

VARIABLES SELECTION FOR QUALITY OF LIFE EVALUATION IN THE EUROPEAN UNION

Z. Pacáková

Received: August 31, 2012

Abstract

PACÁKOVÁ, Z.: *Variables selection for quality of life evaluation in the European Union*. Acta univ. agric. et silvic. Mendel. Brun., 2012, LX, No. 7, pp. 199–206

In 2008 Commission on the Measurement of Economic Performance and Social Progress introduced eight dimensions characterizing the economic performance as well as social well-being without proposing concrete list of variables. The aim of the article is therefore to propose a set of indicators for overall quality of life evaluation in the European Union in accordance to the dimensions introduced by the Commission. Part of the main objective is to propose a methodological tool for selection of indicators reflecting the needs of evaluating quality of life in the European Union. At the beginning a set of 114 variables was completed as a result of the findings listed by the Commission and based on empirical literature. The primary data set was reduced in two consecutive steps. The first step aims to reduce the correlation among the variables. As a result of the second steps, variables reflecting the different levels among consistent groups of states were selected. Therefore, the methodical approach in the second step was based on nonparametric procedures (tests) used to identify variables that are of significantly different levels in clusters identified by cluster analysis. As a result set of 52 variables for quality of life evaluation drawn from 4 different data sources was introduced.

economic performance, well-being, quality of life, correlation analysis, cluster analysis, nonparametric procedures, European Union

Evaluation of economic performance as well as social well-being, the overall quality of life, is a research topic that is of huge interest in recent years. In the second half of the 20th century it has started research that aims to construct a measure or to propose a set of indicators that extend the Gross Domestic Product (GDP) as a measure of economic productivity and quality of life. These activities are known as “activities beyond GDP”.

The need of constructing a supplemental measure that characterizes the quality of life complexly led in different activities which are proceed under organisations such as United Nations Development Programme, European Union or OECD. One of the most recent activities started in 2008 by establishing Commission on the Measurement of Economic Performance and Social Progress (often referred to as the Stiglitz commission). The Commission aimed among others to identify the limits of GDP as an indicator of economic performance and social progress, to consider what additional information

might be required for measuring social progress and to assess the feasibility of alternative measurement tools. (Commission on the Measurement of Economic Performance and Social Progress, 2008) There is more or less agreement among the researchers that quality of life should be measured as a multidimensional event. Many authors (Gönner *et al.*, 2007; Osberg and Sharpe, 2002, 2010; Noorbakhsh, 1998) evaluate the quality of life on the basis of a set of indicators. Some of them modify given indicators such as Human Development Index (HDI) while others construct new summary indicators. Quality of life is than often evaluated both on the basis of monetary and nonmonetary indicators as well as objective and subjective variables.

On the other hand there is no wide agreement on the number of dimensions. As a result of various research activities different numbers of both dimensions and indicators were proposed. Human Development Index combines three

dimensions: life expectancy, education and income. OECD well-being indicators (OECD, 2011) cover eleven dimensions and Legatum Prosperity Index (Legatum Institute, 2010) consists of eight subindices.

The Stiglitz commission (Commission on the Measurement of Economic Performance and Social Progress, 2008) has identified eight dimensions for measuring well-being: i) *Material living standards (income, consumption, wealth)*, ii) *Health*, iii) *Education*; iv) *Personal activities including work*, v) *Political voice and governance*, vi) *Social connections and relationships*, vii) *Environment (present and future conditions)*, viii) *Insecurity, of an economic as well as physical nature*.

The Commission does not propose concrete set of indicators, the aim of the Commission is to motivate discussion about measuring well-being and quality of life and initiate research aiming to construct better and complex measures.

The aim of the article is therefore to propose a set of indicators for overall quality of life evaluation in the European Union in accordance to the dimensions introduced by the Commission. Part of the main objective is to propose a methodological tool for selection of such indicators for the European Union.

MATERIAL AND METHODS

Although the Commission does not provide concrete set of indicators, the dimensions are discussed in detail in the final report of the Commission. The final report presents concrete and detailed recommendations for variables selection and some concrete indicators are discussed. The aim was to find indicators which are easily available and cover the dimensions described above. Other research activities as well as availability of official data (e. g. data available in a set of Structural Indicators) were taken into account when completing the primary list of variables. Due to recommendations of the Commission both objective and subjective indicators were included.

Previously the selection process, requirements on data availability were set down. Due to the aim of overall study of the author the aim was to select variables available in time series covering the period

from 2004 till the last available data (mostly 2009, 2010 if available). Some additional requirements on data availability were set down: i) variables availability for at least 80% of the EU states, ii) no more than 40% values missing in each time series, it means maximum two values missing in time series covering six years.

Due to huge lack of data two member states of the EU were not included in the analysis: Bulgaria and Romania.

The primary data set compiles data from four data sources as it is shown in the Tab. I.

Most of the data come from the Eurostat database, data from this database are well comparable and there is the maximum of existing data. To cover the dimension Political voice and governance, indicators from World Bank database were used. Subjective indicators were drawn from questionnaire surveys and were used to complete the Social connections and relationships dimension above all.

At the beginning the eight dimensions were covered by 114 indicators (Tab. II).

Complete list of variables can be found in Pacáková (2012). The primary set of variables was selected on the basis of recommendations of the Commission, findings of other research activities were also taken into account. To provide that the final set of indicators is of required characteristics discussed below, the primary list of indicators was reduced in two steps.

The variables selection process can be based on either objective or subjective methods, these two approaches can be combined. The objective methods of selection are often based on expert's evaluation (see e.g. Kuprová and Kamenický, 2006). Objective methods (mostly various statistical

I: Data sources and the proportion of variables for each source

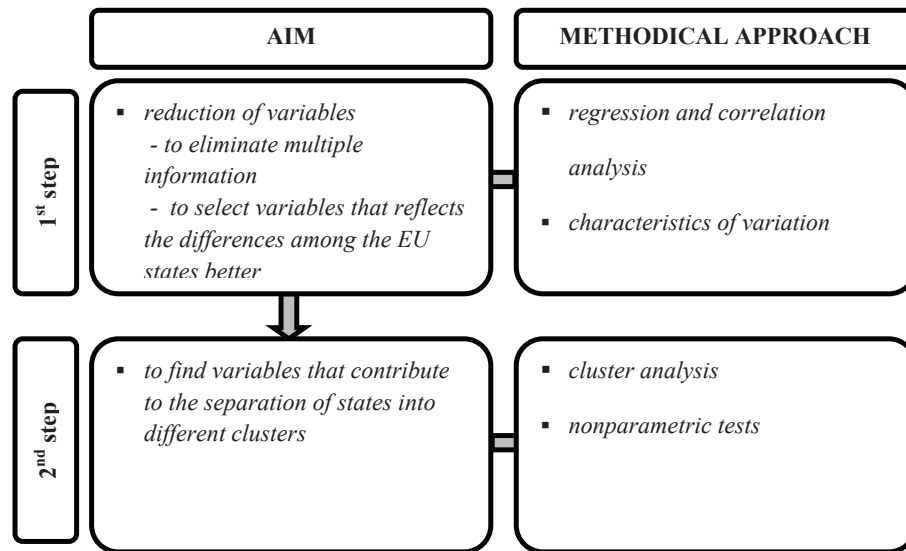
Data source	Proportion of variables from total 114 (%)
Eurostat database	78.2%
European Quality of Life Survey	9.6%
World Bank database	8.7%
Eurobarometer survey	3.5%

Source: own working

II: Proportion of primary set of variables in the dimensions

Dimension	Number of variables	Proportion (%)
Material living standards	22	19.3%
Health	13	11.4%
Education	12	10.5%
Personal activities	19	16.7%
Political voice and governance	12	10.5%
Social connections and relationships	6	5.3%
Environment	10	8.8%
Insecurity	20	17.5%

Source: Commission on the Measurement of Economic Performance and Social Progress, 2008; own working



1: Variables selection for quality of life evaluation – two-steps reduction process
Source: own working

procedures) are the basic for variables selection in many works (e. g. Legatum Institute, 2010, Hlavsa, 2010). *Unsupervised methods*, such as factor analysis, aim to reduce the dimension of a set of variables without target variable. On the other hand by the use of *supervised methods*, such as logistic regression, it is possible to reduce the dataset in relation to target variable (e. g. gross domestic product, happiness). In case of this study, the primary set of variables was proposed on the basis of experts's recommendation. The following steps are based on objective methods. The process of reduction was divided in two consecutive steps of different aims as summed in Fig. 1.

1st step

The statistical procedures in both first and second steps are based on three-years averages (2007–2009). The first step aims to reduce number of indicators to eliminate high correlation among the variables. Pairs of variables with absolute value of correlation coefficient higher than 0.8 were further investigated. On the basis of coefficient of variation computed as $V = s/\bar{x}$, where s is the standard deviation and \bar{x} is the arithmetic mean, variable of higher variation was selected. For the purpose of this step the pairs of variables were sorted descending following the correlation coefficient. In case of extremely high or low values for given indicator, the coefficient of variation was computed without the extreme values to ensure that the coefficient reflects the overall differences among the EU states.

2nd step

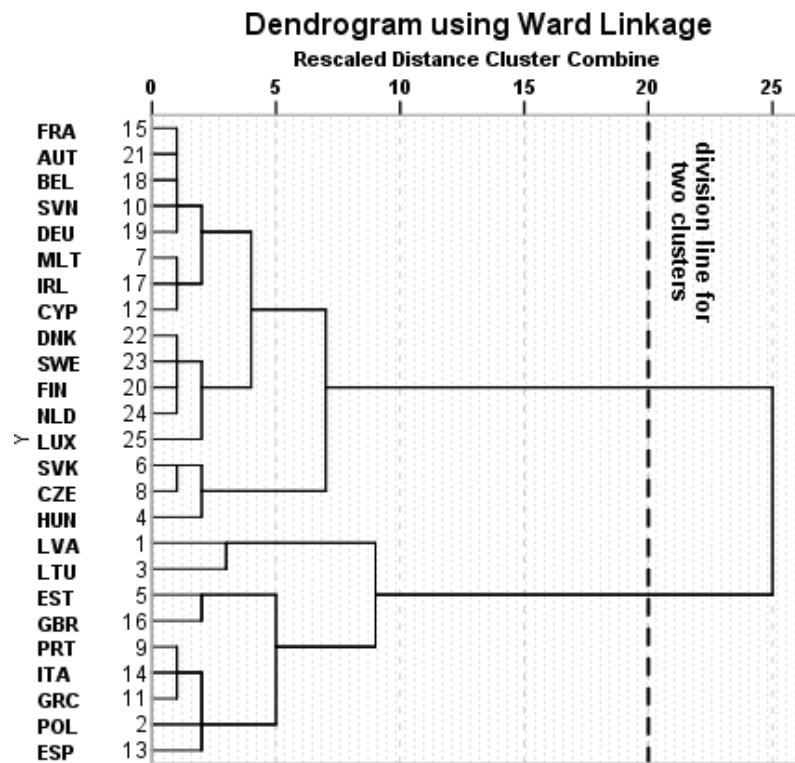
The second step of reduction process aims to find variables that reflect the differences across the European Union the best. There are more or less different groups of states within the European from the view of various dimensions. To identify

groups of more or less unique states, hierarchical cluster analysis was used. Cluster analysis, as well as previous analysis, was performed separately for each dimension on the basis of all variables in each dimension after the first step of reduction. The hierarchical cluster analysis was performed on the basis of all variables in each dimension after the first step of reduction. Each variable was expressed as *z-scores* to ensure the comparability of different indicators. *Euclidean squared distance* was used to quantify the distance between units and between the clusters, states were grouped on the basis of *Ward's method* (for details see e.g. Hebák, 2007).

Two or three obvious groups of states were found for each of the eight dimensions. The criterion for number of clusters was based on further analysis. Each cluster represented one group in hypothesis testing, so no clusters of less than five states were taken into consideration. To identify indicators that are of significantly different level in one cluster compared to another, nonparametric test procedures were used because of limited cluster sizes. Mann-Whitney U-test was used to test the differences between the groups of states in case of two clusters, Kruskal-Wallis test in case of three clusters. Multiple comparison between pairs of clusters were examined using the procedure introduced by Dunn (1964). For all tests performed, 5% level of significance was used. The results are discussed in next chapter.

RESULTS

After the first step of reduction, 71 from the original set of 114 variables were selected. High correlation was often identified between pairs of identical variables referring separately to population of men and women. High correlations were identified between various monetary indicators



2: Dendrogram for dimension *Material living standards*
Source: own working

or poverty indicators in the *Material living standards* dimension as well. In dimension *Health*, correlations between various prevalence of chronic diseases and life expectancy or healthy life years were found. On the other hand no correlation above was identified in dimension *Social connections and relationships*, covered by subjective indicators from Eurofound survey. No correlations above were found in *Environment* dimension also, so the original set of ten variables remain the same for this dimension.

In the second step of reduction, contribution of each variable on the clustering was evaluated. The suggested methodological approach is introduced

for two dimensions from the total of eight. Other dimensions are not discussed in details, all results can be found in Pacáková (2012).

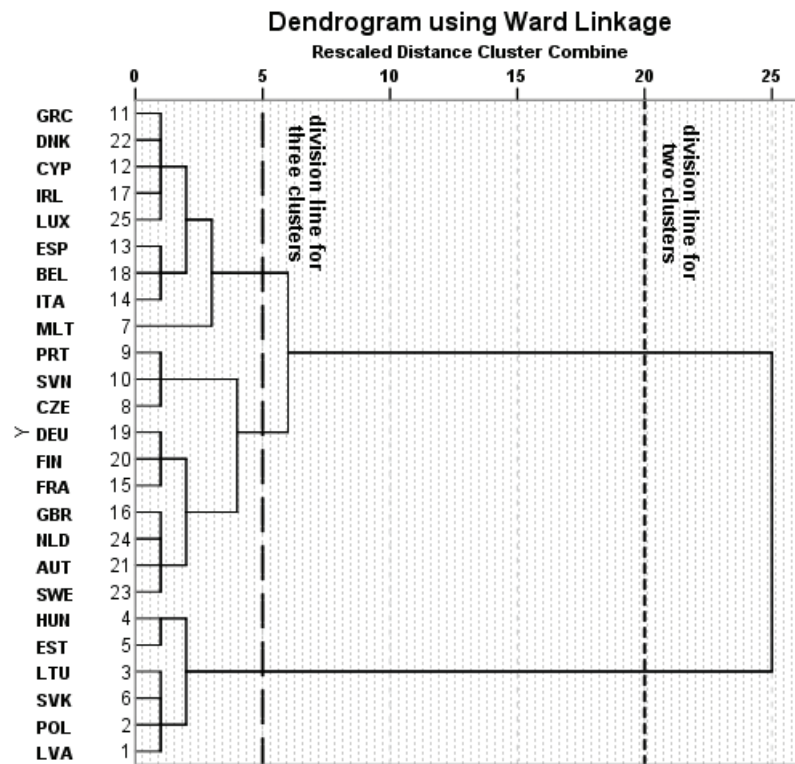
For dimension *Material living standards* two groups of states with obviously different levels were found. The first cluster consists of nine states while the second cluster groups sixteen states together. No further division was taken into consideration (see Fig. 2).

The nine states in the first cluster (Latvia, Lithuania, Estonia, Portugal, Italy, Greece and Poland, Great Britain and Spain) were of worse level of most of the characteristics in this dimension

III: Results of tests for the evaluation of differences between clusters from the view of different indicators – dimension *Material living standards*

Indicator	Test criterion	p-value
Dispersion of regional GDP	56.5	0.693
Bank nonperforming loans to total gross loans	15.0	0.004
Household saving rate	42.5	0.215
In-work at-risk-of-poverty rate	1.0	< 0.001
People at-risk-of-poverty after social transfers	5.0	< 0.001
Net national income at market prices	24.0	0.018
People with unmet needs for medical examination, reason: too expensive	19.5	0.008
Relative median at-risk-of-poverty gap	0.0	< 0.001
Material deprivation rate	24.0	0.018
Level of Internet access - households	15.0	0.004
Inflation rate	47.0	0.333

Source: own working



3: Dendrogram for dimension Health
Source: own working

(higher poverty rate, deeper GDP regional differences, lower household savings).

Mann-Whitney U test was performed to evaluate level of each indicator in the first cluster compared to the second cluster. The eight variables which are of significantly different level in the two clusters (Tab. III) were selected for evaluation of quality of life.

On the basis of results of the cluster analysis for the dimension *Health*, two or three clusters in different distance can be assumed as it is shown in the Fig. 3.

Five of six variables are of significantly different levels in case of either two or three clusters. There is one variable only (*People having a long-standing*

illness or health problem, female) which is of significant differences in case of three groups of states only (Tab. IV).

Level of this variable was found different in two of three pairs of clusters. It can be assumed that the variable is one of important determinants for dividing the biggest cluster (nineteen states) into two smaller clusters. For the reasons stated above, the indicator was found a determinant that reflects the differences among EU states and no reduction was performed in dimension *Health* on the basis of results obtained in the second step.

Cluster analysis for dimension *Education* enables to identify two clusters, one of them consists of five southern states: Portugal, Spain, Italy, Greece and

IV: Results of tests for the evaluation of differences between clusters from the view of different indicators – dimension Health

Indicator	Mann-Whitney U test (2 clusters)		Kruskal-Wallis test (3 clusters)		
	Test criterion	p-value	Test criterion	d.f.	p-value
Self-perceived health: bad and very bad, females	6.0	0.001	10.71	2	0.005
Life expectancy at age 65, females	2.0	< 0.001	12.46	2	0.002
Infant mortality	6.0	0.001	10.54	2	0.005
Healthy life years at birth, males	2.0	< 0.001	15.09	2	0.001
Death rate due to chronic diseases, males	0.0	< 0.001	13.92	2	0.001
People having a long-standing illness or health problem, males	43.5	0.390	15.06	2	0.001

Source: own working

Malta. On the basis of hypothesis testing, five of six variables were selected in this dimension.

Dimension *Personal activities including work* was covered by fourteen indicators (after the first phase of reduction). It was possible to divide the states into two or three clusters on the basis of hierarchical cluster analysis. In case of three clusters it is obvious that these are of different characteristics from the view of most variables. In case of dividing the states into two groups only, one of the clusters consists of states which are of very different characteristics. That is why three clusters were assumed and all the variables which are of different level following Kruskal-Wallis test were selected. For the purpose of next analysis twelve of fourteen variables were selected.

Dimension *Political voice and governance* was described by a set of five variables after the first step of reduction. Two groups of states of different characteristics were proposed on the basis of cluster analysis. Four of five variables were of different levels in the clusters following the results of Mann-Whitney U test. The final set of indicators connected with political situation consists of four variables.

Three clusters can be identified for the dimension *Social connections and relationships*. Following the results of analysis comparing level of each indicator among the three clusters, five of six variables were selected in the second step. Variable characterizing *percentage of people who claims that they have nobody to ask for help in case of illness* is significantly different in one of free pairs only and therefore was not selected for further analysis.

On the basis of suggested approach for the reduction of variables, number of indicators in dimension *Environment* was reduced the most. Five of ten variables were selected for further analysis. The original set of ten variables was of very low correlations between the variables, so no reduction was done in the first step. Combining results of cluster analysis with nonparametric tests allowed to find out that some of the variables are of low differences among the EU states. These variables were not used for further analysis of quality of life.

In case of dimension *Insecurity*, states were divided into two clusters of fourteen and eleven states. Southern states together with the Central European states and Latvia, Lithuania and Estonia are the members of the first cluster while the second cluster groups northern states mostly. Seven of thirteen variables are of significantly different level in the clusters.

After the two-step reduction process, quality of life was characterized by a set of 52 indicators covering all eight dimension proposed by the Commission – see Tab. V.

The list of variables is introduced in Tab. VI separately for each dimension.

CONCLUSION

The paper aims to propose a set of indicators for quality of life evaluation following recommendations of the Commission on the Measurement of Economic Performance and Social Progress. Part of the aim was to propose a methodological approach of variables selection. The selection procedures can be cut into three phases. The first phase was based on the recommendations of the Commission and on the results of previous studies. Therefore this phase was based on subjective methods of variables selection mostly. The primary data set consists of 114 variables drawn from 4 different data sources. Two consecutive steps of variables reduction followed, each one was of different aim. Due to high correlation among the variables, the first step aims to eliminate correlation. On the basis of correlation analysis pairs of high correlation were identified. Variable with higher variation among the states was selected for further analysis. By the first step forty-three variables were eliminated. The second step aims to select variables that reflect the different quality of life level among the European Union states the best. To answer the question, variables that are drivers of clustering of the states into groups of states of different characteristics were selected.

The final data set consists of 52 variables covering all eight dimensions proposed by the Commission. The evaluation of quality of life uses both subjective

V: Proportion of variables in the dimensions after the two-step reduction process

Dimension	after the 1 st step		after the 2 nd step	
	Number of variables	Proportion (%)	Number of variables	Proportion (%)
Material living standards	11	15.5%	8	15.4%
Health	6	8.5%	6	11.5%
Education	6	8.5%	5	9.6%
Personal activities	14	19.7%	12	23.1%
Political voice and governance	5	7.0%	4	7.7%
Social connections and relationships	6	8.5%	5	9.6%
Environment	10	14.0%	5	9.6%
Insecurity	13	18.3%	7	13.5%
Total	71	100%	52	100%

Source: own working

and objective indicators. The proposed data set was used to evaluate quality of life among the European Union states consisting of more unique groups of states.

VI: *Proposed set of indicators for quality of life evaluation in the European Union*

Dimension	Indicators
Material living standards	<ul style="list-style-type: none"> ● Bank nonperforming loans to total gross loans ● In-work at-risk-of-poverty rate ● People at-risk-of-poverty after social transfers ● Net national income at market prices ● People with unmet needs for medical examination, reason: too expensive ● Relative median at-risk-of-poverty gap ● Material deprivation rate ● Level of Internet access - households
Health	<ul style="list-style-type: none"> ● Self-perceived health: bad and very bad, females ● Life expectancy at age 65, females ● Infant mortality ● Healthy life years at birth, males ● Death rate due to chronic diseases, males ● People having a long-standing illness or health problem, males
Education	<ul style="list-style-type: none"> ● Life-long learning, females ● Persons with low educational attainment, from 25 to 64 years ● Unemployment rate, tertiary education, females ● Unemployment rate, tertiary education, males ● Tertiary educational attainment, males, age group 30–34 years
Personal activities	<ul style="list-style-type: none"> ● Part-time workers in % of total employment, females ● Average number of usual weekly hours of work, males ● Average number of usual weekly hours of work, females ● Unemployment rate, males ● Unemployment rate, females ● Employment rate, males ● Employment rate, females ● Involuntary part-time employment as percentage of the total part-time employment ● Labour productivity ● Employees with tertiary education, males ● Subjective evaluation of time spent in hobbies: too little ● Subjective evaluation of time spent in work: too much
Political voice and governance	<ul style="list-style-type: none"> ● Trust in national government ● Control of corruption ● Political stability and absence of violence/terrorism ● Proportion of seats held by women in national parliaments
Social connections and relationships	<ul style="list-style-type: none"> ● Face-to-face contact with family: at least once a week ● Friends outside the family: too little ● Nobody to ask for advice ● Contact with family: just right ● Contact with family: too little
Environment	<ul style="list-style-type: none"> ● Energy intensity of the economy ● Greenhouse gas emissions ● Population suffering from noise from neighbours or from the street ● Urban population ● Standardised death rate - cancer
Insecurity	<ul style="list-style-type: none"> ● Long-term unemployment ● Healthy life years at 65, females ● Unemployment rate, age group less than 25 years ● Median equivalised net income, females aged 65 and over ● Material deprivation rate, age group 65 and over ● Percentage of people who claim that it is very likely that they lose job, males ● Percentage of people who claim that it is very likely that they lose job, females

Source: own working

SUMMARY

The objective of the article is to propose a set of indicators for quality of life evaluation in the European Union following the recommendations introduced by the Commission on the Measurement of Economic Performance and Social Progress. Part of the main objective is to propose a methodological tool for variables selection. At the beginning, set of 114 variables was completed as a result of findings listed by the Commission and based on empirical literature. The primary data set was reduced in two consecutive steps. The first step aims to reduce correlation among the variables. As a result of the second steps, variables reflecting the different levels among consistent groups of states were selected. Therefore, the methodical approach in the second step was based on nonparametric procedures used to identify variables that are of significantly different level in clusters identified by cluster analysis. Each step was performed separately for each of eight dimensions proposed by the Commission. As a result set of 52 variables for quality of life evaluation drawn from 4 different data sources was introduced. The data set combines both objective and subjective variables, is based on open data sources only and it reflects the needs for measuring the quality of life among the European Union states.

REFERENCES

- COMMISSION ON THE MEASUREMENT OF ECONOMIC PERFORMANCE AND SOCIAL PROGRESS, 2008: *Report by the Commission on the Measurement of Economic Performance and Social Progress* [online]. [Accessed 2011-02-12]. URL: <http://www.stiglitz-senfitoussi.fr/documents/rapport_anglais.pdf>.
- DUNN, O. J., 1964.: Multiple Comparisons Using Rank Sums. *Technometrics*, 6: 241–241.
- GÖNNER, C. et al., 2007: Capturing Nested Spheres of Poverty. A Model for Multidimensional Poverty Analysis and Monitoring. *CIFOR*, 46. ISBN 978-979-1412-04-9.
- HLAVSA, T., 2010: Výběr ukazatelů pro hodnocení ekonomiky regionů. *Ekonomická revue*, 13: 145–151. ISSN 1212-3951.
- HEBÁK, P. et al., 2007: *Vícerozměrné statistické metody* (3). Praha: Informatorium, 239 p. ISBN 978-80-7333-001-9.
- KUPROVÁ, L. and KAMENICKÝ, J., 2006: Multikriteriální postavní krajů ČR v letech 2000 až 2004. *Statistika*, 4: 303–315.
- LEGATUM INSTITUTE, 2010: *Legatum Prosperity Index Report* [online]. [Accessed 2011-06-06]. URL: <<http://www.prosperity.com/downloads/2010ProsperityIndexFullReport.pdf>>.
- NOORBAKSH, G., 1998: A Modified Human Development Index. *World Development*, 26, 3: 517–528.
- OECD, 2011: *Compendium of OECD well-being indicators* [online]. [Accessed 2011-06-20]. URL: <http://www.oecd.org/document/28/0,3746,en_2649_201185_47916764_1_1_1_1,00.html>.
- OSBERG, L., SHARPE, A., 2002: An index of economic well-being for selected OECD countries. *Review of Income and Wealth*, 48, 3: 291–316.
- OSBERG, L., SHARPE, A., 2010: How are we doing? The Index of Economic Well-Being. *Challenge*, 53: 25–42.
- PACÁKOVÁ, Z., 2012: *Metodologické nástroje hodnocení kvality života v zemích EU*. Dissertation thesis, Czech University of Life Sciences, Prague. 288 s.
- Data sources
- EUROSTAT ONLINE DATABASE [online database]. Luxembourg: Eurostat. [Accessed 2011-06-20 till 2011-11-20]. URL: <http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database>.
- WORLD DATABANK [online database]. Washington: World Bank. [Accessed 2011-06-20 till 2011-11-20]. URL: <<http://databank.worldbank.org/ddp/home.do>>.
- EUROBAROMETER INTERACTIVE SEARCH SYSTEM [online database]. Brusel: European Commission. [Accessed 2011-06-20 till 2011-11-20]. URL: <http://ec.europa.eu/public_opinion/cf/index_en.cfm>.
- EUROPEAN FOUNDATION FOR THE IMPROVEMENT OF LIVING AND WORKING CONDITIONS, 2006: European Quality of Life Survey [computer file]. Colchester, Essex: UK Data Archive [distributor], February 2006. SN: 5260, <http://dx.doi.org/>.
- EUROPEAN FOUNDATION FOR THE IMPROVEMENT OF LIVING AND WORKING CONDITIONS, 2009: European Quality of Life Survey, 2007 [computer file]. Colchester, Essex: UK Data Archive [distributor], October 2009. SN: 6299, <http://dx.doi.org/>.

Address

Ing. Zuzana Pacáková, Ph.D., Katedra statistiky, Provozně ekonomická fakulta, Česká zemědělská univerzita v Praze, Kamýcká 129, 165 21 Praha 6 – Suchbát, Česká republika, email: pacakova@pef.czu.cz