ISSN 2045-3345

RESEARCH JOURNAL OF ECONOMICS, BUSINESS AND ICT

RICH, AVERAGE AND POOR REGIONS IN THE CZECH AND SLOVAK REPUBLIC – MODEL BASED CLUSTERING

Jitka Bartošová
University of Economics in Prague, Czech Republic

ABSTRACT

One way to analyse the actual state of economics can be done by quantitative illustration of the financial power of households. The current economical crisis has the greatest negative impact on the poorest households below the poverty threshold. Therefore, this paper focuses on quantification of the financial situation of households in individual regions in correlation with the poverty threshold. It contains description of methods used and results of their application with respect to evaluation of spatial distribution of poverty of population on the regional level in the Czech Republic and Slovakia. The methodology is based on finite mixtures of regression models that belong to methods called Model Based Clustering. It concerns special methods of clustering of objects that are based on probability models. The criterion for clustering of regions is the level of risk of poverty rate of households. The regions are divided into three clusters (components) – "rich", "middle", and "poor". The households are scored according to the level of national poverty threshold, in our case according to the level 60% of median of the national equivalised disposable household income. The results of the statistical survey EU SILC (European Union – Statistics on Income and Living Conditions) made by the Czech Statistical Office and the Statistical Office of the Slovak Republic in year 2005 through 2009 form the data base. All calculations have been made in the freeware programming environment R, which is accessible on the internet (http://cran.r-project.org/). For the purpose of modelling of the poverty rate of households using the regression clusters, the upgrade package flexmix was used.

JEL CLASSIFICATION & KEYWORDS

■ C40 ■ F470 ■ CLUSTERING ■ WEALTH DISTRIBUTION

INTRODUCTION

The evaluation of the level of homogeneity of economic potential and wealth distribution across states as well as regions of one state is in the focus of the EU member states. The question of increasing differentiation or even polarisation of economic potential and wealth among states and regions is also of importance. The level of income of population of the region is one of the criteria used for evaluation of the region's potential. This article was inspired by these thoughts; it deals with classification of regions in the Czech Republic and Slovakia based on the risk of poverty rate. The methodology used for the classification is based on finite mixtures of regression models that belong to methods called Model Based Clustering (Longford, 2007).

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 (Pauhofova, 2010). 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 (Bartosova and Forbelska, 2010b, Marek, 2010, Bilkova, 2011).

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.

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, Greer and Thorbecke, 1984), more applications see in (Stankovicova, 2010) or (Zelinsky, 2010a,b)).

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.

Finite Mixtures

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

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

where $f_k(y)$ is the density of the individual component,

$\pi_k > 0$, $k = 1,...,K$ and $\sum_{k=1}^{K} \pi_k = 1$, is its prior probability (weight, proportion).

To illustrate the final mixture of density

$$f_Y(y;\mathbf{\Psi}) = \sum_{k=1}^{K} \pi_k f_k(y;\mathbf{\theta}_k)$$

where $\mathbf{\Psi} = (\pi_1,\ldots,\pi_{K-1},\mathbf{\theta}'_1,\ldots,\mathbf{\theta}'_K)'$ is the vector of unknown parameters of individual components, we first have to determine the type of distribution. On the basis of a random sample $y_1,\ldots,y_N$ the vector parameters is then estimated as $\mathbf{\Psi}$ Unknown individual parameters of the model are determined by the maximum likelihood estimation (MLE) or by restricted maximum likelihood estimation (RMLE). 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

$$\omega_k(y) = \frac{\pi_k f_k(y;\mathbf{\theta}_k)}{f(y;\mathbf{\Psi})} \ , \ k = 1,...,K$$

The subject characterized by *y* is associated with the components *k* whose posterior probability is maximal. Details are available e.g. in a monograph (McLachlan and Peel, 2000), about applications see e.g. in (Bartosova and Forbelska, 2011a,b, or Mala, 2011a,b, Longford and D´urso, 2010, Longford and Pittau, 2006).

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.

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

where $\mathbf{x}$ is the vector of explanatory variables, and

$\mathbf{\Psi} = (\pi_1,\ldots,\pi_{K-1},\beta'_1,\ldots,\beta'_K,\sigma_1^2,...,\sigma_K^2)'$ is a vector of unknown parameters of the model.

If the conditional density distribution is normal, then the density of the *k*-th component has the form

$$f_k(y;\mathbf{x},\beta_k,\sigma_k^2) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left\{-\frac{1}{2\sigma_k^2}\left(y - \mathbf{x}'\beta_k\right)^2\right\}$$

In the case of mixture of generalised linear models (GLMM) we suppose that the densities f_k of particular components are exponential. This general class contains a wide scale of known discrete and continuous distributions, e.g. an alternative, binomial, Poisson's, negative binomial, normal, exponential, gamma distributions etc. The density of these components is generally determined by the formula

$$f_k(y;\mathbf{x},\theta_k,\phi_k) = \exp\left\{\frac{y\theta_k - \gamma(\theta_k)}{\psi(\phi_k)} + d\left(y,\phi_k\right)\right\}$$

Where $\gamma(\theta_k)$ is a known strictly monotone and twice continuously differentiable function, $\psi(\phi_k)$ is a known positive function most frequently formed

$\psi(\phi_k) = \phi_k / \omega$, $d\left(y,\phi_k\right)$

is a known function, θ_k and ϕ_k are parameters (congenial and disturbing).

$\psi(\phi_k) > 0$, $\phi_k > 0$ and $\omega > 0$ are mean, scale factor and known priori weight. It generally applies for the mean value and variance of exponential density that

$$\mu_k = \frac{\partial \gamma(\theta_k)}{\partial \theta_k}$$

$$\sigma_k^2 = \frac{\partial^2 \gamma(\theta_k)}{\partial \theta_k^2}\psi(\phi_k) = V(\theta_k)\psi(\phi_k)$$

where $V(\theta_k)$ is so-called variance function.

Another component of the GLM models is a linear predictor

$$\eta_k = \mathbf{x}'\beta_k$$

which is a linear function of the vector of explanatory variables $\mathbf{X}$. 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 formula

$$\eta_k = g(\mu_k)$$

link function must be differentiated and strictly monotone. If, in addition, it applies that

$g(\mu_k) = \theta_k$, it is so called canonical link function.

(See in McCullagh and Nelder, 1994, Nelder and Wedderburn, 1972, or McCulloch and Searle, 2001). More information about applications, see e.g. in (Forbelska and Bartosova, 2010, Bartosova and Forbelska, 2010b, or Jarosova, 2010).

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

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

where

$$\mathbf{\Psi} = (\pi_1,\ldots,\pi_{K-1},\eta_{11},\ldots,\eta_{KK},\phi_1,\ldots,\phi_K)'$$

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

$$\eta_k = g(\mu_k) \ .$$

If values $\left[y_1,\mathbf{x}_1\right],\ldots,\left[y_N,\mathbf{x}_N\right]$ are available then the maximum likelihood estimation of parameters of

$$\hat{\mathbf{\Psi}}_{MLE} = \left(\hat{\pi}_1,\ldots,\hat{\pi}_{K-1},\hat{\eta}_{1,1},\ldots,\hat{\eta}_{K,K},K,\hat{\phi}_1,\ldots,\hat{\phi}_K\right)$$

can be achieved by application of EM algorithm (see Dempster et al, 1977). Logarithm likelihood function is given by the formula

$$\log L(\mathbf{\Psi}) = \sum_{i=1}^{N} \log\left\{\sum_{k=1}^{K} \pi_k f_k(y_i \left| \mathbf{x}_i,\eta_{i,k},\phi_k)\right.\right\}$$

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 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 ($H \in \{8;14\}$) 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 given by

$$Y_{jh} = \sum_{i=1}^{n_{jh}} U_{jhi} \quad \sim \quad Bi(n_{jh},\mu_{jh})$$

with link function

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

where $j = 1,\ldots,J$, $h = 1,\ldots,H$, $i = 1,\ldots,n_{jh}$

and $\sum_{j=1}^{J}\sum_{h=1}^{H} n_{jh} = N$. A linear predictor of the k-th component of the mixture is simply given as

$$\eta_{jh(k)} = m_k + a_{jk}$$

where $j = 1,\ldots,J$ and $h = 2,\ldots,H$.

Data and equivalent scale of household incomes

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 is a language and at the same time environment suitable for executing statistical calculations and graph outputs. For the purpose of modelling of the poverty rate of households using the regression clusters, the upgrade package flexmix was used.

The household incomes must be considered in correspondence to the assumed household's needs. The needs of individuals differ in dependence on the factors of the number of household members, the age structure of household members and the location of household. For the comparison of various household types and consideration of above-mentioned factors we employ so called equivalence scale. Thus we obtain equalized incomes given by the ratio of total annual net income and the number of consuming units in household enumerated by the equivalence scale. It is given by formula

$$VP = CP/CU \ ,$$

where *VP* is the equalized income, *CP* represents the total annual net income of households (including the social transfers) and *CU* is the number of consuming units in considered household.

In the past in Czech Republic and in Slovakia for mutual comparison of financial power of households, income per person was used. Employing this scale we stem from the assumption that the number of consuming units in household is equal to the number of its members. The main advantage of this method is its simplicity and transparency. But the equalized income computed using this scale can nowadays be very confusing, since the expenditures for housing in both states constantly grow and the expenditure structure gets closer to the structure typical for western European countries. Therefore, nowadays in Czech and Slovak Republic the so called OECD scale is used, which is proposed by Organisation for Economic Co-operation and

Development (OECD). This scale considers the savings stemming from the common running of the household and its age structure. The first adult person is considered fully, i.e. with coefficient 1, every other adult person has lower weight given by coefficient 0.7 and, finally, child defined as a person younger then fourteen is considered with coefficient 0.5. The number of consuming unit according to the OECD definition is thus given by formula

$$CU_{OECD} = 1 + 0.7(n_1 - 1) + 0.5n_2 \; ,$$

where CU_{OECD} stands for the number of consuming units according to the OECD definition, n_1 represents the number of adult household members (i.e. persons in age of 14 or older), n_2 number of children (i.e. persons with the age range of 0 – 13). The member (n_1-1) expresses the count of "other" adult members in the household.

According to the fact that the share of finance spent by households on housing perpetually grows, OECD decided to upgrade this scale in such a manner that it take into account higher economies of scale. The coefficients of increase for other adult household members and for children under fourteen were decreased. The modified OECD scale thus uses coefficient of 0.5 for additional adult person and coefficient 0.3 for a child. Thus the formula for enumeration of modified consuming units is given by formula

$$CU_{OECDmod} = 1 + 0.5n_1 - 1) + 0.3n_2 \; .$$

It is obvious that this scale is very "regardful" particularly to the more numerous households. Modified scale of OECD is accepted and used in all EU countries and EUROSTAT for the computation of so called equalized income, which allows mutual comparison of financial situation of all household types. For comparison of the situation in the Czech Republic and in other EU countries the re-computation of national equalized income (in CZK) to the purchasing power parity (PPP) is used.

Risk of monetary poverty in the Czech and Slovak regions

Tables 1 and 2 show the percentage of households below the poverty line in all regions of the Czech and Slovak Republics. The risk of poverty rates are shown in table 1 and 2. (The regions are identified by numbers used in EU SILC data sets.)

Table 1 and Figure 1 show that households in Capital Prague (id 11) were the wealthiest throughout the entire five-year period. The value of the poverty rate (3.85%) was lowest in Prague in 2007 and highest (6.76%) in the consequent year, when the world crisis took effect. In 2008 the ratio of households living in Prague below the poverty threshold almost doubled. Throughout the entire five-year period (2005 – 2009), 5.28% of households were below the poverty threshold in the capital, which is approximately 1/20 of households. The worst situation was in the Olomouc regions (id 71), in which the rate of poverty risk ranged from 15.12% (in 2007) up to 17.57% (in 2006) and on average was three times higher than in Prague (16.15%).

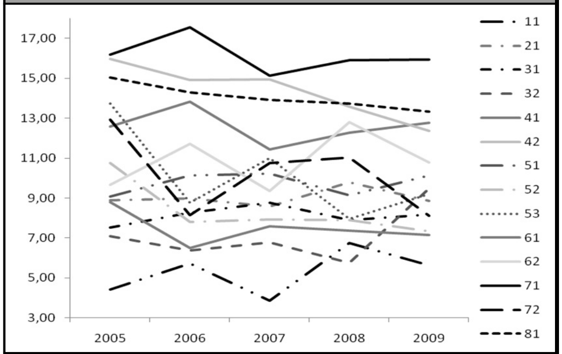
Similarly as in the Czech Republic, also in Slovakia the poverty risk was the lowest in Capital Bratislava (id 1). The average rate of households living in Bratislava below the poverty threshold was 7.10%, minimum was reached in 2007 (5.76%) and maximum was reached in 2005 (7.81%). The highest average rate of Slovakian households below the poverty threshold (15.22%) was in the Presov Region (id 7). Situation gradually improved in this region (from 18.97% in 2005 to 13.90% in 2007). In 2008 and 2009 the Nitra region became the poorest (id 4), in which the rate of households below the poverty threshold reached 15.45% and 15.13%.The results are shown in Table 2 and Figure 2.

Table 1: The percentage of households below the poverty threshold in Czech regions

Regions	2005-9	2005	2006	2007	2008	2009
11 Capital Prague	5,28	4,4	5,72	3,85	6,76	5,62
21 Central Bohemian	9,03	8,88	9	8,59	9,79	8,86
31 South Bohemian	8,12	7,53	8,27	8,75	7,88	8,18
32 Pilsen	7,1	7,08	6,36	6,77	5,73	9,49
41 Carlsbad	12,59	12,59	13,82	11,45	12,29	12,78
42 R. of Usti nad Labem	14,35	15,96	14,91	14,96	13,58	12,38
51 Liberec	9,74	9,08	10,12	10,22	9,15	10,13
52 R. of Hradec Kralove	8,34	10,76	7,81	7,92	7,88	7,34
53 Pardubice	10,11	13,73	8,75	10,99	7,96	9,21
61 Highlands	7,47	8,79	6,5	7,59	7,35	7,15
62 South Moravian	10,88	9,65	11,71	9,36	12,8	10,8
71 Olomouc	16,15	16,18	17,57	15,12	15,9	15,95
72 Zlin	10,19	12,92	8,14	10,75	11,02	8,11
81 Moravian-Silesian	14,06	15,05	14,29	13,91	13,73	13,33
CZ – total	10,17	10,63	10,26	9,81	10,32	9,85

Source: Own calculations

Figure 1: The percentage of households below poverty threshold in Czech regions (2005 - 2006)



Source: Own calculations

Figure 3 compares the ratio of households below the poverty threshold in the Czech Republic and Slovakia in years 2005 – 2009. It shows that until 2007 in both countries the relative poverty was decreasing and homogenizing, i.e. the difference between the Czech Republic and Slovakia was diminishing. However, since the year 2008, the relative poverty increases again and the situation in the Czech Republic and Slovakia differentiates more.

Classification of regions according to the risk of monetary poverty based on the GLMM models

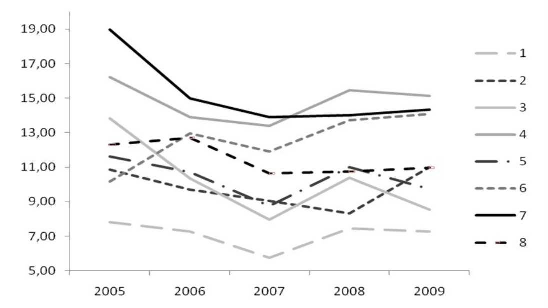
The results of classification based on probability models show Table 3 and Graph 4. We can see, that in the case of the Czech Republic three regions (i.e. 21.4%) rank in the rich group (i.e. group with the lowest risk of monetary poverty), seven regions (i.e. 50%) rank in the middle group (i.e. group 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. 3). 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. 3). The different ranking in the rich and poor groups shows higher concentration of poor regions in Slovakia (see Tab. 3).

Table 2: The percentage of households below the poverty threshold in Slovak regions

Regions	2005-9	2005	2006	2007	2008	2009
1 Capital Bratislava	7,1	7,81	7,25	5,76	7,43	7,27
2 Trnava	9,77	10,86	9,68	9,04	8,33	11
3 Trencin	10,19	13,82	10,36	7,97	10,38	8,52
4 Nitra	14,82	1,22	13,9	13,39	15,45	15,13
5 Zilina	10,35	11,63	10,72	8,75	10,99	9,73
6 R. of Banska Bystrica	12,58	10,17	12,97	11,91	13,72	14,09
7 Presov	15,22	18,97	14,98	13,9	14,02	14,32
8 Kosice	11,48	12,32	12,72	10,65	10,75	10,98
SK - total	11,55	12,82	11,7	0,29	11,52	11,5

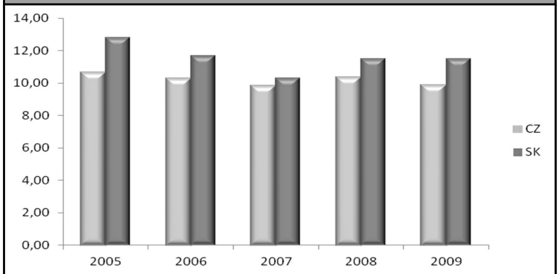
Source: Own calculations

Figure 2: The percentage of households below poverty threshold in Slovak regions (2005 – 2009)



Source: Own calculations

Figure 3: The percentage of households below poverty threshold in the Czech and Slovak Republics (2005 – 2009)



Source: Own calculations

Figure 4 shows dynamics of development of the poverty rate in the individual cluster components in year 2005 through 2009. Whereas in poor regions („Poor") the situation was more or less identical in both countries in the beginning of the period (year 2005), the other two groups („Middle" and „Rich") differed, Czech households being richer. Over the period of the five years, the differentiation between Czech and Slovak regions with respect to the „Middle" and „Rich" groups diminished. In 2007 the difference was at its minimum. Afterwards, the differentiation of the „Middle" group inreased again, while differentiation in the „Rich" groups remained low.

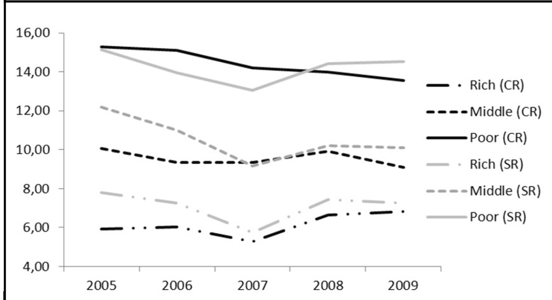
The situation of the „Poor" groups evolved in a different manner. The relative poverty in the Czech regions that were included into this component, decreased steadily. The relative poverty of the Slovak regions in this component, first decreased quicky (the minimum was reached in 2007, and consequently, the differeniation between Czech and Slovak regions increased, Slovak regions being richer. However, afterwards the poverty rate in the Slovak regions increased up to the initial value, resulting into Czech households being richer.

Table 3: 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	9,1	21, 31, 51, 52, 53, 62, 72
Poor	15,3	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	9,18	10,2	10,09	2, 3, 5, 8
Poor	15,14	13,94	13,07	14,42	14,53	4, 6 ,7

Source: Own calculations

Figure 4: The percentage of households below poverty threshold in Slovak regions (2005 – 2009)



Source: Own calculations

Conclusion

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" and "Poor") according to their risk of monetary poverty and to estimate the dynamics of the poverty development within each of the groups. The results show that cluster analysis methods, known as Model Based Clustering, enable to form clusters of households with similar characteristics, based on which households can be successfully classified into individual classes.

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

Bartosova, Jitka, and Marie Forbelska,"Mixture Model Clustering for Household Incomes." Aplimat – Journal of Applied Mathematics 3(2010a): 163-172.

Bartosova, Jitka, Marie Forbelska, and eds. "Comparison of Regional Monetary Poverty in the Czech and Slovak Republic." Social Capital, Human Capital and Poverty in the Regions of Slovakia. Kosice : Technical University of Kosice, 2010b: 76-84.

Bilkova, Diana. "Analysis of the Development in Wage Distributions of Men and Women in the Czech Republic in Recent Years." Statistika – Journal for Economy and Statistics 48.1(2011): 40-57.

Dempster, Arthur P., et al. "Likelihood from Incomplete Data via the EM Algorithm." Journal of the Royal Statistical Society: Series B 39.1(1977): 1-38.

Forbelska, Marie, Jitka Bartosova, and eds. "Clustering of Czech Household Incomes Over Very Short Time Period." Proceedings of COMPSTAT'2010. Paris : Physica Verlag HD, 2010: 1023-1030.

Foster, James, Joel Greer, and Erik Thorbecke. "A Class of Decomposable Poverty Measures." Econometrica 52.3(1984): 761-766.

Jiang, Jiming. Linear and Generalized Linear Mixed Models and Their Applications. Berlin : Springer Verlag, 2007.

Jarosova, Eva, and eds. "Estimation with the Linear mixed effects model." Proceedings of Aplimat'2010. CD-ROM. Bratislava : Slovak university of technology in Bratislava, 2010: 913–920.

Longford, Nicholas T., et al., eds. "Discussion of 'Model-based clustering for social networks' by Handcock, M.S., Raftery, A.E., and Tantrum, J.M." Journal of the Royal Statistical Society: Series B 170.2(2007): 329.

Longford, Nicholas T., and M. Grazia Pittau. "Stability of household income in European countries in the 1990's." Computational Statistics and Data Analysis 51.2(2006): 1364-1383.

Longford, Nicholas T., and Pierpaolo D´urso. "Mixtures with an improper components." Journal of Applied Statistics 38.11(2011): 2511-2521.

Mala, Ivana. "Conditional Distributions of Finite Mixtures of Probability Distributions." AMSE'2011. Wroclaw : Uniwersytet Ekonomiczny, 2011a: 18.

Mala, Ivana. "Conditional Distributions of Incomes and their Characteristics." Proceedings of ISI'2011. Dublin : ISI, 2011b: 1–6.

Marek, Lubos. "The Trend of Income Distributions in Czech Republic in the Years 1995-2008 Analysis." Politicka ekonomie. 2010. 2(2010): 186-206.

McCullagh, Peter, and John Ashworth Nelder. Generalized Linear Models. London : Chapman and Hall, 1994.

McCulloch, Charles E., and Shayle Robert Searle. Generalized, Linear, and Mixed Models. New York : Wiley & Sons, 2001.

McLachlan, Geoffrey J., and David Peel. Finite mixture models. New York : Wiley & Sons, 2000.

Nelder, John A., and Robert William Maclagan Wedderburn. "Generalized linear models." Journal of the Royal Statistical : Society Series A 135.3(1972): 370-384.

Pauhofova, Iveta, and eds. "Regional Income Stratification of Slovak Population." Social Capital, Human Capital and Poverty in the Regions of Slovakia. Kosice : Technical University of Kosice, 2010: 22-30.´

Stankovicova, Iveta, and eds. "Regional aspects of monetary poverty in Slovakia." Social Capital, Human Capital and Poverty in the Regions of Slovakia. Kosice : Technical University of Kosice, 2010, 67-75.

Vecernik, Jiri. Czech Society in the 2000s: a report on socio-economic policies and structures. Praha : ACADEMIA, 2009.

Zelinsky, Tomas. "Comparison of Alternative Approaches to Estimation of Individual Welfare of Poor Households." Ekonomicky casopis / Journal of Economics. 58. (2010a): 251-270.

Zelinsky, Tomas. "Analysis of Poverty in Slovakia Based on the Concept of Relative Deprivation." Politicka ekonomie. 58.3(2010b): 542-565.