Anna Sączewska-Piotrowska

University of Economics in Katowice
Faculty of Economics
Department of Labour Market Forecasting
and Analysis
anna.saczewska-piotrowska@ue.katowice.pl

Jaroslaw Wąsowicz

University of Economics in Katowice
Faculty of Economics
Department of Labour Market Forecasting
and Analysis
jaroslaw.wasowicz@ue.katowice.pl

PROFILES OF RICHNESS IN POLAND

Summary: The aim of this paper was to analyze the dynamics of income richness in Poland. There were used simple approach that counts number of periods above the richness line and approach examining both the length and number of richness periods. The results show that the highest proportions of never rich households were in groups of households with low-educated head, single households and households living on unearned sources. Based on estimated model it can be stated that the odds of being in higher richness categories versus in lower categories differed in statistically significant way depending on education of household's head and socio-economic group of household.

Keywords: income richness, richness persistence, recurrence of richness, ordinal logistic regression.

JEL Classification: D31, D33.

Introduction

The richness is analyzed very rarely. Rich people are not a problem, therefore analyses pay attention to poverty and methods of combating this negative phenomenon. It must be noted that rich people should be also interest of analyses. They can consume and save larger amounts of money than the others in society, therefore they should be interest of banks, companies providing different kinds of services and companies producing different kinds of goods. The answer to the question which households are rich and how long the richness lasts is very important to these banks, companies, etc. The aim of this paper was to analyze the dynamics of income richness in Poland. There were applied methods used in poverty dynamics study. There were used simple approach that counts number of

periods above the richness line and approach that creates richness profiles examining both the length and number of richness periods. There were formed several research hypotheses for the purpose of this paper. There was verified the hypothesis claiming that in 2007-2015 most of households were never rich. The second hypothesis assumes that the education of the household's head and socio-economic group of household differentiates the number of periods spent in richness. It can be predicted that proportion of never rich households is the highest in group with low-educated household's head and in group of households living on unearned sources. The third hypothesis claims that the odds of being in higher richness categories versus in lower categories differ in statistically significant way depending again on education of the household's head and socio-economic group of household.

1. Data and methods

The analysis of richness dynamics was conducted for Poland based on the balanced panel 2007-2015 (five waves: 2007, 2009, 2011, 2013, 2015) in the framework of "Social Diagnosis" project [Council for Social Monitoring, 2015]. The panel contains almost 1100 households.

In the ordinary sense to be 'rich' means to have a lot of money. In the scientific sense the meaning of the term 'rich' must be precisely defined. Poverty is a very often analyzed phenomenon all over the world. This negative phenomenon may be defined in different ways – focusing only on material needs or also on the social aspects. The opposite of poverty is richness and this phenomenon may be also interpreted in different ways. In this paper the richness is interpreted very narrowly and is considered only through the prism of income. Therefore, there is used a term 'income richness'. The analyze of income richness was based on the relative richness line that would be calculated as a double median equivalized income in each of studied year. There are also used in literature another relative richness lines – 300% and 400% of the median equivalized income [Brzeziński, 2010]. The another way of defining the richness line bases on income of the richest p% of households (empirically mostly common set as 5% or 1%, for example Atkinson and Piketty [eds., 2007], Leigh [2009]). In Polish empirical study there were used richness lines based on absolute amount of income per capita or per household. For example, Teresa Słaby [ed., 2006] used the term 'the economic elite' to describe the group high-income households with monthly income 5000 PLN per capita. In study conducted by KPMG [KPMG w Polsce,

2014] there were assumed that 'people rich and wealth' have monthly income higher than 7100 PLN (before tax). Besides Słaby and KPMG, in Polish literature the richness was discussed, i.a., by Brzeziński [2010] and Sączewska-Piotrowska [2015].

We assume that the simplest method of analysis of richness dynamics is richness hit rate (in analogy to poverty hit rate), which counts the number of times when households are rich within the observation period. This approach is so called 'x out of n' approach (the proportion of households with income above the richness line in x out of n time periods), where $0 \leq x \leq n$. This measure carries some information about duration of the richness. The prevalence of persistent versus transitory richness is then evaluated by comparing the proportion of households that were rich at least half of periods (the persistently rich) with the proportion of households rich less than half of periods (the transitorily rich). In this paper attention is focused on the poverty profiles defined by Muffels, Fouarge and Dekker [2000] and continued by Fouarge and Layte [2003], Layte and Fouarge [2004], Fouarge and Layte [2005], Muffels and Fouarge [2010]. In the paper this approach is adapted to analysis of richness and there are defined richness profiles. It can be stated that the time dependent nature of richness, as in the case of poverty, is characterized by four dimensions: the length of the observation period, the length of the richness spell, the extent of recurrent spells and the volatility and stability of richness statuses over time. These four dimensions of longitudinal richness together determine the pattern or profile of richness for each household over time. Four types of income richness profiles are considered:
− the persistent non-rich (never rich in the observation period);
− the transient rich (once rich in the observation period);
− the recurrent rich (more than once rich but never longer than $x - 1$ consecutive years);
− the persistent rich (rich for a consecutive period of at least x consecutive years).

The value for x (the number of periods) in the definition of recurrent and persistent poverty has to be chosen by the researcher or the politician and depends on the length of the observation period and the distribution of poverty risks over time in society [Muffels, Fouarge, 2010]. For the five-wave panels there were usually used $x = 3$ and in our richness analysis this value is also used.

The richness profiles combine the information on prevalence and on duration of richness, but in the analysis left and right censored spells are included. This is the difference between richness profiles and event history analysis in which left censored spells (unknown starting date) are excluded.

The previous research [e.g. Fouarge, Layte, 2005] shows that there is an ordering among the poverty profile categories, therefore, there will be estimated parameters of ordered logit model to predict the different richness profiles. In the model, certain dummy variables referring to the personal and household characteristics are included: gender, age and education of the household's head, the place of residence, number of household's members, biological family type, socio-economic group, labor force status and the presence of disabled person in household. The variables were measured at the beginning of the richness spell.

In the analysis there was used the ordered logistic regression model also called proportional odds model. Let the response variable Y_i (for the i-th individual) has J ordered categories $j = 1, 2, \ldots, J$. Associated probabilities are defined as $p_{ij} = P(Y_i = j)$ and cumulative probabilities are defined as $F_{ij} = P(Y_i \leq j) = p_{i1} + p_{i2} + \ldots + p_{ij}$. The last one cumulative probability always equals 1 (there are needed first $J - 1$ cumulative probabilities). Then a cumulative logit is defined as

$$\text{logit}(F_{ij}) = \log\left(\frac{F_{ij}}{1 - F_{ij}}\right) = \log\left(\frac{p_{i1} + p_{i2} + \cdots + p_{ij}}{p_{i,j+1} + \cdots + p_{iJ}}\right)$$

and describes the log-odds of two cumulative probabilities, one less-than and the other greater-than type. The cumulative logit measures the probability of being at or below a category divided by its complimentary probability, i.e. probability of being above that category [Liu, 2016]. Cumulative logits contrast the lower levels of response variable with higher levels of response variable or compare higher values to the lower values [Derr, 2013]. For example, a response variable is ordinal with four levels ($J = 1, 2, 3,$ and 4). We can compare higher values to the lower values: the probabilities of category 2, 3, 4 versus 1; the probabilities of category 3 and 4 versus 1 and 2; probabilities of category 4 versus 1, 2, and 3. We can also compare the lower levels with higher levels: probabilities between category 1 and categories 2, 3, and 4; probabilities of being in category 1 and 2 versus 3 and 4; probabilities of categories 1, 2, and 3 versus 4. The interpretation is related to chosen option of cumulative probabilities: going from the lowest to the highest or going from the highest to the lowest.

The cumulative logits are related to covariates in the following logistic regression model:

$$\text{logit}(F_{ij}) = \alpha_j + \mathbf{x}_i'\boldsymbol{\beta},$$

for $j = 1, 2, \ldots, J - 1$, where $\mathbf{x}_i$ is a vector of covariates, $\boldsymbol{\beta}$ is a vector of unknown parameters and α_j is a vector of intercepts between response levels (also

called thresholds). This model implies that the cumulative logits j and j', logit(F_{ij}) and logit($F_{ij'}$), have the same slopes β, but the intercepts α_j differ. This is equal slopes assumption (also called parallel lines assumption). In practice, several binary logistic regression models are estimated simultaneously, for each $J - 1$ categories, in which the intercepts are different, but the slopes are the same. The ordinal logistic regression is estimated using maximum likelihood. To estimate the parameters of the model R program with MASS package [Ripley et al., 2015] was used.

2. Results

The first step of the analysis was to compare proportions of rich households in different groups for x waves (in 'x out of n' approach). The households were divided by personal and household characteristics (Figure 1).
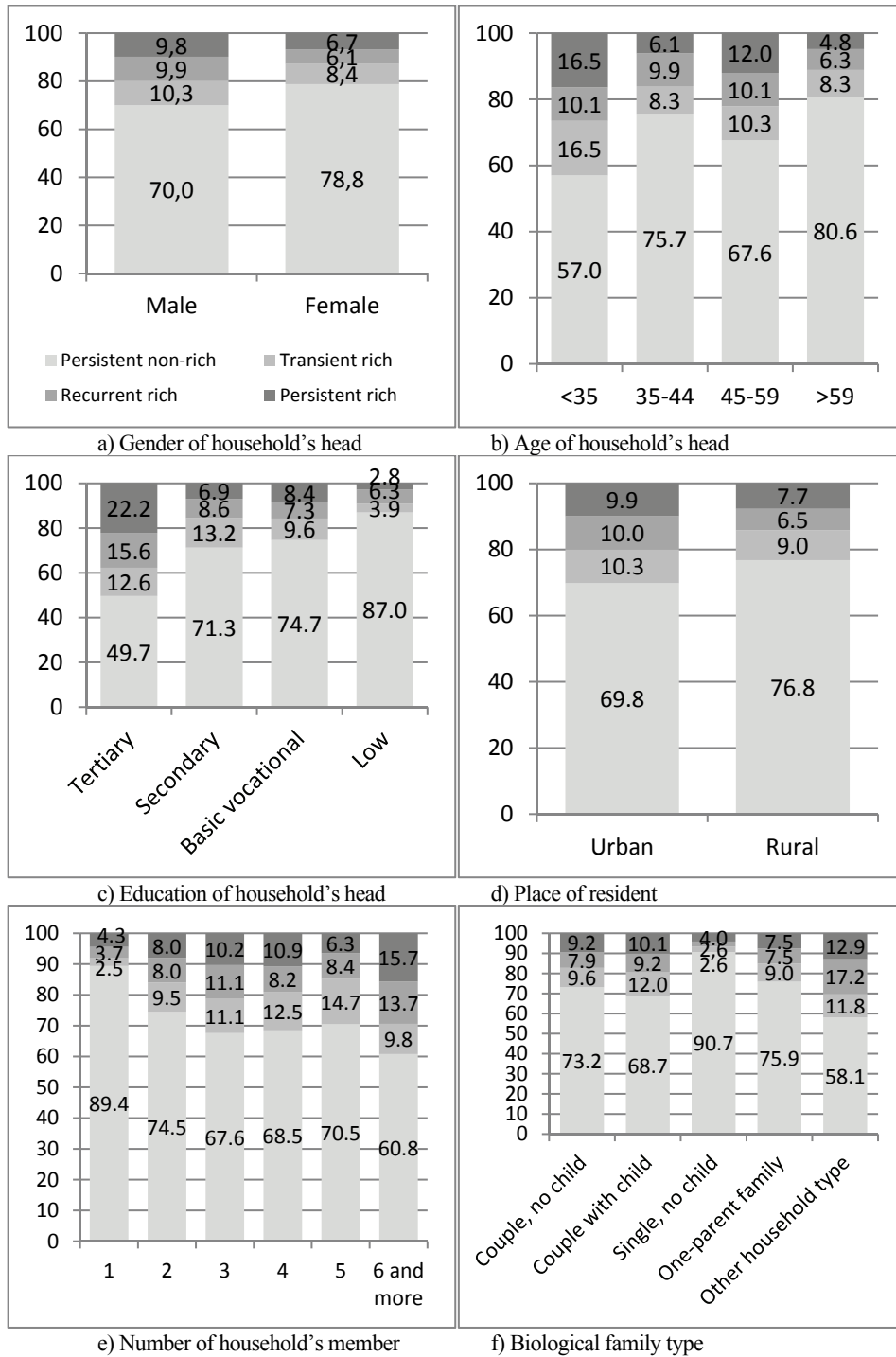
Between 2007 and 2015 in most groups more than half of the households were never rich, the exceptions were two groups: households with high-educated head (49.7%) and households of the self-employed (45.3%). The largest percentage of never rich were in groups of households with low-educated head (87.0%), single households (89.4%) and households living on unearned sources (90.5%). In 'x out of n' approach the group of households of the self-employed and group of households with high educated head were characterized by the highest percentage of persistent rich, i.e. households rich three or more times rich (30.2% and 27.6%, respectively). The lowest proportions of persistent rich were in groups of households: of retirees and pensioners (4.5%), of living on unearned sources (4.8%), with low-educated head (4.4%). It should be noted that based on 'x out of n' approach we cannot answer the question concerning recurrence of the richness. We can only say that e.g. household was three times above the richness line, but we do not know whether the richness lasts continuously or whether periods of richness are interspersed with non-richness periods. For this purpose there were created richness profiles (Figure 2).

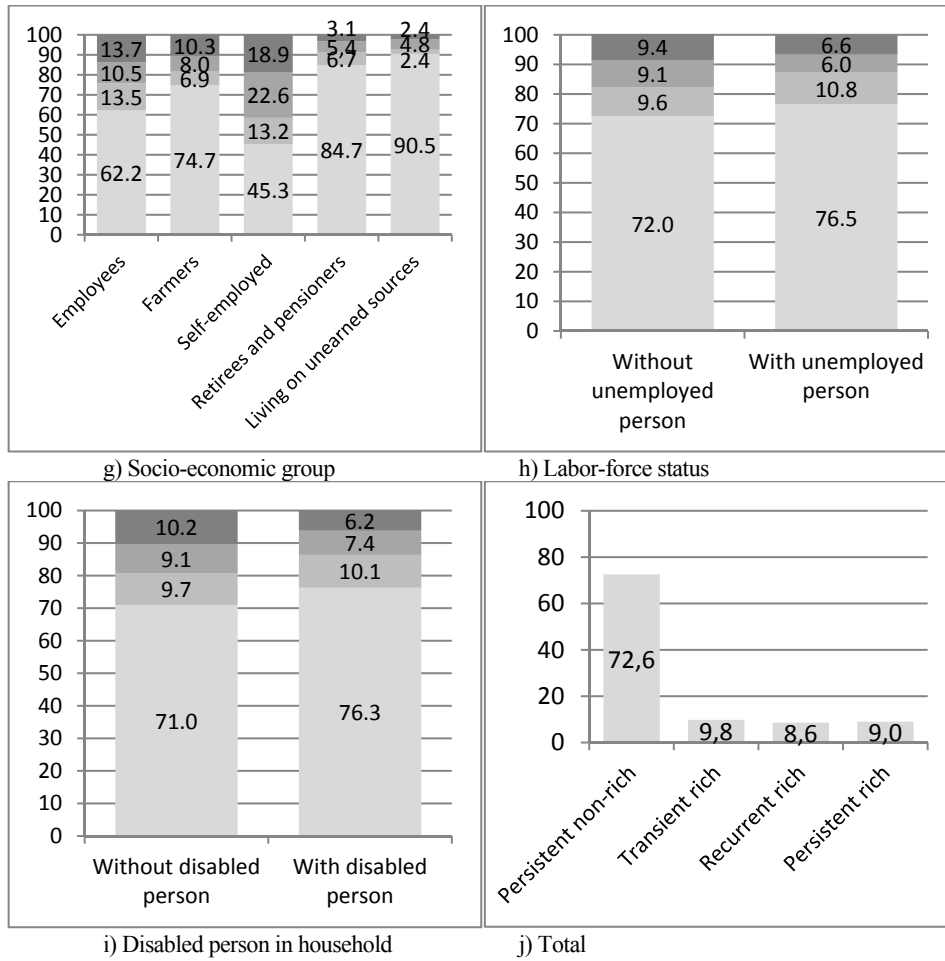**Figure 1.** Proportion of the rich households for *x* waves in Poland
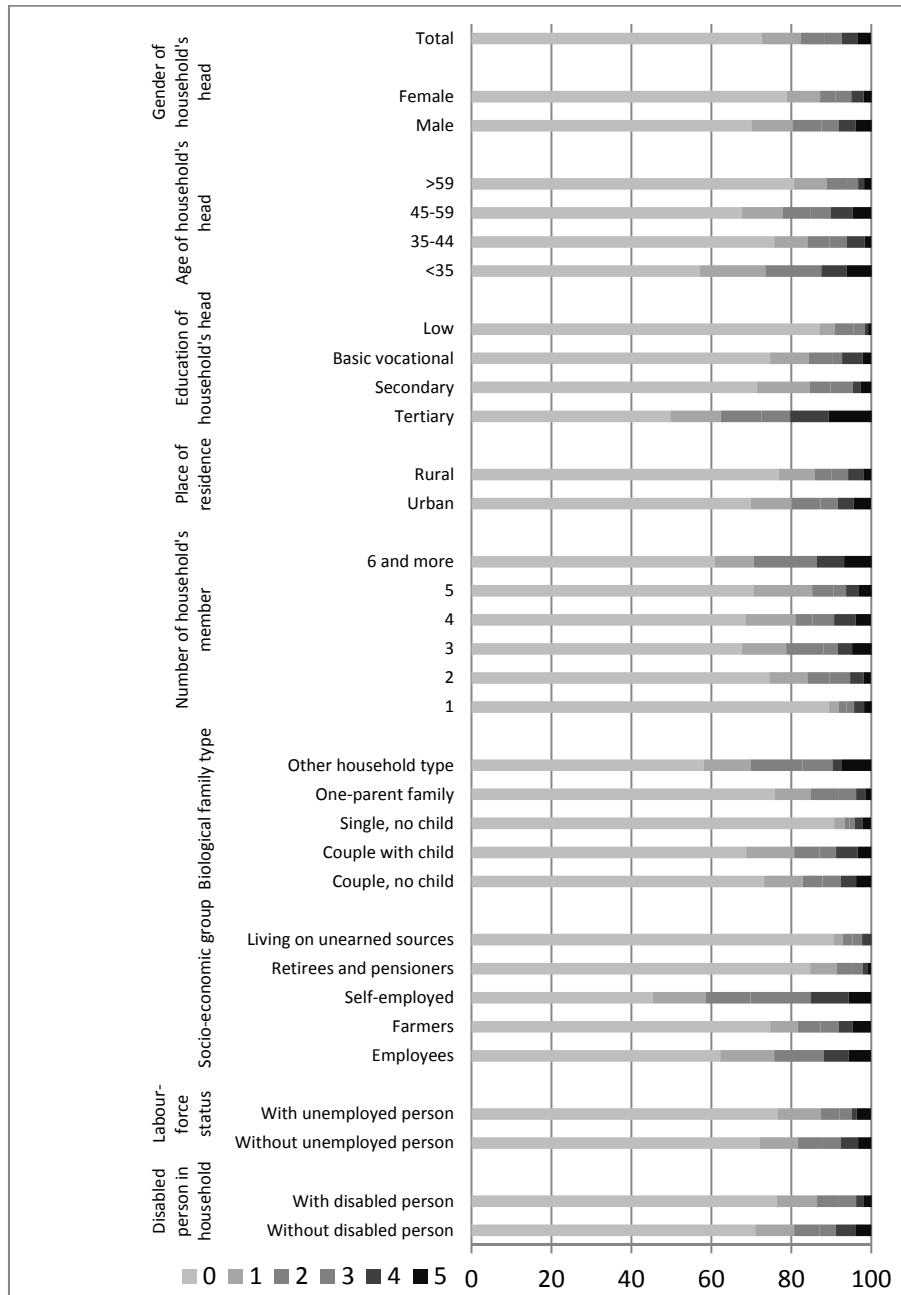
Source: Calculation based on data from: Council for Social Monitoring [2015].

Figure 2. Richness profiles in Poland

a) Gender of household's head

b) Age of household's head

c) Education of household's head

d) Place of resident

e) Number of household's member

f) Biological family type

g) Socio-economic group

h) Labor-force status

i) Disabled person in household

j) Total

Source: Calculation based on data from: Council for Social Monitoring [2015].

The households of the self-employed and households with high-educated head are the most often recurrent rich (22.6% and 15.6%, respectively). These households are also very often persistent rich (18.9% and 22.2%, respectively). Households with head 35- and big size households (6 members and more) are also quite often persistent rich (16.5% and 15.7% respectively). The least often in persistent richness were households of living on unearned sources (2.4%), households with low-educated head of household (2.8%) and households of retirees and pensioners (3.1%).

In the next step there were estimated parameters of the ordinal logistic regression model (Table 1). The outcome in model is:

$$Y_i = \begin{cases} 1 \text{ if household is persistent non rich,} \\ 2 \text{ if household is transient rich,} \\ 3 \text{ if household is recurrent rich,} \\ 4 \text{ if household is persistent rich.} \end{cases}$$

Table 1. Results of ordinal logistic regression model for richness profiles in Poland

Reference group: never rich	Coefficient	Standard Error	Odds Ratio
Persistent non rich|Transient rich	–0.870*	0.377	x
Transient rich|Recurrent rich	–0.201	0.585	x
Recurrent rich|Persistent rich	0.681	0.380	x
Gender of household's head:			
Male head	ref.		
Female head	–0.353	0.194	0.703
Age of household's head:			
Head aged <35	ref.		
Head aged 35-44 (ref: <35)	–0.969**	0.291	0.379
Head aged 45-59	–0.193	0.254	0.824
Head aged >59	–0.119	0.320	0.888
Education of household's head:			
Head has tertiary education	ref.		
Head has secondary education	–1.011***	0.204	0.364
Head has basic vocational education	–1.184***	0.212	0.306
Head has low education (lower secondary, primary and incomplete primary)	–1.734***	0.269	0.177
Place of residence:			
Urban	ref.		
Rural	–0.197	0.174	0.821
Number of household's members	0.069	0.092	1.072
Biological family type:			
Couple, no child	ref.		
Couple with child	–0.193	0.257	0.825
Single, no child	–0.897*	0.353	0.408
One-parent family	0.055	0.300	1.056
Other household type	0.655	0.377	1.925
Socio-economic group:			
Households of employees	ref.		
Households of farmers	–0.317	0.298	0.728
Households of the self-employed	0.596*	0.281	1.815
Households of retirees and pensioners	–1.145***	0.242	0.318
Households of living on unearned sources	–1.463*	0.625	0.232
Labour-force status:			
Household without unemployed person	ref.		
Household with unemployed person	–0.357	0.220	0.700
Disabled person in household:			
Household without disabled person	ref.		
Household with disabled person	–0.069	0.171	0.933
AIC	1734.825		

* Significance codes: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Source: Calculation based on data from: Council for Social Monitoring [2015].

The proportional odds assumption appears to have held because the significance of Chi-Square statistics is 0.868 (this means not to reject the null hypothesis which states that the slope coefficients in the model are the same across response categories)[1]. Based on the proportional odds assumption the odds ratio is the same, regardless of category. The odds of being in the higher richness categories versus in lower richness categories is about 62% lower for households with head aged 35-44 relative to households with head aged <35, given the other variables are held constant in the model. The odds are about 63.5%, 69.5% and 82% lower for households with head having secondary, basic vocational and low education relative to households with head having tertiary education, when the other variables are held constant. The odds of single, no child households relative to couple, no child households are 59% lower controlling for the other explanatory variables. Households of retirees and pensioners, and households of living on unearned sources relative to households of employees have about 68% and 77% (respectively) lower odds of being in better, not in worse richness category, when the other variables are held constant. The odds of being in higher versus in lower richness categories are 1.82 times greater for households of the self--employed relative to households of employees net of the other variables.

Conclusions

The conducted analysis confirms the earlier suspicions concerning the dynamics of the richness. It can be concluded that in 2007-2015 most of households were never rich and only in group of households with high-educated head and in group of households of the self-employed the percentages of never rich were lower than 50%. Households of living on unearned sources and single households were in the worst income situation (percentages of never rich are higher than 90%). Using the ordinal logistic regression model allowed to compare the odds of being in higher richness categories versus in lower categories. Based on the results it can be concluded that the odds are lower in households with head having secondary, basic vocational and low education relative to households with highly-educated head. Almost all categories of socio-economic group are significantly differentiated due to odds of being in better, not in worse richness category, but only households of the self-employed versus households of the employees have higher odds. Summarizing, tertiary education of the

---

[1] Rejection of the null hypothesis indicates that ordered logit coefficients are not equal across the level of the outcome and it should be fit a less restrictive model (e.g., multinomial logit model) [Institute for Digital Research and Education, s.a.].

household's head and self-employment allow to stay the household more number of periods in richness and these periods last longer than in other groups of households.

Further research on richness dynamics will focus on event history models which allow to exclude left censored spells and take into consideration only the spells with known starting date. The analysis will allow to estimate the survival and hazard functions and to determine the factors decreasing and increasing the odds to exit and to enter to richness.

References

Atkinson A., Piketty T., eds. (2007), *Top Incomes over the Twentieth Century*, Oxford University Press, Oxford.

Brzeziński M. (2010), *Income Affluence in Poland*, "Social Indicators Research", Vol. 99, No. 2, pp. 285-299.

Council for Social Monitoring (2015), *Social Diagnosis: Integrated Database*, http://www.diagnoza.com (accessed: 28.12.2015).

Derr B. (2013), *Ordinal Response Modelling with the LOGISTIC Procedure*, Paper 446-2013, SAS Global Forum, San Francisco.

Fouarge D., Layte R. (2003), *Duration of Poverty Spells in Europe*, European Panel Analysis Group (EPAG) Working Papers, 47, University of Essex, Colchester.

Fouarge D., Layte R. (2005), *Welfare Regimes and Poverty Dynamics: The Duration and Recurrence of Poverty Spells in Europe*, "Journal of Social Policy", Vol. 34, No. 3, pp. 407-426.

Institute for Digital Research and Education (s.a.), *Annotated SPSS Output. Ordered Logistic Regression*, http://www.ats.ucla.edu/stat/spss/output/ologit.htm (accessed: 24.08.2016).

KPMG w Polsce (2014), *Rynek dóbr luksusowych w Polsce. Edycja 2014*, Warszawa.

Layte R., Fouarge D. (2004), *The Dynamics of Income Poverty* [in:] R. Berthoud, M. Iacovou (eds.), *Social Europe. Living Standards and Welfare States*, Edward Elgar Publishing, Cheltenham-Northampton, pp. 202-224.

Leigh A. (2009), *Top Incomes* [in:] W. Salverda, B. Nolan, T. Smeeding (eds.), *The Oxford Handbook of Economic Inequality*, Oxford University Press, Oxford.

Liu X. (2016), *Applied Ordinal Logistic Regression Using Stata: From Single-Level to Multilevel Modeling*, Sage Publications, Singapore.

Muffels R., Fouarge D. (2010), *Cyclical Poverty in Matured Welfare States* [in:] S. del Campo, T. Hamada, G. Barbiroli, S. Sassen, E. Barbieri-Masini, P. Nchoji Nkwi, O. Sichone, A. Momoh (eds.), *Social and Economic Development*, Vol. VII, Eolss Publishers, Paris, pp. 11-34.

Muffels R., Fouarge D., Dekker R. (2000), *Longitudinal Poverty and Income Inequality: A Comparative Panel Study for the Netherlands, Germany and the UK*, Working Papers, WP2000-6, OSA Institute for Labour Studies, Tilburg.

Ripley B., Venables B., Bates D.M., Hornik K., Gebhardt A., Firth D. (2015), *MASS: R package version 7.3-37*, https://cran.r-project.org/web/packages/MASS (accessed: 14.08.2016).

Sączewska-Piotrowska A. (2015), *Identyfikacja determinant bogactwa dochodowego z zastosowaniem modelu logitowego*, „Zarządzanie i Finanse", vol. 13, nr 4, cz. 2, pp. 241-259.

Słaby T., ed. (2006), *Konsumpcja elit ekonomicznych w Polsce – ujęcie empiryczne*, SGH, Warszawa.

PROFILE BOGACTWA W POLSCE

Streszczenie: Celem artykułu była analiza dynamiki bogactwa dochodowego w Polsce. Zastosowano prostą metodę, która bazuje na zliczaniu okresów spędzonych ponad linią bogactwa oraz podejście badające liczbę i długość okresów bogactwa. Uzyskane wyniki pozwalają stwierdzić, że największym odsetkiem gospodarstw, które nigdy nie były bogate, cechowała się grupa gospodarstw z nisko wykształconą głową, grupa gospodarstw jednoosobowych oraz grupa utrzymujących się z niezarobkowych źródeł. Na podstawie oszacowanego modelu logitowego można stwierdzić, że szanse pobytu w wyższych kategoriach dochodowych, a nie w niższych kategoriach dochodowych różnią się w statystycznie istotny sposób w zależności od wykształcenia głowy gospodarstwa domowego oraz grupy społeczno-ekonomicznej gospodarstwa.

Słowa kluczowe: bogactwo dochodowe, trwałość bogactwa, powtarzalność bogactwa, porządkowa regresja logistyczna.