DETERMINANTS OF THE STATE OF POVERTY USING LOGISTIC REGRESSION

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY
Nr 16(22)

ISSN 1644-6739
e-ISSN 2449-9765

Anna Sączewska-Piotrowska

Univeristy of Economics in Katowice

e-mail: anna.saczewska-piotrowska@ue.katowice.pl

ORCID: 0000-0001-9760-3649

DOI: 10.15611/sps.2018.16.04
JEL Classification: D31, I32

Abstract: The aim of the paper was to identify the determinants of the state of poverty using logistic regression. The analysis focused on economic poverty considered through the prism of income. Three states of poverty were considered: poverty, near poverty (household's income from 100% to 125% of the adopted poverty threshold) and above near poverty (income higher than 125% of poverty threshold), using the ordinal logit model and – after the rejection of the proportional odds assumption – the multinomial logit model. The analysis was preceded by a presentation of the basic facts concerning three states of poverty. Based on the conducted analysis it can be stated that the education of the household's head, place of residence, labour-force status and socio-economic group were very important factors of the state of poverty, and they change the odds of being in above near poverty relative to poverty and the odds of being in near poverty relative to poverty.

Keywords: state of poverty, ordinal logistic regression, multinomial logistic regression, poverty threshold, household's income.

1. Introduction

The actions of social policy are most often focused on poor and socially excluded households. There are also households living on the edge of poverty (near-poverty state) and they are not supported by government and non-governmental organizations. These households are at risk of entering poverty and attention should be paid to these households to help them to prevent the worsening of their financial situation. It is very important to indicate the factors associated with the state of poverty and forecast which groups of households are most likely to be in poverty and in the near-poverty states.

The aim of the paper was to identify the determinants of a state of poverty using logistic regression. The analysis focused on economic poverty considered through the prism of income. There were considered three poverty states: poverty, near poverty (household's income from 100% to 125% of the adopted poverty threshold) and above near poverty (income higher than 125% of poverty threshold), estimating

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 16(22)

the parameters of the ordinal logistic regression model and – in the case of rejection of the proportional odds assumption – the multinomial logistic regression model. The independent variables were related to the personal characteristics of the household's head and to the household's characteristics.

2. States of poverty

In the standard analysis two states of poverty are studied: poor and non-poor. It is popular in poverty literature to focus on movements between these states using transition matrices and Markov chain models (e.g. [Baulch 2013]).

Three states (low income, middle income and high income class) are considered in income polarization studies to answer the question whether the middle income class is disappearing or not [Wolfson 1994; Kot 2008; Panek 2017]. A different kind of division into three states of poverty is also used in poverty studies – poor, near poor and above near poor. This division allows to focus on groups of households living on the edge of poverty. The income of a household living in near poverty is close to, but not below, the poverty threshold. The term "close to" is not clearly defined. The first idea of near poverty was proposed by Orshansky [1966]. She defined the near poor as those living from 100% to 133% of the poverty threshold. Since 1971, the U.S. Census Bureau reports have contained information about near poor persons defined as those living from 100% to 125% of the poverty threshold. Other authors proposed different solutions concerning the definition of near poverty: Ben-Shalom, Moffitt and Scholz [2011] from 100% to 150%, Short and Smeeding [2012] from 100% to 200%, Hokayem and Heggeness [2014] from 100% to 125%. The first analysis of near poverty in Poland was conducted by Sączewska-Piotrowska [2016a; 2016b]. The near poor were defined as those living from 100% to 125% of the poverty threshold.

Some authors take into account more states of poverty. For example, Baulch [2013] enlarged transition matrices and constructed the extended poverty transition matrices, which divided the welfare distribution into categories based on fractions and multiples of the poverty line. Sometimes in the movement analyses the quintile transition matrices are used [Haughton, Khandker 2009]. Other authors define the states of poverty based on average income. For example, Rendtel, Langeheine and Berntsen [1998] defined three states of poverty: less than 40%, between 40% and 60%, and above 60% of average equivalised income.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 16(22)

3. Data

The facts about the states of poverty in Poland are presented based on panel 2000-2015 (eight waves: 2000 and from 2003 every two years) in the framework of the "Social Diagnosis" project [Council for Social Monitoring 2015]. Table 1 contains the information on the number of households surveyed in subsequent waves of the panel.

Table 1. Number of households in the database of the Social Diagnosis project

Year	2000	2003	2005	2007	2009	2011	2013	2015
Wave	I	II	III	IV	V	VI	VII	VIII
Number of households	3005	3962	3881	5532	12380	12359	12343	11738

Source: own calculation based on [Council for Social Monitoring 2015].

The parameters of ordinal and multinomial logistic regression models are estimated based on the newest wave of the panel.

The poverty analysis adopted the economic definition of poverty. It was assumed that the indicator for poverty measurement is the net income of households in Poland in March/June in subsequent waves of the panel. In order to take into account the differences in a household's size and its composition an equivalised income was calculated by dividing the household's income by its equivalent size. The modified OECD (Organization for Economic Co-operation and Development) equivalence scale was used. This scale assigns 1 to the first adult of the household, 0.5 to each subsequent adult aged 14 or more and 0.3 to children (each person under 14). In the analysis the poverty threshold was set at 60% of the national median income in 2000 and in the subsequent years the inflated thresholds were used . The term "near poverty" referred to income between 100% and 125% of the poverty threshold and the term "above near poor" to income above 125% of the threshold.

4. Ordinal and multinomial logistic regression

In the analysis the ordered logistic regression model also called proportional odds model was used. Let the response variable Y have J ordered categories $j = 1,2, \dots, J$. Associated probabilities are defined as $p_j = P(Y = j)$ and cumulative probabilities are defined as $F_j = P(Y \leq j) = p_1 + p_2 + \dots + p_j$. The last one, cumulative probability, always equals 1 (there are needed first $J - 1$ cumulative probabilities). Then a cumulative logit is defined as

$$\text{logit}(F_j) = \log\left(\frac{F_j}{1-F_j}\right) = \log\left(\frac{p_1 + p_2 + \cdots + p_j}{p_{j+1} + \cdots + p_J}\right)$$

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. the probability of being above that category [Liu 2016]. Cumulative logits contrast the lower levels of response variable with higher levels of response variable. It should be noted that we can also compare the 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 the 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 [Agresti 2002; Derr 2013; Stanisz 2016]:

$$\text{logit}(F_j) = \alpha_j + \mathbf{X}'\boldsymbol{\beta},$$

for $j = 1,2, \ldots, J-1$, where $\mathbf{X} = (X_1, X_2 \ldots, X_p)$ is a vector of covariates, $\boldsymbol{\beta}$ is a vector of unknown slope parameters and α_j is an unknown intercept parameter between response levels (also called threshold). This model has $J - 1 + p$ parameters. The model implies that the cumulative logits j and j', $\text{logit}(F_j)$ and $\text{logit}(F_{j'})$, have the same slopes β, but the intercepts α_j differ. This is an equal slopes assumption (also called parallel lines or proportional odds 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 and to test the proportional odds assumption, R program [R Core Team 2016] with ordinal package [Christensen 2015b] was used. The likelihood ratio tests comparing the likelihoods of a model proportional odds for all terms and models with non-proportional odds for each term in turn were performed.

Under the hypothesis that the two models are equivalent, the statistic asymptotically follows a χ^2 distribution. Further details are given in Christensen [2015a].

Rejection of the parallel lines assumption means that it should fit a less restrictive model. This kind of model is a multinomial (or polytomous) logit model[1]. In multinomial logistic regression, dependent variable Y has more than two nominal categories which do not have an ordinal structure. In brief, it is an expanded form of binary logistic regression model for J categories. One group is chosen to be a reference category for the other groups (estimates equations for $J-1$ groups) [Liao 1994]. For example, if the response variable has 4 categories ($J = 1, 2, 3, \text{and } 4$) and 1 is the base category, then three models are estimated: 2 versus 1, 3 versus 1, and 4 versus 1. Therefore, multinomial logistic regression is a series of binary logit models [Agresti 2002; Derr 2013]:

$$\text{logit}[P(Y = j)] = \log\left(\frac{p_j}{p_{j^*}}\right) = \alpha_j + \mathbf{X}'\boldsymbol{\beta}_j, \text{j} \neq j^*$$

for $j = 1,2, \dots, J-1$, where $\mathbf{X} = (X_1, X_2 \dots, X_p)$ is a vector of covariates, $\boldsymbol{\beta}_j$ is a vector of unknown slope parameters and α_j is an unknown intercept parameter, j^* is a reference category. The j subscripts on both the intercept (α_j) and slope (β_j) indicate that there is an intercept and a slope for the comparison of each category to the reference category. There are $(J-1)(1+p)$ parameters in the model. The multinomial logistic regression is estimated using maximum likelihood. Package nnet [Ripley, Venables 2016] in R program [R Core Team 2016] was used to estimate the parameters of the multinomial logit model.

5. Results

In the first step we calculated the near poverty, poverty and above near poverty rates (percentages of poor, near poor and above near poor households) from 2000 to 2015 in Poland. The thresholds were anchored at 2000 and for the next years the thresholds were multiplied by the inflation factor. The results are shown in Figure 1.

In 2000-2015 the percentage of near poor households was lower than the percentage of poor households, but the difference was decreasing. Definitely the highest values were in the case of above near poverty.

[1] Other models may also be used, e.g. the partial proportional odds model. Examples of using this model in [Dudek 2012; Dudek, Landmesser 2013].

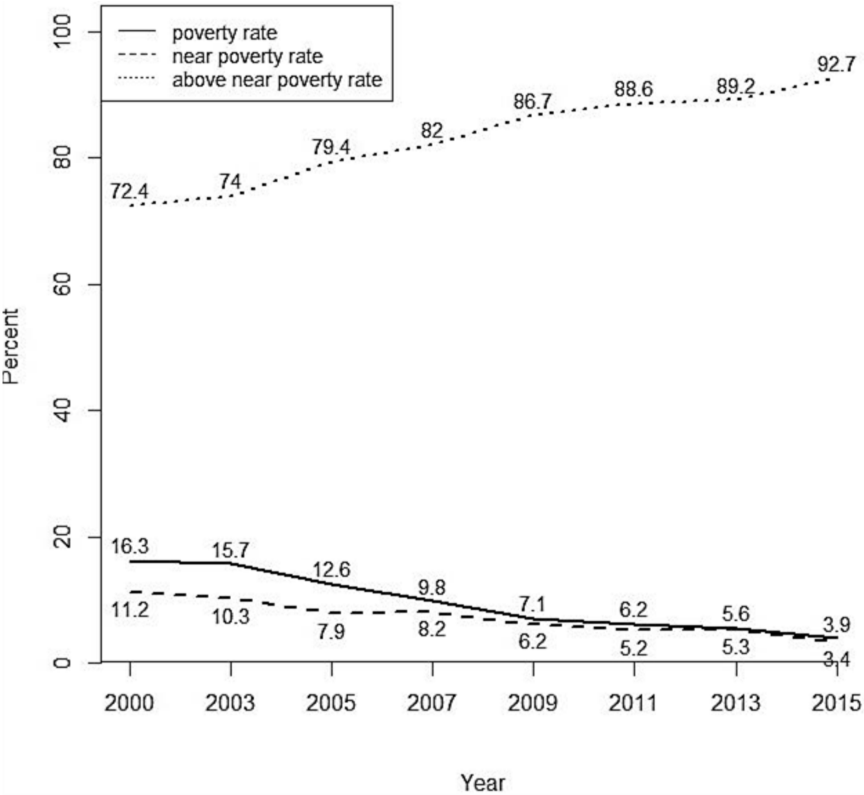

Fig. 1. Poverty, near poverty and above near poverty rates in Poland 2000-2015 calculated for base year 2000, adjusted for inflation

Source: own study based on [Council for Social Monitoring 2015].

During the observation period the rates were changing – poverty and near poverty rates were decreasing and above near poverty rate was increasing. The smallest increase in above near poverty rate and an increase in near poverty rate were in 2013, which means that 2013 was characterized by the least favourable changes in economic situation. That was an echo of the slight recession which took place in Poland in late 2012 and early 2013. Another improvement of economic situation took place in 2015.

In 2015, using the thresholds from 2015, poverty and above near poverty rates are definitely higher (15.4% and 11.1%, respectively), and near poverty rate is lower (73.5%) than using the thresholds from 2000. The estimated models used the thresholds from 2015.

In the estimated logistic regression models for three states of poverty the following variables were included:

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 16(22)

- SEX – sex of household's head: male is a reference category (ref.),
- AGE – age of household's head: AGE1 – 34 and less (ref.), AGE2 – 35-44, AGE3 – 45-59, AGE4 – 60 and above,
- EDU – education of household's head: EDU1 – tertiary (ref.), EDU2 – secondary, EDU3 – basic vocational, EDU4 – low,
- PLACE – place of residence: urban is a reference category,
- NUMBER – number of household's members,
- BIOL – biological family type: BIOL1 – couple, no child (ref.), BIOL2 – couple with child, BIOL3 – single, no child, BIOL4 – one-parent family, BIOL5 – other household type,
- GROUP – socio-economic group: GROUP1 – employees (ref.), GROUP2 – farmers, GROUP3 – self-employed, GROUP4 – retirees and pensioners, GROUP5 – living on unearned sources,
- UNEMP – labour-force status: household without unemployed person is a reference category,
- DIS – disabled person in household: no disabled person is a reference category.

Table 2. Results of ordinal logistic regression model for three states of poverty

Variable	Coefficient	Standard Error	Odds Ratio
Intercept: Above near poor or near poor *vs*. poor Above near poor *vs*. near poor or poor	-5.2903*** -4.4247***	0.1775 0.1750	x x
SEX	-0.38853***	0.06432	0.6781
AGE2	-0.34033**	0.12492	0.7115
AGE3	0.03830	0.11667	1.0390
AGE4	0.63656***	0.13158	1.8900
EDU2	-1.35282***	0.12354	0.2585
EDU3	-1.95641***	0.12289	0.1414
EDU4	-2.43414***	0.12767	0.0877
PLACE	-0.60137***	0.05445	0.5481
NUMBER	-0.15907***	0.03327	0.8529
BIOL2	-0.18129	0.10069	0.8342
BIOL3	-1.11981***	0.09220	0.3263
BIOL4	-0.49347***	0.10289	0.6105
BIOL5	0.35394*	0.14085	1.4247
GROUP2	-0.66184***	0.09914	0.5159
GROUP3	0.50052**	0.16375	1.6496
GROUP4	-0.67724***	0.08568	0.5080
GROUP5	-1.73886***	0.12579	0.1757
UNEMP	-1.34951***	0.07033	0.2594
DIS	-0.43113***	0.05452	0.6498
AIC	12909.88		
McFadden	0,176		

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Source: own calculation based on [Council for Social Monitoring 2015].

ŚLĄSKI PRZEGLĄD STATYSTYCZNY
Nr 16(22)

The results of the estimated ordinal logistic regression model are shown in Table 2.

In the ordinal logistic regression model the higher values to the lower values there were compared. Most variables have a significant influence associated with the odds of being in a state of poverty. For example, the odds of being in the higher poverty categories *vs*. in lower poverty categories is about 91% lower for households with a poorly educated head relative to households with a highly educated household's head. According to Table 3, most variables break the assumption of parallel lines, this means to reject the null hypothesis which states that the slope coefficients in the model are the same across response categories. If the parallel lines assumption does not hold, interpretations about the results will be wrong, therefore in order to find correct results alternative models were used instead of ordinal logit regression models. We decided to use the multinomial logit model.

Table 3. Results of the proportional odds assumption test

Variable	χ^2	*p*
SEX	16.5423	0.000***
AGE2 AGE3 AGE4	0.8023 24.9827 27.7884	0.370 0.000*** 0.000***
EDU2 EDU3 EDU4	0.7155 5.4609 6.9873	0,398 0.019* 0.008**
PLACE	6.1976	0,013*
NUMBER	3.0040	0.083
BIOL2 BIOL3 BIOL4 BIOL5	2.5404 23.5006 3.1915 1.1523	0,111 0.000*** 0.074 0.283
GROUP2 GROUP3 GROUP4 GROUP5	12.3266 0.0134 18.4276 8.8174	0.000*** 0.908 0.000*** 0.003**
UNEMP	11.8647	0.001***
DIS	0.2487	0.618

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Source: own calculation based on [Council for Social Monitoring 2015].

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 16(22)

Results of the estimation of a multinomial logit model are shown in Table 4.

Table 4. Results of multinomial logistic regression model for three states of poverty

Variable	Coefficient	Standard Error	Odds Ratio
Intercept: Near poor *vs*. poor	0.58136	0.31635	x
SEX	0.10574	0.10289	1.1115
AGE2	-0.30601	0.20783	0.7364
AGE3	-0.27235	0.19530	0.7616
AGE4	0.47482*	0.21668	1.6077
EDU2	-0.48837*	0.24170	0.6136
EDU3	-0.59281*	0.23896	0.5528
EDU4	-0.74148**	0.24371	0.4764
PLACE	-0.36185***	0.08942	0.6964
NUMBER	0.04060	0.05262	1.0414
BIOL2	0.34539*	0.17076	1.4125
BIOL3	0.16054	0.15519	1.1741
BIOL4	-0.02325	0.17189	0.9770
BIOL5	0.68128**	0.23090	1.9764
GROUP2	-0.58070***	0.16287	0.5595
GROUP3	0.28806	0.29178	1.3338
GROUP4	-0.38143**	0.13832	0.6829
GROUP5	-1.38735***	0.19600	0.2497
UNEMP	-0.85138***	0.11014	0.4268
DIS	-0.14459	0.08756	0.8654
Intercept: Above near poor *vs*. poor	5.51118***	0.25138	x
SEX	-0.38304***	0.08214	0.6818
AGE2	-0.45551**	0.16331	0.6341
AGE3	-0.03192	0.15229	0.9686
AGE4	0.83045***	0.17128	2.2943
EDU2	-1.62991***	0.19364	0.1959
EDU3	-2.28757***	0.19189	0.1015
EDU4	-2.86360***	0.19674	0.0571
PLACE	-0.75851***	0.07081	0.4684
NUMBER	-0.15022***	0.04196	0.8605
BIOL2	-0.06868	0.12905	0.9336
BIOL3	-1.08447***	0.11878	0.3381
BIOL4	-0.50941***	0.12917	0.6008
BIOL5	0.60558***	0.18151	1.8323
GROUP2	-0.83042***	0.11974	0.4359
GROUP3	0.61986**	0.22836	1.8587
GROUP4	-0.84344***	0.10928	0.4302
GROUP5	-2.06608***	0.15132	0.1269
UNEMP	-1.61273***	0.08563	0.1993
DIS	-0.50595***	0.06990	0.6029
AIC	12859.81		
McFadden	0,181		

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Source: own calculation based on [Council for Social Monitoring 2015].

In the estimated models, 'poor state' was the base category, then two models were estimated: 'near poor' state *vs*. 'poor state' and 'above near poor' state *vs*. 'poor state'. The multinomial logit model compared to ordinal model had a higher value of McFadden's pseudo-R^2 which means the multinomial model is better fitted to the data. We can see that definitely more variables had a significant influence in the case of above near poor state *vs*. poor state than in the case of near poor *vs*. poor state. Female-headed households relative to male-headed households had 31.8% lower odds of being above near poor relative to poor (given all other variables in the model are held constant). Households with a head of 60 and above relative to households with a head aged 34 and less had higher odds of being above near poor than poor, and higher odds of being near poor than poor (2.29 and 1.61 times, respectively). Poorly educated head of a household relative to a highly educated household's head reduced the odds of being above near poor relative to near poor and the odds of being near poor relative to poor. For households in rural areas (the variable urban evaluated at zero) the odds for being in above near poverty relative to poverty were 53.2% lower and the odds for being in near poverty relative to poverty were 30.4% lower (the other variables are held constant). A one unit increase in the number of household's members reduced the odds of being above near poor relative to poor by almost 14% (the other variables in the model are held constant). Households living on unearned sources relative to households of employees had definitely lower odds of being above near poor relative to poor and the lower odds of being near poor relative to poor (87.3% and 75% lower odds, respectively). Households with an unemployed person (the variable without unemployed person evaluated at zero) had lower odds of being in above near poverty relative to poverty and lower odds of being in near poverty relative to poverty.

It should be noted that the term "forecasting" in relation to cross-sectional data refers to a unit of observation not a unit time. Micro forecasts may refer to units in the sample as well as non-sample units. The multinomial logit model was used to determine the probability forecasts of the state of poverty for households with different socio-economic characteristics. Probability forecasts for six households from the database are shown in Table 5.

For all six analysed households the highest probability was in the case of above near poverty, but the level of probability was different

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 16(22)

for households. For example, the probability of being in above near poverty state for household number 2, 3 and 5 was more than twice higher than for household number 6.

Table 5. Probability forecasts of the state of poverty for selected households from the database

Number of household	Poor	Near poor	Above near poor
1	0.13159	0.12936	0.73905
2	0.02438	0.02551	0.95011
3	0.03120	0.05589	0.91291
4	0.10912	0.12306	0.76782
5	0.01919	0.05121	0.92960
6	0.26546	0.29207	0.44247

Source: own calculation based on [Council for Social Monitoring 2015].

The example probability forecasts for non-sample households with a different educational level of household's head (categorical variables fixed at zero and number of households fixed at one) are shown in Table 6.

Table 6. Probability forecasts of the state of poverty for households with different educational level of household's head (SEX, AGE, PLACE, BIOL, GROUP, UNEMP, DIS at zero, NUMBER = 1)

Education level	Poor	Near poor	Above near poor
Tertiary	0.00463	0.00863	0.98674
Secondary	0.02280	0.02605	0.95115
Basic vocational	0.04229	0.04354	0.91417
Low	0.07124	0.06321	0.86555

Source: own calculation based on [Council for Social Monitoring 2015].

Definitely the lowest probability was in the case of the lowest category – in the state of poverty. However the probability of being in poverty is the lowest for households with a highly educated head and the highest for households with poorly educated head.

6. Conclusions

The aim of the paper was to identify the determinants of the state of poverty (poverty, near poverty and above near poverty) using logistic regression, using the ordinal and multinomial logit models; the multinomial model was a better fit to data. Therefore, the state of poverty should not be treated as an ordinal variable.

The conducted analysis is a wider analysis of poverty than in the traditional view and allows to indicate the factors of living on the edge of poverty (and not receiving social assistance – "too rich" households) and factors of living in above near poverty (away from poverty). This analysis provides the opportunities for preventive action and to help households with the lowest odds of being in a good income situation. Based on the analysis it can be stated that education is a very important factor of the state of poverty. Low education of household's head reduces the odds of being in above near poverty relative to poverty and the odds of being in near poverty relative to poverty. For households with one or more unemployed person the odds of being in above near poverty relative to poverty and the odds of being in near poverty relative to poverty are definitely lower than for households without an unemployed person. Households living in rural areas had lower odds of being in above near poverty relative to poverty as well the odds of being in near poverty relative to poverty. Attention should also be paid to households living on unearned sources of income, which have much lower odds (relative to households of employees) to be in above near poverty and in near poverty relative to poverty.

References

Agresti A., 2002, *Categorical Data Analysis. Second edition*, John Wiley & Sons, Hoboken, New Jersey.

Baulch B., 2013, *Understanding Poverty Dynamics And Economic Mobility,* [in:] A. Shepherd, J. Brunt (eds), *Chronic Poverty: Concepts, Causes And Policy*, Palgrave Macmillan, Basingstoke, United Kingdom, pp. 38-59.

Ben-Shalom Y., Moffitt R.A., Scholz J.K., 2011, *An Assessment of the Effectiveness of Anti-Poverty Programs in the United States*, National Bureau of Economic Research Working Paper no. 17042.

Christensen R.H.B., 2015a, *A Tutorial on Fitting Cumulative Link Models with the Ordinal Package*, https://cran.r-project.org/web/packages/ordinal/vignettes/clm_tutorial.pdf.

Christensen R.H.B., 2015b, *Ordinal – Regression Models for Ordinal Data*, R package version 2015.6-28, http://www.cran.r-project.org/package=ordinal.

Council for Social Monitoring, 2015, *Integrated database*, http://www.diagnoza.com.

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

Dudek H., 2012, *Subiektywne skale ekwiwalentności – analiza na podstawie danych o satysfakcji z osiąganych dochodów*, [in:] K. Jajuga, M. Walesiak (eds.), Taksonomia 19, *Klasyfikacja i analiza danych – teoria i zastosowania*, UE, Wrocław, pp. 153-162.

Dudek H., Landmesser J., 2013, *Wpływ relatywnej deprywacji na subiektywne postrzeganie dochodów,* [in:] K. Jajuga, M. Walesiak (eds.), Taksonomia 20, *Klasyfikacja i analiza danych – teoria i zastosowania*, Uniwersytet Ekonomiczny we Wrocławiu, Wrocław, pp. 142-150.

Haughton J., Khandker S.R., 2009, *Handbook on Poverty and Inequality*, The World Bank, Washington, DC.

Hokayem C., Heggeness M.L., 2014, *Factors influencing transitions into and out of near poverty: 2004-2012*, SEHSD Working Paper 2014-05, U.S. Census Bureau, Washington DC.

Kot S.M., 2008, *Polaryzacja ekonomiczna. Teoria i zastosowanie*, Wydawnictwo Naukowe PWN, Warszawa.

Liao T.F., 1994, *Interpreting probability models*, Quantitative Applications in the Social Sciences, no. 101, Thousand Oaks, CA, Sage.

Liu X., 2016, *Applied ordinal logistic regression using Stata: From single-level to multilevel modeling*, SAGE Publications, Singapore.

Orshansky M., 1966, *Recounting the poor – a five-year review*, Social Security Bulletin, 29, pp. 20-37.

Panek T., 2017, *Polaryzacja ekonomiczna w Polsce*, Wiadomości Statystyczne, vol. 1, pp. 41-61.

R Core Team, 2016, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, http://www.r-project.org.

Rendtel U., Langeheine R., Berntsen R., 1998, *The estimation of poverty dynamics using different measurements of household income*, Review of Income and Wealth, Series 44, no. 1, pp. 81-98.

Ripley B., Venables B., 2016, *Nnet: feed-forward neural networks and multinomial log-linear models*, R Package Version 7.3-12, https://cran.r-project.org/web/packages/nnet.

Sączewska-Piotrowska A., 2016a, *Near poverty – definition, factors, predictions*, Ekonometria, vol. 4, no. 54, pp. 82-94.

Sączewska-Piotrowska A., 2016b, *Transitions Into and Out of Near Poverty in Urban and Rural Areas in Poland,* [in:] M. Boďa, V. Mendelová (eds.), *Applications of Mathematics and Statistics in Economics 2016,* Conference Proceedings, Banská Bystrica, Občianske združenie Financ., pp. 302-311.

Short K., Smeeding T., 2012, *Understanding income-to-threshold ratios using the supplemental poverty measure*, U.S. Census Bureau Social, Economic, and Housing Statistics Division Working Paper, no. 2012-18.

Stanisz A., 2016, *Modele regresji logistycznej. Zastosowania w medycynie, naukach przyrodniczych i społecznych*, StatSoft Polska, Kraków.

Wolfson M.C., 1994, *When inequalities diverge*, The American Economic Review, vol. 84, pp. 353-358.

ŚLĄSKI
PRZEGLĄD
STATYSTYCZNY

Nr 16(22)

DETERMINANTY STANU PRZYNALEŻNOŚCI DO SFERY UBÓSTWA Z WYKORZYSTANIEM REGRESJI LOGISTYCZNEJ

Streszczenie: Celem artykułu jest wyznaczenie determinant stanu przynależności do sfery ubóstwa z zastosowaniem regresji logistycznej. Analiza skupia się na ubóstwie ekonomicznym analizowanym przez pryzmat dochodów. Rozważane są trzy stany przynależności: stan ubóstwa, stan blisko ubóstwa (dochody gospodarstwa od 100% do 125% przyjętej granicy ubóstwa) i stan poza zagrożeniem ubóstwem (dochody wyższe niż 125% granicy ubóstwa). Wykorzystano porządkowy model logistyczny oraz – po odrzuceniu założenia proporcjonalnych szans – wielomianowy model logistyczny. Analiza jest poprzedzona prezentacją podstawowych faktów dotyczących wyróżnionych stanów ubóstwa.

Słowa kluczowe: stan przynależności do sfery ubóstwa, porządkowy model logistyczny, wielomianowy model logistyczny, granica ubóstwa, dochód gospodarstwa domowego.