niższe w przypadku gospodarstw domowych, których głowa jest w wieku 60 lat i więcej, w porównaniu z gospodarstwami, których głowa jest w wieku 34 lat i mniej,
- ok. 8,4 razy, 9,5 razy oraz 11 razy wyższe dla gospodarstw, których głowa ma wykształcenie odpowiednio średnie, zawodowe oraz co najwyżej gimnazjalne, w porównaniu z gospodarstwami, których głowa ma wykształcenie wyższe,
- ok. 1,65 razy, 1,69 razy i 1,89 razy wyższe odpowiednio dla gospodarstw rolników, emerytów i rencistów oraz utrzymujących się z niezarobkowych źródeł w porównaniu z gospodarstwami pracowników,
- ok. 1,41 razy wyższe dla gospodarstw domowych z przynajmniej jedną osobą bezrobotną w porównaniu z gospodarstwami bez osób bezrobotnych.

Dla modelu Coxa i modelu komplementarnego log-log interpretacja wyników odnosi się do ryzyka, a nie do szansy. Analizując wyniki oszacowań parametrów modelu Coxa, można przykładowo stwierdzić, że ryzyko wejścia do sfery ubóstwa gospodarstw domowych z głową mającą 60 lat i więcej jest o 69% niższe niż ryzyko wejścia gospodarstw z głową w wieku 34 lat i mniej.

**Tabela 3. Determinanty wejścia do sfery ubóstwa – wyniki estymacji modeli**

| Zmienna               | Cox                 |                         | Logit               |                         | Cloglog             |                         |
|-----------------------|---------------------|-------------------------|---------------------|-------------------------|---------------------|-------------------------|
|                       | <i>b</i> (SE)       | <i>exp</i> ( <i>b</i> ) | <i>b</i> (SE)       | <i>exp</i> ( <i>b</i> ) | <i>b</i> (SE)       | <i>exp</i> ( <i>b</i> ) |
| Stała                 |                     |                         | -4,5943<br>(0,7344) | 0,0101***               | -4,5809<br>(0,7276) | 0,0102***               |
| <i>P</i>              | 0,0667<br>(0,1209)  | 1,0689                  | 0,0521<br>(0,1286)  | 1,0535                  | 0,0485<br>(0,1210)  | 1,0497                  |
| <i>W</i> <sub>2</sub> | 0,0905<br>(0,2282)  | 1,0947                  | 0,2754<br>(0,2442)  | 1,3171                  | 0,2460<br>(0,2274)  | 1,2789                  |
| <i>W</i> <sub>3</sub> | -0,2823<br>(0,2162) | 0,7541                  | -0,1115<br>(0,2289) | 0,8945                  | -0,1110<br>(0,2140) | 0,8949                  |
| <i>W</i> <sub>4</sub> | -1,1704<br>(0,2618) | 0,3103***               | -0,7793<br>(0,2732) | 0,4587**                | -0,7377<br>(0,2561) | 0,4782**                |
| <i>E</i> <sub>2</sub> | 2,1572<br>(0,7183)  | 8,6472**                | 2,1296<br>(0,7235)  | 8,4112**                | 2,0869<br>(0,7183)  | 8,0597**                |
| <i>E</i> <sub>3</sub> | 2,2657<br>(0,7151)  | 9,6379**                | 2,2541<br>(0,7195)  | 9,5262**                | 2,2089<br>(0,7156)  | 9,1056**                |
| <i>E</i> <sub>4</sub> | 2,4591<br>(0,7191)  | 11,6937***              | 2,3996<br>(0,7242)  | 11,0183***              | 2,3357<br>(0,7187)  | 10,3368**               |
| <i>K</i>              | 0,0643<br>(0,1159)  | 1,0664                  | 0,0979<br>(0,1227)  | 1,1029                  | 0,0898<br>(0,1155)  | 1,0940                  |
| <i>G</i> <sub>2</sub> | 0,5178<br>(0,1621)  | 1,6783**                | 0,4976<br>(0,1746)  | 1,6448**                | 0,4575<br>(0,1621)  | 1,5802**                |
| <i>G</i> <sub>3</sub> | -0,0307<br>(0,3001) | 0,9698                  | -0,1481<br>(0,3159) | 0,8623                  | -0,1371<br>(0,3000) | 0,8719                  |
| <i>G</i> <sub>4</sub> | 0,6476<br>(0,1662)  | 1,9110***               | 0,5272<br>(0,1737)  | 1,6941**                | 0,4944<br>(0,1624)  | 1,6395**                |
| <i>G</i> <sub>5</sub> | 0,5586<br>(0,2272)  | 1,7483*                 | 0,6391<br>(0,2482)  | 1,8948*                 | 0,5942<br>(0,2256)  | 1,8116**                |
| <i>B</i>              | 0,3491<br>(0,1248)  | 1,4178**                | 0,3447<br>(0,1342)  | 1,4115*                 | 0,3202<br>(0,1248)  | 1,3773*                 |

poziom istotności: \*\*\*0,001, \*\*0,01, \*0,05, .0,1

Źródło: opracowanie własne.

## 6. Determinanty wyjścia ze sfery ubóstwa

W celu określenia determinant wyjścia ze sfery ubóstwa oszacowano parametry modelu Coxa oraz modeli o czasie dyskretnym (tabela 4).

**Tabela 4. Determinanty wyjścia ze sfery ubóstwa – wyniki estymacji modeli**

| Zmienna               | Cox                 |                         | Logit               |                         | Cloglog             |                         |
|-----------------------|---------------------|-------------------------|---------------------|-------------------------|---------------------|-------------------------|
|                       | <i>b</i> (SE)       | <i>exp</i> ( <i>b</i> ) | <i>b</i> (SE)       | <i>exp</i> ( <i>b</i> ) | <i>b</i> (SE)       | <i>exp</i> ( <i>b</i> ) |
| Stała                 |                     |                         | -0,1101<br>(0,2821) | 0,8957                  | -0,4060<br>(0,2221) | 0,6663.                 |
| <i>P</i>              | -0,0770<br>(0,0782) | 0,9259                  | -0,0934<br>(0,0960) | 0,9109                  | -0,0824<br>(0,0790) | 0,9209                  |
| <i>W</i> <sub>2</sub> | -0,2163<br>(0,1557) | 0,8055                  | -0,3398<br>(0,1963) | 0,7119.                 | -0,2820<br>(0,1569) | 0,7543.                 |
| <i>W</i> <sub>3</sub> | -0,1577<br>(0,1432) | 0,8541                  | -0,1544<br>(0,1823) | 0,8569                  | -0,1292<br>(0,1444) | 0,8788                  |
| <i>W</i> <sub>4</sub> | -0,3477<br>(0,1712) | 0,7063*                 | -0,1757<br>(0,2127) | 0,8389                  | -0,1463<br>(0,1712) | 0,8639                  |
| <i>E</i> <sub>2</sub> | -0,1122<br>(0,1941) | 0,8938                  | -0,1518<br>(0,2452) | 0,8592                  | -0,1368<br>(0,1949) | 0,8722                  |
| <i>E</i> <sub>3</sub> | -0,0047<br>(0,1882) | 0,9953                  | -0,0090<br>(0,2387) | 0,9910                  | -0,0238<br>(0,1894) | 0,9765                  |
| <i>E</i> <sub>4</sub> | -0,0531<br>(0,1964) | 0,9482                  | -0,1579<br>(0,2472) | 0,8539                  | -0,1412<br>(0,1973) | 0,8683                  |
| <i>K</i>              | -0,2144<br>(0,0760) | 0,8070**                | -0,2705<br>(0,0929) | 0,7630**                | -0,2229<br>(0,0766) | 0,8002**                |
| <i>G</i> <sub>2</sub> | 0,2088<br>(0,1176)  | 1,2323.                 | 0,1870<br>(0,1484)  | 1,2056                  | 0,1543<br>(0,1192)  | 1,1668                  |
| <i>G</i> <sub>3</sub> | -0,1921<br>(0,2328) | 0,8252                  | -0,2125<br>(0,2832) | 0,8086                  | -0,1755<br>(0,2343) | 0,8391                  |
| <i>G</i> <sub>4</sub> | -0,1399<br>(0,1093) | 0,8695                  | -0,3863<br>(0,1322) | 0,6795**                | -0,3204<br>(0,1093) | 0,7259**                |
| <i>G</i> <sub>5</sub> | -0,4858<br>(0,1296) | 0,6152***               | -0,6148<br>(0,1538) | 0,5408***               | -0,5122<br>(0,1303) | 0,5992***               |
| <i>B</i>              | 0,0785<br>(0,0767)  | 1,0817                  | 0,0878<br>(0,0949)  | 1,0917                  | 0,0667<br>(0,0776)  | 1,0690                  |

poziom istotności: \*\*\*0,001, \*\*0,01, \*0,05, .0,1

Źródło: opracowanie własne.

Na podstawie testu ilorazu wiarygodności można stwierdzić, że w każdym oszacowanym modelu łącznie wszystkie zmienne były statystycznie istotne (na poziomie 0,05). Wyniki oszacowań parametrów w modelu Coxa są odmienne od wyników oszacowań modeli o czasie dyskretnym – inne zmienne są statystycznie istotne. Z jednej strony statystycznie istotne w modelu Coxa są zmienne *W*<sub>4</sub> i *G*<sub>2</sub>, które w pozostałych modelach okazały się nieistotne, natomiast z drugiej strony zmienna *G*<sub>4</sub> jest istotna w modelu logitowym i komplementarnym log-log, a wyniki oszacowań modelu Coxa wskazują na jej nieistotność statystyczną.

Na podstawie modelu logitowego można stwierdzić, że szanse wyjścia ze sfery ubóstwa są:

- o niecałe 24% niższe w przypadku gospodarstw domowych zamieszkujących na wsi w porównaniu z gospodarstwami zamieszkującymi w miastach,
- o ok. 32% i ok. 46% niższe odpowiednio dla gospodarstw domowych emerytów i rencistów oraz gospodarstw utrzymujących się z niezarobkowych źródeł w porównaniu z gospodarstwami pracowników.

Odnosząc się do modelu Coxa, można stwierdzić, że ryzyko wyjścia ze sfery ubóstwa gospodarstwa domowego z głową mającą 60 lat i więcej jest o 30% niższe niż ryzyko wyjścia gospodarstwa z głową w wieku 34 lat i mniej.

## 7. Podsumowanie i kierunki dalszych badań

Wyniki przeprowadzonej analizy wskazują na fakt, że prawdopodobieństwo pobytu gospodarstwa domowego w sferze ubóstwa przez długi czas jest znacznie mniejsze niż prawdopodobieństwo długookresowego pobytu poza nią. Ryzyko wyjścia ze sfery ubóstwa i ryzyko wejścia do niej są najwyższe na początku okresu pobytu, a następnie stopniowo maleją. Wyniki przeprowadzonej analizy potwierdzają więc wyniki wcześniejszych badań świadczące o negatywnym wpływie czasu na możliwość wejścia do sfery ubóstwa oraz wyjścia z niej.

Zgodnie z oczekiwaniami szanse wejścia do sfery ubóstwa i wyjścia z niej różnią się statystycznie w zależności od wybranych cech gospodarstwa domowego i jego głowy. Można zauważyć, że na szanse te istotnie wpływają zmienne odnoszące się do grupy społeczno-ekonomicznej gospodarstwa domowego. Szczególną uwagę należy zwrócić na gospodarstwa utrzymujące się z niezarobkowych źródeł, które z jednej strony z trudem opuszczają sferę ubóstwa, a z drugiej – z łatwością do niej wchodzi. Te gospodarstwa domowe mają więc bardzo małe szanse na poprawę swojej sytuacji ekonomicznej, w wielu przypadkach ich członkowie żyją z zasiłków, na marginesie społeczeństwa, blisko degradacji biologicznej. Kolejne pokolenia, nie widząc innej możliwości, powielają wzorzec wyniesiony z domu rodzinnego. Dotarcie z pomocą do tych gospodarstw domowych jest zatem niezwykle trudne, ponieważ wiąże się przede wszystkim ze zmianą wzorca zachowania, niejednokrotnie przekazywanego z pokolenia na pokolenie.

Przeprowadzoną analizę determinant można poszerzyć, uwzględniając w modelach dodatkowe cechy gospodarstw domowych i ich członków, mogące

istotnie wpływać na ryzyko wejścia do sfery ubóstwa i ryzyko wyjścia z niej, np. narodziny członka gospodarstwa domowego.

## Bibliografia

- Allison P.D., *Survival analysis*, w: *The reviewer's guide to quantitative methods in the social sciences*, red. G.R. Hancock, R.O. Mueller, Routledge, New York 2010, s. 413–425.
- Andriopoulou E., Tsakoglou P., *The determinants of poverty transitions in Europe and the role of duration dependence*, IZA Discussion Paper no. 5692, 2011.
- Bane M.J., Ellwood D.T., *Slipping into and out of poverty: the dynamics of spell*, „The Journal of Human Resources” 1986, vol. 21, no. 1, s. 1–23.
- Bieszk-Stolorz B., Markowicz I., *Modele regresji Coxa w analizie bezrobocia*, CeDeWu, Warszawa 2012.
- Biewen M., *Who are the chronic poor? An econometric analysis of chronic poverty in Germany*, „Research on Economic Inequality” 2006, vol. 13, no. 1, s. 31–62.
- Biewen M., *Who are the chronic poor? Evidence on the extent and the composition of chronic poverty in Germany*, IZA Discussion Paper no. 779, 2003.
- Collett D., *Modelling survival data in medical research*, Chapman and Hall/CRC, Boca-Raton 2004.
- Cox D.R., *Partial likelihood*, „Biometrika” 1975, vol. 62, s. 269–276.
- Cox D.R., *Regression models and life tables*, „Journal of the Royal Statistical Society” 1972, vol. 34, s. 187–220.
- Devicienti F., *Poverty persistence in Britain: a multivariate analysis using the BHPS, 1991–1997*, „Journal of Economics” 2002, vol. 9, no. 1, s. 307–340.
- Frątczak E., Gach-Ciepiela U., Babiker H., *Analiza historii zdarzeń*, Oficyna Wydawnicza SGH, Warszawa 2005.
- Hosmer D.W., Lemeshow S., May S., *Applied survival analysis. Regression modeling of time-to-event data*, John Wiley & Sons Inc., Hoboken 2008.
- Kaplan E.L., Meier P., *Nonparametric estimation from incomplete observations*, „Journal of the American Statistical Association” 1958, vol. 53, s. 457–481.
- Mills M., *Survival and event history analysis*, Sage Publications, Los Angeles–London–New Dehli–Singapore–Washington 2011.
- Nehrebecka N., *Analiza ubóstwa w Polsce w latach 1997–2000 z wykorzystaniem modeli hazardu*, „Ekonomista” 2010, nr 1, s. 95–116.
- Panek T., *Ubóstwo, wykluczenie społeczne i nierówności. Teoria i praktyka pomiaru*, Oficyna Wydawnicza SGH, Warszawa 2011.



- Sączewska-Piotrowska A., *Analysis of poverty transitions in Poland using multilevel discrete-time event history models*, w: *Applications of Mathematics and Statistics in Economics*, red. Z. Rusnak, B. Zmyślona, Proceedings of the 17th AMSE, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2014, s. 219–228.
- Sączewska-Piotrowska A., *Poverty duration of households of the self-employed*, „*Ekonomometria*” 2015, nr 1(47), s. 44–55.
- Stevens A.H., *Climbing out of poverty, falling back in: measuring the persistence of poverty over multiple spells*, „*Journal of Human Resources*” 1999, vol. 34, no. 3, s. 557–588.
- Więckowska B., *Podręcznik użytkownika – PQStat*, PQStat Software, 2015.

## Źródła sieciowe

- Diagnoza społeczna: zintegrowana baza danych*, Rada Monitoringu Społecznego, 2015, <http://www.diagnoza.com> (odczyt: 28.12.2015).
- Steele F., Washbrook E., *Discrete-time event history analysis. Lectures*, Centre for Multilevel Modelling, University of Bristol, 2013, <http://www.bristol.ac.uk/media-library/sites/cmm/migrated/documents/discrete-time-eha-july2013-combined.pdf> (odczyt: 15.06.2015).

\* \* \*

## Study of poverty dynamics of households using selected event history analysis models

### Summary

The aim of this paper is to identify factors increasing and decreasing chances to poverty entry and to poverty exit. There are used event history models: semiparametric Cox model and discrete-time models (logit and complementary log-log). In the models there are included covariates: gender, age and education of household's head, class of locality, socio-economic group and labour force status of the household. The analysis is extended by descriptive analysis of survival time in poverty and out of poverty for all households using Kaplan – Meier estimator. The probability of survival for a long time in poverty is less than in the case of survival out of poverty. Households living on unearned sources are in the worst economic situation, because they have small odds to exit from poverty and high odds to enter to poverty.

**Keywords:** poverty dynamics, Cox model, logit model, complementary log-log model, Kaplan–Meier estimator

