

Józef Dziechciarz (ed.): *ECONOMETRICS. EKONOMETRIA. ADVANCES IN APPLIED DATA ANALYSIS*. WUE 2018, vol. 22, no 3.

This volume of the journal contains seven articles devoted to various important aspects of the economy in Poland and in Europe.

The article by Maria M. Grzelak, Iwona Laskowska and Elzbieta Roszko-Wójtowicz brings up the subject of the influence of the research and development (R&D) expenditures on the competitiveness of manufacturing enterprises in Poland. These expenditures finance innovative activities which are considered as those with major impact on the competitiveness of both the country and the enterprises. The authors report in brief on the results of econometric modelling for 22 manufacturing divisions based on the panel data for the period of 2009–2016. They also compare estimation results of the fixed effects model (FEM) with those of the random effects model (REM). During the chosen data period we could observe a very dynamic increase in internal expenditure on R&D from PLN 9.07 billion in 2009 to maximum PLN 18.07 billion in 2015 and only slightly less in 2016. Three characteristics describing the competitiveness of manufacturing divisions were chosen as endogenous variables: goods sold, gross value added and labour productivity. The Cobb-Douglas production function was chosen with employment, investment expenditures and R&D expenditures – current and lagged by one year. The FEM models were finally interpreted by the authors – according to the Hausman test results the null hypothesis that individual effects for manufacturing divisions were independent of explanatory variables was rejected for both models. The main result is that a 1% increase of current R&D expenditures cause around 0.037% (0.047%) increase in goods sold (gross value added) and respectively, a 1% increase of expenditures in the previous year was responsible for around 0.044% (0.052%) increase in sold production (gross value added). Labour productivity was defined by the authors as a ratio of the gross value added and employment in the manufacturing divisions. The REM version of the model proved to be superior. The estimation results presented confirm the positive impact of R&D expenditures on productivity – a 1% increase in current R&D expenditures may cause a 0.04% increase in labour productivity, while for lagged R&D – approximately a 0.05% increase. The presented results are interesting, based on the standard Cobb-Douglas production function. It would be informative to give the reader additional idea about the strength of the possible collinearities among the explanatory variables which might affect the undervaluation of estimation errors.

The second article by Artur Mikulec and Małgorzata Misztal touches on the problem of small firms survival. Data were collected for firms in the counties of Łódzkie voivodship. The 21,163 firms established in 2010 were observed – 44.3% of them went into liquidation by the end of 2015. The authors distinguish two factors hypothetically determining the duration of the firm – type of business activity and location. The Kaplan-Meier estimation, the Cox proportional-hazards

regression model as well as the regression tree approach were used to answer the question: does the type of business activity and location of the enterprise affect its duration? The structure of the sample showed that more than 36 percent of firms started in 2010 their activity in the city of Łódź. The authors summarize in brief the methodology of the analysis and the results. The proposed algorithm to build survival tree is CTree (Conditional Inference Tree). Prediction error curves and the bootstrap cross-validation were used to describe the precision of the predicted survival probabilities measured by the Brier score. The results are richly illustrated by informative graphs. The authors also discuss qualitative problems corresponding to the economic activity of small firms. For instance, after two years firms are often liquidated because of the changing conditions of social security payments or the lack of subsidies granted to new entrepreneurs. The Kaplan-Meier survival curves distinguish the following sectors: “human health and social work activities”, “real estate activities” and “professional, scientific and technical activities” as dominating from the point of view of the probabilities of survival and such sectors as: “agriculture, forestry and fishing”, “arts, entertainment and recreation” and “financial and insurance activities” as being dominated. Statistical analysis confirmed the hypothesis that “the higher level of development of a given county and higher level of urbanization translate into the higher probability of surviving”. The authors provide us with an extensive bibliography – methodological and empirical. It would be of great importance to present the results of the spatial analysis for the Łódź voivodship on the maps with some additional information about e.g. population density, unemployment rate, etc.

The article by Witold Bartkiewicz and Zbigniew Gontar “Cloud computing migration barriers and advantage in manufacturing – an analysis of ambiguity and dependencies in the assessment criteria” takes up the problem of managers’ decision regarding the adoption of cloud computing. The authors questioned IT managers from 400 enterprises in Poland who implemented the cloud computing solutions about their opinion on the advantages and barriers of the process. They present the results of the respondents’ selection of pre-proposed answers as a multiple query. The survey results are then analysed using similarity matrices. Hierarchical agglomerative clustering algorithm was applied to discover clusters of decision criteria according to the level of similarity coefficients. Based on the survey results, the authors illustrate binary association rules – this part of the article is, in my opinion *‘art for art’s sake’* – to obtain final clusters first of all one has to decide on the direction of the implication between categories. It seems to be too subjective and the authors did not prove that the results are not sensitive to the results of this first step. Nevertheless, the article is an interesting contribution to the discussion on decision-making criteria and measuring relationships among them.

The article by Justyna Brzezińska “Latent class analysis in the evaluation of items in survey research” presents the elements of item response theory (IRT) and its application. In the first part of the work the historical background of the theory

is presented. The theory is applied, for an example, to the results of the recent survey of 242 students, who were to answer 26 questions about their attitudes towards their professional future choosing “Yes” or “No”. Based on the correlation coefficients between the pairs of answers, the two pairs of questions with the highest correlation were chosen. The log-linear modelling was applied to analyse structures of dependency in the thus obtained contingency table – the author presented likelihood ratio statistics, Akaike information criteria and Bayesian information criteria values for chosen models. The simplest, and at the same time, providing sufficient values of information criteria was the one which related quality of life, success in life, upgrading competencies and future studies. Latent class analysis was then conducted – the respondents were classified finally into three latent classes – each class was characterized according to the most frequent answer to the discriminated four questions. The article is interesting from the methodological point of view but, in my opinion, the description and the interpretation of the results are too modest compared with the large dose of history and definitions of measures.

The article by Katarzyna Cymbranowicz “The ‘working poor’ phenomenon in Europe – a taxonomic analysis” offers a discerning survey of literature and official statistics concerning the problem that people who are working can still suffer from poverty. The author presents various definitions of ‘working poor’. In general, the problem is identified as the result of labour market disfunction, job instability as well as the household structure of the person working. Among main indicators of ‘working poverty’ based on Eurostat, we find: at-risk-of-poverty rate, severe material deprivation, people living in households with very low work intensity. Statistical data for 28 European countries were collected from Eurostat database. The data period of 2005–2016 was divided into three four-year sub-periods. The values of diagnostic variables were tabulated but not commented on by the author. The data were then standardized. The Ward agglomeration method based on the Euclidean distance between the EU-28 Member States was applied to distinguish groups of countries with a similar level of ‘working poverty’ in sub-periods. The author discusses the content of clusters obtained using the Ward method. It would be worthy to complement the analysis by a comment on the changes observed in sub-periods, at least for Poland. The article contains a large number of bibliographic positions.

The next article, by Marcin Pełka, “Analysis of innovations in the European Union via ensemble symbolic density clustering” proposes another look at the classification methods and rankings of countries and regions of EU from the point of view of ability and strength of innovations creation. The author presents some distance measures for symbolic interval-valued data available in R software. DBSCAN (density-based clustering) algorithm was used – we can find a summarizing description of the methodology in the article; 18 symbolic interval-valued variables measuring, e.g. level of education, scientific publications, expenditures on various type of innovations, employment and exports in high-tech

manufacturing companies and knowledge-intensive services, sales of innovations were taken into account. The research covered 220 European regions in 2017. Regional symbolic data were then aggregated on a country level. Clustering algorithm was applied with different initial parameters and different distance measures. Ensemble clustering based on co-association matrix was used to improve the accuracy and stability of the classification algorithms' results. Rousseeuw's Silhouette internal cluster quality index was the optimization criterion for the number of clusters and the adjusted Rand index was used to validate the stability of the clustering procedure. Four clusters of countries were finally identified: innovative leaders (Belgium, Denmark, Germany, Ireland, France, Netherlands, Austria, Slovenia, Finland, Sweden, United Kingdom, Norway), mid-low innovative countries (Bulgaria, Croatia, Hungary, Poland, Romania, Slovakia), moderate innovators (Spain, Italy, Czech Republic), and the fourth cluster containing Greece and Portugal – also moderate innovators. The description of the clustering methodology and research conducted is quite clear and detailed enough.

The article by Włodzimierz Kołodziejczak "Criteria used in LFS to identify the population's economic activity status vs. the respondents' subjective views: implications for unemployment research" concerns the analysis of various indicators of the labour market based on the official statistics and also on the subjective opinion of respondents to the Labour Force Survey (LFS) on their own economic activity status: employed, unemployed or professionally inactive. Among the indicators the author chose: labour force participation rate, employment rate, and unemployment rate. There exists a problem with defining the term used by the Polish Central Statistical Office in BAEL surveys: "economically active population" – it is understood as the entire population aged 15 and above. The EU Labour Force Survey uses the term "active and inactive population 15 years and above" in the same sense. The author summarizes in the article the methodology of the surveys used by the EU-LFS and the method of measuring flow rates between employment, unemployment and inactivity. Several interesting characteristics of the Polish population's economic activity in 2016 are presented: unemployment outflow rate, unemployment-to-employment flow rate, unemployment inflow rate, average duration of unemployment, employment rate, economic activity rate, unemployment rate, equilibrium employment rate. The estimates differ depending on the choice of objective or subjective approach. The author presents the results for the whole population and distributed with sex and urban/rural criterion. In general, the subjective opinion is that the unemployment rate exceeds the value presented in the official statistics, especially in rural areas (the difference in the estimates is greater than 4 p.p.). The article underlines the need to identify the reasons behind the found differences in the estimation results. Such information should be an important guideline for labour and economic policy.

*Iwona Konarzewska, University of Łódź*