

## FROM THE EDITOR

A set of twelve articles included in this issue is structured conventionally in three major sections. It is appended by short reports from two international conferences – a selection of papers from these conferences constitutes the last section of this issue.

The first section – sampling methods and estimation – contains three articles. It starts with **G. N. Singh's, Mukti Khetan's, Shweta Maurya's** paper *Some Effective Estimation Procedures Under Non-Response in Two-Phase Successive Sampling*. The authors discuss an import issue of how to assess the effect of non-response in estimation of the current population mean in two-phase successive sampling on two occasions. They use the sub-sampling technique of non-respondents and propose exponential methods of estimation under two-phase successive sampling arrangement. Properties of the proposed estimation procedures have been examined along with some suggestions on estimation procedures for survey practitioners concerning the use of auxiliary information (in the form of exponential methods of estimation).

The next paper, *Transmuted Kumaraswamy Distribution* by **Muhammad Shuaib Khan, Robert King, Irene Lena Hudson** is devoted to one of the most widely applied statistical distribution in hydrological problems and many natural phenomena. The authors propose a generalization of the Kumaraswamy distribution referred to as the transmuted Kumaraswamy (*TKw*) distribution using the quadratic rank transmutation map (studied by Shaw et al., 2009). They provide a comprehensive account of the mathematical properties of the new distribution with specific expressions for the moments, moment generating function, entropy, mean deviation, Bonferroni and Lorenz curves, and formulated moments for order statistics. The *TKw* distribution parameters are estimated by using the method of maximum likelihood, and Monte Carlo simulation is performed in order to investigate the performance of MLEs – the usefulness of the proposed model is illustrated using the flood and HIV/AIDS data. In conclusions, the authors indicate on the better performance of the model (in terms of a better fit) than the *Kw* distribution.

**Ranjita Pandey's** and **Anoop Chaturvedi's** paper *Bayesian Inference for State Space Model with Panel Data* explores panel data set-up in a Bayesian state space model. The authors apply the conditional posterior densities of parameters to determine the marginal posterior densities using the Gibbs sampler. They believe that the theoretical framework they developed is generally more effective and useful for applied researchers and practitioners, especially in terms of a more precise panel data-based prediction.

The research paper section starts with an article by **Mauro Mussini**, *On Measuring Income Polarization: An Approach Based on Regression Trees* in which the application of regression trees for analysing income polarization is proposed. Using an approach to polarization based on the analysis of variance, the author shows how the regression trees can uncover groups of homogeneous income receivers in a data-driven way. Since the regression tree can deal with nonlinear relationships between income and the characteristics of income receivers, it can detect which characteristics and their interactions actually play a role in explaining income polarization. In consequence, the author believes the regression tree is a useful flexible statistical tool to explore whether income receivers concentrate around local poles. Some interesting partition of income receivers is demonstrated for the case of Italy. For instance, an empirical analysis of Italian income data shows that the interactions among employment status, educational qualification and age form well-identified groups of income receivers, whereas the other characteristics do not play a clear role in explaining income polarization.

**Oleksandr Osaulenko's** article *Quality of Life and Poverty in Ukraine – Preliminary Assessment Based on the Subjective Well-Being Indicators* provides a first-hand account of the research on the topic in Ukraine. It starts with an overview of the database and methodology, followed by main results of the quality of life and poverty research conducted by the national statistical office. A system of subjective well-being indicators is based on self-evaluation of the attained level of well-being, of the level of meeting the basic living needs and the levels of deprivation of consumption. In addition, methodological approaches to analyzing economic and infrastructure deprivation (due to the geographic limitations of services accessibility) are briefly described. The paper reviews the factors that underlie the Ukrainian list of deprivations and define the percentage of population that is particularly affected by multiple deprivation. The data covers a period of several years allowing for exploration of the distribution of deprivations by different population groups.

In the last article of this section, *On the Relationships Between Smart Growth and Cohesion Indicators in the EU Countries* by **Beata Bal-Domańska, Elżbieta Sobczak**, the problem of evaluation of the relationships between smart growth and economic and social cohesion factors is discussed from the perspective of the Europe 2020 strategy's objectives toward improving the situation in education, digital society and research and innovation. The authors employ aggregate measures for these three phenomena based on panel data models. Social cohesion is described by the level of employment rate as one of the conditions essential to the well-being and prosperity of individuals, and economic cohesion is defined by the level of GDP per capita in PPS. The study covered the group of 27 European Union countries during the period of 2002-2011. Some of the conclusions have policy implications. For instance, it was found that an increase in the employment rate was related to the increasing role of employment in smart specialization sectors.

The conference papers' section is opened by ***Heteroscedastic Discriminant Analysis Combined with Feature Selection for Credit Scoring*** in which **Katarzyna Stapor, Tomasz Smolarczyk, Piotr Fabian** propose an approach for building a credit scoring model based on the combination of heteroscedastic extension of classical Fisher Linear Discriminant Analysis and a feature selection algorithm that retains sufficient information for classification purpose. Starting with an observation that credit granting is a fundamental and one of the most complex question being faced by credit institutions, the authors attempt to develop an effective classification model that would be helpful for managers. To this aim, they focused on the feature selection algorithm that retains sufficient information for classification purpose and tested five feature subset selection algorithms: two filters and three wrappers. In order to evaluate the accuracy of the proposed credit scoring model and to compare it with the existing approaches, the German credit data set was used (Chen, Li, 2010). In the conclusions they found that the proposed hybrid approach is an effective and promising method for building credit scoring models. The analysis results in better prediction accuracy, also due to applying variable importance analysis for identifying the most relevant variables for the classification purpose.

In **Iwona Skrodzka's** article ***Knowledge-Based Economy in the European Union – Cross-Country Analysis*** spatial differences in the level of development of the knowledge-based economy in the European Union countries are discussed using a soft modelling approach. The estimation of a synthetic measure of KBE, as well as the arrangement and classification of the UE-27 countries into typological groups for the years 2000 and 2013 are provided. For instance, the highest level of development of the knowledge-based economy was observed for Sweden, Denmark, Finland and Luxembourg, whereas the lowest one for Greece, Bulgaria and Romania. Eleven countries, including Poland, improved their ranking in 2013 compared to 2000, while nine countries reduced their positions (the highest increase was in Hungary, while the largest fall in Italy in 2013).

In the next article, ***New Method of Variable Selection for Binary Data Cluster Analysis*** by **Jerzy Korzeniewski** the problem of the level of measurement in the variable selection procedures is discussed, with intention of improving the efficiency of the existing methods, with special reference to the marketing type data. The author proposes that a variable selection method be based on connecting the filtering of the input set of all variables with grouping of sets of variables similar with respect to analogous groupings of objects. The new method allows for linking good features of two entirely different approaches to variable selection in cluster analysis, i.e. *filtering* methods and *wrapper* methods. The proposed method of variable selection yields best results when the classical *k*-means method of objects grouping is slightly modified.

In **Dominik Krężolek's** paper, ***The GlueVar Risk Measure and Investor's Attitudes to Risk – An Application to the Non-Ferrous Metals Market***, the issue of risk in economic investment decisions is discussed and a new risk measure is proposed – the GlueVaR risk measure. It can be defined as a linear combination of VaR and GlueVaR and is aimed at helping to deal with uncertainty and

volatility that is characteristic to the economic investment decisions. The most commonly used risk measure, Value-at-Risk, suffers from a significant drawback, which is the lack of subadditivity, but is crucial in terms of portfolio diversification. The proposed GlueVaR measure allows for calculating the level of investment loss depending on investment's attitudes to risk while meeting the needed requirement, therefore it may be used in portfolio risk assessment. The application of the GlueVaR risk measure is presented for the non-ferrous metals market. Compared to classical measures, the most useful feature of the proposed new risk measures is that for a particular investor it is possible to implicitly define the set of adverse events and determine the importance of such events.

**Marta Hozer-Koćmiel, Christian Lis** present the results of *Examining Similarities in Time Allocation Amongst European Countries*. Time allocation has been defined as the daily distribution of time to various activities. Professional work time, domestic work time and leisure time appear to be the most important for the economic approach. It has been proved that there are coherent groups of countries with similar structure of time allocation. The taxonomic methods used in order to verify the thesis included: cluster analysis, k-means method, generalised distance measure GDM and interval taxonomic method TMI. The analysis was performed on the basis of HETUS survey data. In conclusions, two groups of countries that show strong similarity in setting the time budget of the population have emerged from the analysis: (i) the 'new' European countries (that has undergone economic transformation), characterized by a distinctly longer professional work time and shorter leisure time; and (ii) the Scandinavian countries and the more developed countries of Western Europe, where the basic variables of time allocation were opposite, i.e., relatively short professional work time and long leisure time.

**Joanna M. Landmesser's** paper, *Decomposition of Differences in Income Distributions Using Quantile Regression*, deals with microeconomic techniques useful for the study of differences between groups of objects. Using the Machado-Mata quantile regression approach, the empirical decomposition of the inequalities in income distributions of one-person households in urban and rural areas was performed using data from the Household Budget Survey for Poland in 2012. It was found that the tendency towards increased income inequalities between urban and rural residents when moving to the right of the income distribution can be observed. The rural residents are at a disadvantage. The decomposition of the inequalities revealed a growing share of the part explained by different characteristics of people (especially in educational level), and a declining share of the unexplained part, associated with the evaluation of those characteristics.

**Włodzimierz Okrasa**

Editor