POVERTY RATE IN CZECH HOUSEHOLDS DEPENDING ON THE AGE, SEX AND EDUCATIONAL LEVEL

Jitka Bartošová – Marie Forbelská

Abstract

The paper presents a quantitative study of at-risk-of-poverty Czech households depending on the age and sex of the householder and educational level of the family. At-risk-of-poverty rates of age categories are calculated as the proportion of households with an equivalised income below the poverty threshold, which is set at 60% of the national median equivalised income. The equivalised household income is used to allow comparisons between households of different sizes and composition. The data base for the study at-risk-of-poverty rate Czech households is made up of the results of the sample survey EU-SILC (The European Union Statistics on Income and Living Conditions) between 2005 and 2009. We use Generalized Linear Mixed Models (GLMM) to model at-risk-of-poverty rates between 2005 and 2009. GLMM have become a very powerful and widely used statistical tool. The R environment (R Development Core Team, 2010) is used for GLMM analysis. The modelling results are presented graphically.

Key words: equivalised household income, EU-SILC, generalized linear mixed models (GLMMs), poverty rate

JEL Code: D31, I32, R10

Introduction

The current economic crisis is having an important impact on social surveys. There is a pressure on statistics to provide updated information to monitor the extent of the crisis in the social field. It is the case of EU-SILC, the main source of comparable information on income and living conditions across Europe. As a consequence of the crisis, there has be increase in unemployment and hence of poverty in the Czech Republic and elsewhere. The crisis has mostly worsen financial situation of juniors and seniors. In these two categories, we can observe increase in unemployment the most and therefore we can also expect the highest at-risk-of-poverty rates. The threat of unemployment and poverty rate in European households significantly depends also on the educational level of the family and sex of the householder.

The article deals with modelling the development of risk of poverty in Czech households depending on these factors.

1 Financial power and risk of monetary poverty of Czech households

1.1 EU-SILC

The EU-SILC (European Union Statistics on Income and Living Conditions) is an instrument aiming at collecting timely and comparable cross-sectional and longitudinal multidimensional microdata on income, poverty, social exclusion and living conditions. Investigation is carried out by the so-called rotating panel, where the same households were re-interviewed in the annual intervals for four years. After this time are replaced by other households living in the newly visited homes that are added to the investigation file continuously by the random selection. Longer monitoring of a household permits building image of their social situation, not only in the year, but also the changes and developments over time (see e.g. Marek, 2010, Bílková, 2012, Bílková and Malá, 2012). Further analysis regarding the financial situation of households led to field of modeling the income distribution, risk of monetary poverty and unemployment, search of the factors that influence it or crowding into homogeneous subgroups – clusters. These include works of Malá, 2011, Löster and Langhamrová, 2011, Pacáková, Linda and Sipková, 2012, Řezanková, Löster and Húsek, 2011, Stankovičová, 2010, Šimpach, 2012, Šimpach and Langhamrová, 2012, Želinský, 2012 and 2010.

To construct the models in this paper, the data files coming from EU-SILC survey conducted by Czech Statistical Office in 2005 – 2009 were used. These files contain information about the Czech households for the period 2004 – 2008.

1.2 Equivalised household income and monetary poverty

In the recent past and in present the significant part of research is devoted to the question of measurement and elimination of poverty. In literature dedicated to such issues we encounter several different views of poverty. But there are two primary approaches – the objective and subjective one. The objective approach defines poverty by means of certain criteria concerning income or assets of a person. In contrary, the methods of subjective approach investigate whether the person considers himself/herself poor and perceives the symptoms of poverty or ranks himself/herself among poor.

Within the framework of objective approach we further distinguish the absolute and relative methodologies. The absolute methodologies define poverty using some fixed value.

The relative ones define the poverty via relationship to an important characteristics (average or median income, distribution of income categories, etc.). The importance of social context on determination of poverty is emphasized also in definition accepted by European commission in 1984. According to this definition, as poor can be considered a person, family or group of individuals whose resources (material, cultural and social) are so limited that they disqualify such people from minimally acceptable way of life in member states they live in. For the evaluation of poverty or abundance of household or individuals in EU member states the European commission chose so called monetary poverty. Among the basic comparison criteria ranks beforehand given "typical" level of income separating households (or individuals) endangered by monetary poverty from the others. Such a line, so called threshold of risk of monetary poverty, is prescribed by EU on 60% of median of national equivalent income scaled on single currency Euro in purchasing power parity. Household is thus considered as "monetary poor" if it's disposable income scaled by consuming unit (so called equalized income) lies beneath the poverty threshold. The proportion of households, with equalized income below this threshold is called *risk-of-poverty-rate* (or *head count index*) (see e.g. Bartošová and Želinský, 2013). Calculation of this relative measure of poverty in EU is based on equalized income, whose number of consuming units CU is defined as

$$CU = 1 + 0{,}3 \cdot n_1 + 0{,}5 \cdot n_2, \qquad (1)$$

where n_1 is the count of children between 0 – 13 years of age and n_2 is the number of other household members (except the head of household). The consuming unit represents the size of the household involving only one person (head of household) with the full weight (1). Weights of other members are lowered (for more information, see Bartošová and Bína, 2012).

2 Modelling of poverty rate

Fixed effects models which assume that all observations are independent of each other are not appropriate for analysis of several types of correlated data structures, in particular, for clustered data.

2.1 Generalized linear mixed models (GLMMs)

Generalized linear models (GLMs) represent a class of fixed effects regression models for several types of dependent variables (i.e., continuous, dichotomous, counts). Common Generalized linear models (GLMs) include linear regression, logistic regression, and Poisson regression (see e.g. McCulloch and Searle, 2001).

In clustered designs, subjects are observed nested within larger units. For analysis of such data, random cluster effects can be added into the regression model to account for the correlation of the data. The resulting model is a mixed model including the usual fixed effects for the regressors, and the random effects.

Let i denote the number of clusters and let j denote the nested observation. Assume there are $i = 1,...,N$ clusters and $j = 1,...,n_i$ repeated observations nested within each cluster. A random – intercept model, which is the simplest mixed model, augments the linear predictor with a single random effect for subject i, $\eta_{ij} = \mathbf{x}_{ij} \cdot \boldsymbol{\beta} + v_i$, where v_i is the random effect (one for each cluster). These random effects represent the influence of cluster I , on its repeated observations that is not captured by the observed covariates. These are treated as random effects because the sampled clusters are thought to represent a population of clusters, and they are usually assumed to be distributed as $N(0,\sigma_v^2)$ .The parameter σ_v^2 indicates the variance in the population distribution, and therefore the degree of heterogeneity of clusters. The model can be easily extended to include multiple random effects. The model is now written as

$$\eta_{ij} = \mathbf{x}_{ij} \cdot \boldsymbol{\beta} + \mathbf{z}_{ij} \cdot \mathbf{v}_i \ . \qquad (2)$$

The vector of random effects v_i is assumed to follow a multivariate normal distribution with mean vector $\mathbf{0}$ and variance – covariance matrix $\boldsymbol{\Sigma}_v$. Note that the conditional mean μ_{ij} is now specified as $E(Y_{ij} | \mathbf{v}_i, \mathbf{x}_{ij})$, namely, in terms of the vector of random effects. Parameter estimation in GLMMs typically involves maximum likelihood (ML) or variants of ML. Additionally, the solutions are usually iterative ones that can be numerically quite intensive.

2.2 Logistic ANOVA model with fixed and random effects

The choice of model depends on the types of variables. Our aim is to describe the dynamic of monetary poverty development of Czech households in years 2004 – 2008 depending on sex and age group of the householder and of gained educational level. Dependent variable (regressant) – *risk-of-poverty-rate* – has dichotomous character; independent variables (regressors) are categoricall. They are:

- *year* – five levels (EU-SILC 2005 – 2009)

- *sex* – two categories (M – male, F – female)
- *AgeGroup* – five categories (0,25], (25,30], (30,55], (55,60], (60,100]
- *EducationalLevel* – three levels (1 – primary, 2 – secondary, 3 – tertiary)

With all requirements corresponds logistic ANOVA model with fixed and random effects. This model is specificied by logit link function, so called logit, so that linear predictor is defined by formula

$$\eta_{ij} = \text{logit}(\mu_{ij}) = \frac{\log(\mu_{ij})}{1 - \log(\mu_{ij})}. \qquad (3)$$

Here, the conditional expectation $\mu_{ij} = \mathrm{E}\left(Y_{ij} \middle| \mathbf{v}_i, \mathbf{x}_{ij}\right)$ equals $\mathrm{P}\left(Y_{ij} = 1 \middle| \mathbf{v}_i, \mathbf{x}_{ij}\right)$, namely, the conditional probability of a response given the random effects (and covariate values). This model can also be written as

$$\mathrm{P}\left(Y_{ij} = 1 \middle| \mathbf{v}_i, \mathbf{x}_{ij}, \mathbf{z}_{ij}\right) = \text{logit}^{-1}(\mu_{ij}) = \mathrm{F}_{\text{logist}}(\eta_{ij}) = \frac{1}{1 + \exp(-\eta_{ij})}, \qquad (4)$$

where the inverse link function $\mathrm{F}_{\text{logist}}(\eta_{ij})$ is the logistic cumulative distribution function.

This logistic regression model with fixed and random effects is defined by linear predictor

$$\eta_{ij,k} = \alpha_{sex} \cdot sex_{ij,k} + \left(\beta_{year} + v_{year,AgeGroup} + v_{year,EducationalLevel}\right) \cdot year_{ij,k}, \qquad (5)$$

where i ($i = 1, 2, 3$) indicates gain educational level of household, j ($j = 1, ..., 5$) indicates age group of head of household, and k indicates k-th household in ij-th subgroup.

2.3 Results of construction of model

Percentages of households below threshold of monetary poverty according to used factors are in table 1. The first part of table shows the effect of age category in combination with the sex of a person acting at the head of the household. The second part of table shows the combined effect of the age category with the achieved level of education.

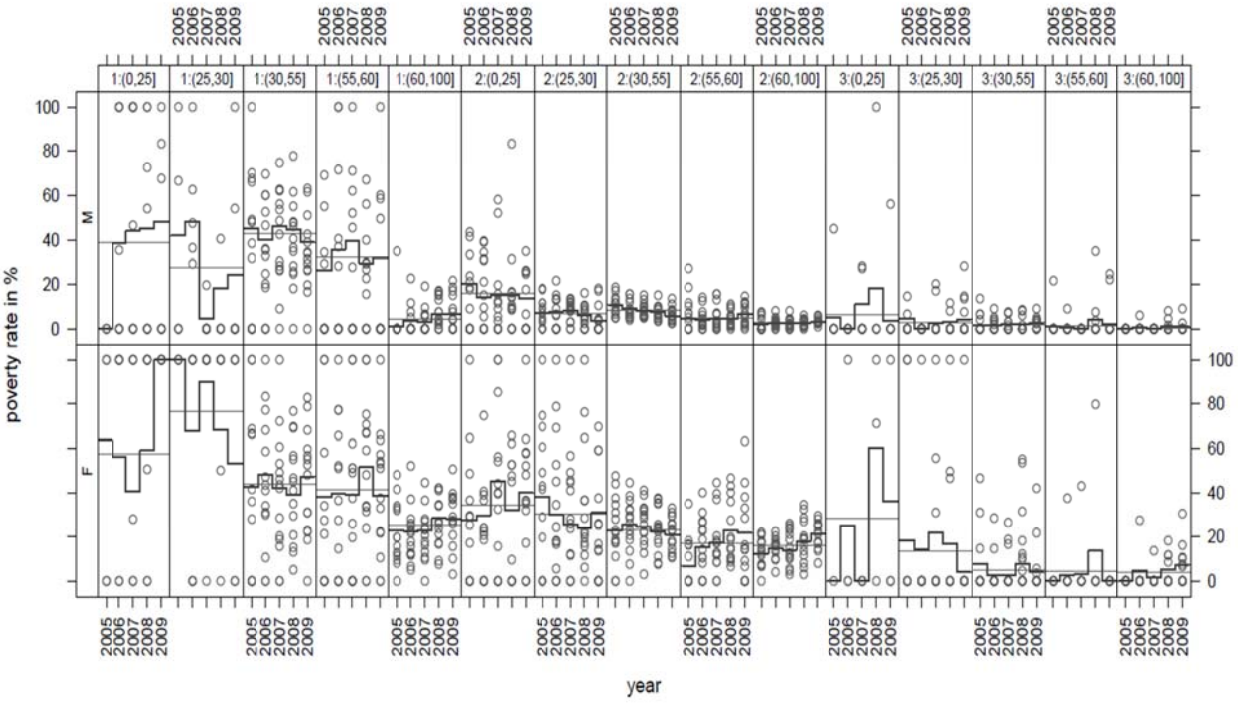
Figure 1 shows the results of the estimated logistic ANOVA model with fixed and random effects. It can be seen that the categories of the youngest households (0,25] whose head has primary or tertiary education are insufficiently occupied. Therefore it would be desirable to omit these subgroups, or to unite these households with households higher age category (25,30].

Tab. 1: At-risk-of-poverty-rate according to age and sex of householder and educational level of the family.

EU-SILC	2005	2006	2007	2008	2009
***Age group* (0,25]**					
male	17,32%	15,85%	18,38%	18,99%	15,72%
female	**30,36%**	**32,10%**	**39,72%**	**39,30%**	**42,32%**
primary level	**27,61%**	**44,04%**	**42,89%**	**48,98%**	**57,66%**
secondary level	21,95%	18,74%	23,04%	18,99%	21,64%
tertiary level	4,98%	9,13%	8,11%	40,41%	11,03%
***Age group* (25,30]**					
male	7,93%	7,20%	7,03%	5,87%	4,71%
female	**39,31%**	**30,13%**	**30,24%**	**26,29%**	**25,16%**
primary level	**55,80%**	**55,02%**	**27,43%**	**34,55%**	**30,40%**
secondary level	11,79%	11,03%	10,48%	8,48%	8,47%
tertiary level	6,29%	1,63%	4,66%	5,06%	4,42%
***Age group* (30,55]**					
male	10,46%	8,80%	8,36%	8,06%	6,51%
female	**23,75%**	**25,58%**	**23,89%**	**22,56%**	**21,97%**
primary level	**44,10%**	**43,55%**	**44,37%**	**42,43%**	**42,49%**
secondary level	12,83%	11,75%	10,93%	10,24%	8,44%
tertiary level	2,33%	1,72%	2,15%	2,93%	3,09%
***Age group* (55,60]**					
male	5,56%	4,79%	5,75%	5,81%	7,26%
female	**12,67%**	**20,31%**	**20,73%**	**27,87%**	**24,80%**
primary level	**31,04%**	**38,04%**	**39,31%**	**41,23%**	**35,62%**
secondary level	5,12%	5,96%	6,98%	7,99%	9,28%
tertiary level	0,99%	0,91%	0,33%	5,60%	2,09%
***Age group* (60,100]**					
male	1,99%	2,45%	2,25%	3,02%	3,38%
female	**15,94%**	**17,18%**	**16,46%**	**20,75%**	**22,70%**
primary level	**16,68%**	**17,49%**	**17,38%**	**22,17%**	**21,91%**
secondary level	5,50%	6,50%	5,95%	7,72%	8,66%
tertiary level	0,00%	1,29%	0,27%	1,80%	2,23%

Source: own calculation

Fig. 1: Results of the estimated logistic ANOVA model. (EU-SILC 2005 – 2009)



Source: own calculation

Conclusion

It has been proven that all three observed factors (age, gender and the level of education) have big impact on a poor level of household income (monetary poverty). Especially in young households, the risk of monetary poverty is much higher than in the middle generation. It also increases along with higher age but more slowly. If the head of the household is a male, then for example in the age category (0,25] the percentage of households with incomes below the poverty line varies from 15,72% (in 2009) to 18,99% (in 2008). If in charge of the household is a woman, this percentage rises to 30,36% (in 2005) – 42,32% (in 2009). For middle-aged households (30 – 55 years) is the situation much better – it is from 6,51% (in 2009) to 10,46% (in 2005) for male, and from 21,97% (in 2009) to 25,58% (in 2006) for female.

Also, the achieved level of education significantly affects the percentage of households that are below the poverty line, regardless of age or gender of the person running the household. For example, with young households (0 – 25 years) the risk of poverty is reduced with the in education received gradually from 27,61% to 21,95% and 4,98% in 2005, and from 27,61% to 21,64% and 11,03% in 2009. In the middle generation (30 – 55 years) it is from 44,10% to 12,83% and 2,33% in 2005, and from 42,49% to 8,44% and 3,09% in 2009.

And in the older households (55 – 60 years) then from 31,04% to 5,12% and 0,99% in 2005, and from 35,62% to 9,28% and 2,09% in 2009.

Acknowledgment

The data set was supported by project VEGA 1/0127/11 "*Spatial Distribution of Poverty in the European Union*" of the Slovak Scientific Grant Agency.

References

Bartošová, Jitka, and Tomáš Želinský. 2013. "Extent of poverty in the Czech and Slovak Republics fifteen years after split." *Post-Communist Economies, 25(1)*: 119 –131.

Bartošová, Jitka, and Vladislav Bína. 2012. "Sensitivity of monetary poverty measures on the setting of parameters concerning equalization of household size." In *Proceedings of 30th International Conference Mathematical Methods in Economics 2012*, ed. Jaroslav Ramík, and Daniel Stavárek, 25–30. Karviná: Silesian University, School of Business Administration.

Bílková, Diana. 2012. „Recent Development of the Wage and Income Distribution in the Czech Republic". *Prague Economic Papers, 2012(2)*: 233-250,

Bílková, Diana, and Ivana Malá. 2011. „Modelling income distribution in the Czech Republic with a special focus on income distribution of professionals in informatics." In *International Days of Statistics and Economics*, ed. Tomáš Löster and Tomáš Pavelka, 51–60. Slaný: Melandrium.

Löster, Tomáš, and Jana Langhamrová. 2011. *Analysis of long-term unemployment in the czech republic* . In Loster Tomas, Pavelka Tomas (Eds.), *International Days of Statistics and Economics* (pp. 307-316). ISBN 978-80-86175-77-5

Malá, Ivana. 2011. „The use of finite mixtures of lognormal distributions in the modelling of incomes of the Czech households." In *International Days of Statistics and Economics*, ed. Tomáš Löster and Tomáš Pavelka, 348–358. Slaný: Melandrium.

Marek, Luboš. 2010. „Analýza vývoje mezd v ČR v letech 1995–2008". *Politická ekonomie* 58(2): 168–206 .

McCulloch, Charles E., and Shayle R. Searle. 2001. *Generalized, Linear, and Mixed Models*. New York: Wiley.

Pacáková, Viera, Linda, Bohdan, and Ľubica Sipková. 2012. "Rozdelenie a faktory najvyšších miezd zamestnancov v Slovenskej republike." *Ekonomický časopis 60(9)*: 935–948.

Řezanková, Hana, Löster, Tomáš, and Dušan Húsek. 2011. „Evaluation of categorical data clustering.“ In *Advances in Intelligent Web Mastering – 3*, 173–182. Berlin: Springer Verlag.

Stankovičová, Iveta. 2010. “Regional Aspects of Monetary Poverty in Slovakia.” In *Social Capital, Human Capital and Poverty in the Regions of Slovakia*, ed. Iveta Pauhofová, Oto Hudec, and Tomáš Želinský, 67–75. Košice: Technical University of Košice.

Šimpach, Ondřej. 2012. „Faster convergence for estimates of parameters of Gompertz-Makeham function using available methods in solver MS Excel 2010.“ In *Proceedings of 30th International Conference Mathematical Methods in Economics 2012*, ed. Jaroslav Ramík, and Daniel Stavárek, 870–874. Karviná: Silesian University, School of Business Administration.

Šimpach, Ondřej, and Jitka Langhamrová. 2012. „Czech household computer facilities as a reliable variable in a life expectancy forecast model up to the year 2060.“ In IDIMT-2012, 143–152. Linz: Trauner.

Želinský, Tomáš. 2012. „Changes in Relative Material Deprivation in Regions of Slovakia and the Czech Republic.“ *Panoeconomicus 59(3)*: 335-353.

Želinský, Tomáš. 2010. “Analysis of Poverty in Slovakia Based on the Concept of Relative Deprivation.” *Politická ekonomie, 58(4)*: 542–565 .

Contact

Jitka Bartošová
University of Economics in Prague, Faculty of Management, Department of Exact Methods
Jarošovská 1117/II, 37701 Jindřichův Hradec, Czech Republic
bartosov@fm.vse.cz

Marie Forbelská
Masaryk University in Brno, Faculty of Science, Department of Mathematics and Statistics
Kotlářská 2, 611 37 Brno, Czech Republic
forbel@math.muni.cz