QUANTITATIVE STUDY OF MONETARY POVERTY IN THE CZECH REPUBLIC AND SLOVAKIA

Jitka Bartošová – Marie Forbelská

Abstract

It is difficult to describe the actual state of economics. However, economics of a country is clearly reflected by the financial situation of households. Therefore, one way to analyse the actual state of economics can be done by quantitative illustration of the financial potential of households. The current economical crisis has the greatest negative impact on the poorest households below the poverty line. Our paper focuses on quantification and modelling of the financial situation of households in correlation with the poverty line. For modelling we use GLM-mixtures. Our research is based on data sets EU-SILC.

Key words: financial potential, GLM-mixtures, household, monetary poverty

JEL Code: 62H30, 30C40

Introduction

Poverty currently presents serious social and economic problem in both developing and developed countries. The current financial crises affect the poorest households who have incomes under the line of relative poverty or near it. Poverty in the Czech Republic and Slovakia affect the 'lower' strata of our society, those with a worse approach to the labour market. It is understandable that there are regional differences in poverty due to the fact that in regions with a higher concentration of these risk factors it must be expected to find a higher rate of poverty and unemployment. Regional disparity concerning the financial potential and poverty of its inhabitants is connected to the development of the individual regions, their economic and demographic structure. It is necessary to recognize that though there may be many hidden cause of poverty, which will be shown in the problems in which we will successfully classify the limits of the subgroups (cluster) with similar financial situations. This will enable us to forecast the whole spectrum of factors which affects the unfortunate situation in the regions, and consequently find a way for the leaders to improve or eliminate the problem.

The aim of this contribution is to illustrate classifications of individual regions in the Czech Republic and Slovakia for the perspective of the at-risk poverty rates in the relatively affluent, middle and poor regions and track the dynamic development of poverty in these three groups, followed by an evaluation and comparison of the situation and development of individual regions in both countries.

1 Risk-of-Poverty Rate

When comparing poverty rates in advanced countries, and now in all countries of the EU, at-risk-of-poverty rate is most frequently used. This is represented by a percentage share equivalent disposable income lower than the poverty line of all the given number of groups of individuals. It acts as a relative measurement which evaluates the financial security of households (individuals) with respect to the national level; (more on this e.g. (Foster, Greet and Thorbecke, 1984), more applications see in (Stankovičová, 2009) or (Želinský, 2010)).

Basic to the so-called poverty threshold which is set by the European bank to be 60% of the national median of the so-called equivalent disposable income (including social transfers) displayed on the modified OECD scale. (Old-age pensions and retirements are taken in this case as income before transfers (see www.czso,cz).) For calculating the limits of poverty (the poverty threshold), there is a given method on which basis it is possible to carry out an international comparison inside the EU.

As it is known, the Czech Republic and Slovakia belong within the borders of the EU as countries with the lowest relative levels of threatened monetary poverty. This is partly due to the fact that the modified OECD scale used isn't quite in accordance with the Czech and Slovak realities because relative spending for bigger households is underestimated.

2 Model Based Clustering

One method which is used when clustering and classifying objects are the methods based on the probability models (Model Based Clustering). There is a range of these methods – regression mixtures, linear models and general linear models – which are built upon a mixture of dense finite mixtures, linear models or general linear models. Bearing this in mind, it is our job to model the risk of monetary poverty i.e. binary quantity, and we must use generalised linear models and logistic regression.

2.1 Finite Mixtures

The density $f_Y(y)$ randomly dividend by the quantity *Y* is the final mixture of density, if

$$Y \sim f_Y(y) = \sum_{k=1}^{K} \pi_k f_k(y) \tag{1}$$

where $\pi_k > 0$, $k = 1,...,K$, $\sum_{k=1}^{K} \pi_k = 1$. $f_k(y)$ is the density of the individual components and π_k, $k = 1,...,K$, is their priori probability (weight, proportion).

For forecasting values of unknown individual parameters of the model the expectation-maximization (EM) algorithm is usually used, see (Dempster et al., 1977). When definition classifying the object y to one of the final number of classes, we use the so-called posterior probabilities. Details are available e.g. in a monograph (McLachlan and Peel, 2000).

2.2 Regression Mixtures

Finite mixtures of density are also applicable for regression mixtures (RM) by replacing the unconditional density distribution by the conditional one, i.e.

$$Y \mid \mathbf{x} \sim f_Y(y; \mathbf{x}, \boldsymbol{\Psi}) = \sum_{k=1}^{K} \pi_k f_k(y; \mathbf{x}, \beta_k, \sigma_k^2) \tag{2}$$

where $\mathbf{x}$ is the vector of explanatory variables, and $\boldsymbol{\Psi} = (\pi_1, \ldots, \pi_{K-1}, \beta_1', \ldots, \beta_K', \sigma_1^2, ..., \sigma_K^2)'$ is a vector of unknown parameters of the model (see McCullagh and Nelder, 1994, Nelder and Wedderburn, 1972, or McCulloch and Searle, 2001). More information about applications, see e.g. in (Forbelská and Bartošová, 2010) or (Bartošová and Forbelská, 2010).

In the case of mixture of generalised linear models (GLMM) the density of conditional distribution of a random variable *Y* has the form

$$Y \mid \mathbf{x} \sim f_Y(y; \mathbf{x}, \boldsymbol{\Psi}) = \sum_{k=1}^{K} \pi_k f_k(y; \mathbf{x}, \eta_{ik}, \phi_k) \tag{3}$$

where $\boldsymbol{\Psi} = (\pi_1, \ldots, \pi_{K-1}, \eta_{11}, \ldots, \eta_{KK}, \phi_1, ..., \phi_K)'$ is a vector of unknown parameters of the model. The mean value of the *k*-th component $\eta_k = \mathbf{x}'\beta_k$ is bound up with a linear predictor by so called link function by the formula $\eta_k = g(\mu_k)$. The maximum likelihood estimation of parameters of the model can be achieved by application of EM algorithm (see Dempster et al, 1977),

Considering that our task is to model the risk of monetary poverty, i.e. binary variable ($U \in \{0;1\}$) it is necessary to choose logistic regression from the generalized linear models. As a link function it is necessary to use so called logit function $g(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$ which is a logarithm of chances, it has the values within the interval $(-\infty, \infty)$ and is canonical (see Jiang, 2007).

Since we have at our disposal data achieved within five years $J = 5$) from 14 regions of the Czech Republic and 8 regions from Slovakia ($H = 14$, and $H = 8$) and the classification will be executed in three types according to the measure of risk of poverty ($K = 3$), instead of alternative random variables ($U \sim A(\mu)$, where $\mu = P(U = 1)$), we will consider the binomial variables $Y_{jh} = \sum_{i=1}^{n_{jh}} U_{jhi} \sim Bi(n_{jh}, \mu_{jh})$ with link function

$$g(\mu_{jh}) = \log\left(\frac{\mu_{jh}}{1-\mu_{jh}}\right) = \log\left(\frac{n_{jh}\mu_{jh}}{n_{jh} - n_{jh}\mu_{jh}}\right) \tag{4}$$

where $j = 1,..., J$, $h = 1,..., H$, $i = 1,..., n_{jh}$, $\sum_{j=1}^{J}\sum_{h=1}^{H} n_{jh} = N$. A linear predictor of the *k*-th component of the mixture is simply given as $\eta_{jh(k)} = m_k + a_{jk}$ where $j = 1,..., J$, $h = 2,..., H$.

3 Classification of regions at the risk-of-monetary-poverty measure

3.1 Data and calculation basis

The data basis was formed by collections achieved from statistical sampling of EU SILC (European Union – Statistics on Income and Living Conditions) containing representative data on income distribution of particular household types, data on type, quality and financial demands of dwelling, household provision with objects of long-term consumption, and on working, material and health conditions of adults living in a joint household. The data from EU SILC investigation between 2005 and 2009 were used for the analysis.

All the calculations were executed by means of freeware R easily available on the internet (http://cran.r-project.org/). R means the language and computing environment suitable for executing of statistical calculations and creating graphical output. It originates in the language S and thanks to a wide users' basis it disposes of an amount additional sets enlarging its capacity. An additional set *flexmix* will be used for modelling the risk of monetary poverty of households by the means of regression mixtures (see Gruen and Leisch, 2007, 2008 or Leisch, 2004).

3.2 Resulting classification of regions in the Czech and Slovak Republics according to the risk of monetary poverty based on the GLMM models

Table 1 shows the percentage of households below the poverty line in all regions of the Czech and Slovak Republics.

Tab. 1: The percentage of households below the poverty line in Czech and Slovak regions

Regions	2005	2006	2007	2008	2009
Czech Republic					
11 Capital Prague R.	4.40	5.72	3.85	6.76	5.62
21 Central Bohemian R.	8.88	9.00	8.59	9.79	8.86
31 South Bohemian R.	7.53	8.27	8.75	7.88	8.18
32 Pilsen R.	7.08	6.36	6.77	5.73	9.49
41 Carlsbad R.	12.59	13.82	11.45	12.29	12.78
42 R. of Ústí nad Labem	15.96	14.91	14.96	13.58	12.38
51 Liberec R.	9.08	10.12	10.22	9.15	10.13
52 R. of Hradec Králové	10.76	7.81	7.92	7.88	7.34
53 Pardubice R.	13.73	8.75	10.99	7.96	9.21
61 Highlands R.	8.79	6.50	7.59	7.35	7.15
62 South Moravian R.	9.65	11.71	9.36	12.80	10.80
71 Olomouc R.	16.18	17.57	15.12	15.90	15.95
72 Zlín R.	12.92	8.14	10.75	11.02	8.11
81 Moravian-Silesian R.	15.05	14.29	13.91	13.73	13.33
CZ – total	**10.63**	**10.26**	**9.81**	**10.32**	**9.85**
Slovak Republic					
1 Capital Bratislava R.	7.81	7.25	5.76	7.43	7.27
2 Trnava R.	10.86	9.68	9.04	8.33	11.00
3 Trenčín R.	13.82	10.36	7.97	10.38	8.52
4 Nitra R.	16.22	13.90	13.39	15.45	15.13
5 Žilina R.	11.63	10.72	8.75	10.99	9.73
6. R. of Banská Bystrica	10.17	12.97	11.91	13.72	14.09
7. Prešov R.	18.97	14.98	13.90	14.02	14.32
8. Košice R.	12.32	12.72	10.65	10.75	10.98
SK – total	**12.82**	**11.70**	**10.29**	**11.52**	**11.50**

Source: Own calculations; data from CZSO.

The results of classification based on probability models show that in the case of the Czech Republic three regions (i.e. 21.4%) rank in the rich group (i.e. group with the lowest risk of monetary poverty), seven regions (i.e. 50%) rank in the middle group (i.e. group with the medium risk of monetary poverty) and four regions (i.e. 28.6%) rank in the poor group (i.e. group with the highest risk of monetary poverty) (see Tab. 2). In Slovakia only one region ranks in the rich group (12.5%), four regions rank in the middle group (i.e. 50%) and three regions rank in the poor group (37.5%). In both countries 50% of regions rank in the middle group (see Tab. 2). The different ranking in the rich and poor groups shows higher concentration of poor regions in Slovakia (see Tab. 2).

Tab. 2: The GLMM models – the estimation of poverty in the model components and the classification of the regions

Component	2005	2006	2007	2008	2009	Classification
Czech Republic						
Rich	5.91	6.03	5.28	6.64	6.83	11, 32, 61
Middle	10.06	9.34	9.34	9.91	9.10	21, 31, 51, 52, 53, 62, 72
Poor	15.30	15.11	14.21	13.99	13.55	41, 42, 71, 81
Slovak Republic						
Rich	7.81	7.25	5.76	7.43	7.27	1
Middle	12.18	11.00	9.18	10.20	10.09	2, 3, 5, 8
Poor	15.14	13.94	13.07	14.42	14.53	4, 6, 7

Source: Own calculations; data from CZSO.

Conclusions

The aim of this contribution was not only to execute cluster analysis in the Czech and Slovak regions by the means of probability models, specifically by the means of the mixture of GLM models, but first and foremost to execute the classification of the regions in three groups (rich, middle, poor) according to their risk of monetary poverty and to estimate the dynamics of the poverty development within each of the groups. The results of the classification by the means of the GLMM models have proved higher proportion of poor regions in Slovakia in comparison with the Czech Republic. However, neither in the Czech Republic nor in Slovakia the results show any tendency to promote differentiation of regions considering the monetary poverty risks. They rather indicate the contradictory tendency.

References

BARTOŠOVÁ, J., FORBELSKÁ, M.: Mixture Model Clustering for Household Incomes. *Aplimat - Journal of Applied Mathematics*, Bratislava : Slovak University of Technology in Bratislava, 163-172, 2010.

DEMPSTER, A. P., LAIRD, N. M. RUBIN, D. B.: Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological) 39 (1)*, 1-38, 1977.

FORBELSKÁ, M., BARTOŠOVÁ, J.: Clustering of Czech Household Incomes Over Very Short Time Period. In *Proceedings of COMPSTAT'2010. Book of abstracts*. 19th International Conference on Computational Statistics, 1023-1030, 2010.

FOSTER, J., GREER, J., THORBECKE, E.: A Class of Decomposable Poverty Measures. *Econometrica*. *52 (3)*, 761-766, 1984.

GRUEN, B., LEISCH, F.: Fitting finite mixtures of generalized linear regressions in R. *Computational Statistics & Data Analysis, 51(11)*, 5247-5252, 2007.

GRUEN, B., LEISCH, F.: FLEXMIX: Version 2. Finite mixtures with concomitant variables and varying and constant parameters. *Journal of Statistical Software, 28(4)*, 1-35, 2008.

JIANG, J.: *Linear and Generalized Linear Mixed Models and Their Applications*. Springer, Berlin. 2007.

LEISCH, F.: FlexMix: A general framework for finite mixture models and latent class regression in R. *Journal of Statistical Software, 11(8)*, 1-18, 2004.

McCULLAGH, P., NELDER, J. A.: *Generalized Linear Models*. Chapman and Hall, London 1994.

McCULLOCH, C. E., SEALRE, S. R.: *Generalized, Linear, and Mixed Models*. Wiley, New York, 2001.

McLACHLAN, G. J. , PEEL, D.: *Finite mixture models*. New York: Wiley & Sons, 2000.

NELDER, J. A., WEDDERBURN, R. W. M.: Generalized linear models. *Journal of the Royal Statistical Society - Series A 135(3)*: 370384. 1972.

R Development Core Team: R: A language and environment for statistical computing. R. Foundation for Statistical Computing, Vienna, Austria. 2010. URL http://www.R-project.org.

STANKOVIČOVÁ, I. Analýza činiteľov monetárnej chudoby v domácnostiach Českej republiky. *Forum Statisticum Slovacum 5(7)*. 151-156, 2009.

ŽELINSKÝ, T.: Porovnanie alternatívnych prístupov k odhadu individuálneho blahobytu domácností ohrozených rizikom chudoby. *Ekonomický časopis* 58 (3). 251 – 270, 2010.

Contact

Jitka Bartošová
University of Economics in Prague, Faculty of Management, Department of Information Management, 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