Development of Population in Predominantly Rural Regions of the Czech Republic and Slovakia

Ondřej Šimpach[1] and Marie Pechrová[2]

[1] University of Economics Prague, Faculty of Informatics and Statistics, Department of Statistics and Probability, W. Churchill sq. 4, 130 67 Prague 3, Czech Republic
[2] Institute of Agricultural Economics and Information, Modelling of the Impacts of Agricultural Policy, Mánesova 1453/75, 120 00 Prague 2, Czech Republic

[1]ondrej.simpach@vse.cz, [2]pechrova.marie@uzei.cz

Abstract: The aim of the paper is to assess the population development in predominantly rural regions of the Czech Republic and Slovakia and to project further trends. Rural areas in both countries underwent relatively similar development and currently suffer by the changes in agriculture. The demographic data and statistics about NUTS 3 regions for period 2002–2014 were obtained from Eurostat database. We consider the predominantly rural regions as defined by the Eurostat. There are 6 those regions in the CR and 4 in SR.

Firstly, a hierarchical cluster analysis was used to group the regions according to the number of live birth, deaths, net migration and total population change. Three groups emerged where Prešovský (alternated by Plzeňský) outstand in separate group due to its high number of inhabitants and birth rates. First cluster had worse development conditions (number of deaths higher than live births, emigration higher than immigration, and hence negative total population change during the whole observed period) than the second cluster, where the population development was more positive. The affiliation of regions to particular group was changing during the period, but Jihočeský region stayed always in the first cluster.

Secondly, cohort-component method was used to project the population development in five-year intervals from 2014 to 2029. The population projection shows that the population bases are downsizing due to generation ageing and increasing average age of economically productive persons in observed predominantly rural regions.

Key words: cluster analysis, population projections, rural development, Czech Republic, Slovakia

JEL classification: J11, J16, Q01

1 Introduction

Rural areas in both countries underwent relatively similar development and currently suffer by the changes in agriculture. "Over the past decades, major changes have taken place in Europe's rural areas. These changes include contrasting developments like depopulation and land abandonment in some regions, and urbanisation and agricultural intensification in others." (Westhoek et al., 2006) In the strategic policy documents of the Czech Republic are highlighted especially population ageing, unfavourable age structure, flee of the young, educated inhabitants and competent entrepreneurs, and low potential of economic diversification as the main threats. According to the Rural Development Programme (RDP, Ministry of Agriculture, 2013), the major problem faced by rural areas is the stabilization of the rural population as it is aging more rapidly than the rest of the country because of the emigration of young people to urban centres to obtain jobs and better social infrastructure. Agriculture – with its relatively low share in the total employment (11 % in rural areas and 3.8 % overall) – has a limited capacity to reverse this trend (Tvrdon, 2011). In many parts of rural Europe the main cause of ageing and depopulation is selective out-migration of young adults from remote and economically weak regions to prosperous urban areas (Leibert and Wiest, 2011). This brain-drain may lead to the situation when the rural regions are lagging behind the developed. Therefore, European Union implemented special policies, programmes and measures to support the growth of the lagging regions. However,

"Understanding the drivers of regional growth in an ex-post integrated economic environment is an important assignment for both theory and policy," (Petrakos et al., 2011).

For example Becker et al. (2010) observed the development of average annual growth of GDP per capita at purchasing power parity and average annual employment growth at NUTS 2 and 3 levels during a programming period and tried to assess the effects of EU's structural policy (particularly of the Objective 1 facilitating convergence and cohesion within the EU regions). They found that "Objective 1 treatment status does not cause immediate effects but takes, in the average programming period and region, at least four years to display growth effects on GDP per capita". (Becker et al., 2010) Pechrová and Šimpach (2013) grouped the regions in the EU according to their development potential which was described by four categories of indicators: economic and social area, equipment and agricultural characteristics.

Agarwal et al. (2009) were searching for key factors which enhance the competitiveness of the rural areas in UK and came to the conclusion that the essential development drivers are economic and human capital. Dufek and Minařík (2010) also utilized the indicator of human resources in the regions of the CR as one of the most important factor for their development potential. They used indicators from: demographic area, economic and employment field, standards of living and social level, and education, health and criminality. The demographic development was assessed by population density, share of rural inhabitants, age structure, age index, life expectancy at birth, live born and death on 1 000 inhabitants, and migration increment on 1 000 inhabitants. Urbancová and Hlavsa (2014) examined the usage of Age Management measures (which may help to improve the situation on the labour market and encourage young people to work in the agriculture) by Czech agrarian businesses and found that the farmers were not quite familiar with those practices.

The aim of the paper is to assess the population development in predominantly rural regions of the CR and Slovakia and to project further trends. Future state of the population is calculated based on knowledge of the previous state of population and prediction of the number of deaths and born, and the number of immigrants and emigrants from particular regions. The paper is structured as follows: First, there are introduced the data from Eurostat, the methods of cluster analysis and demographic projection. Then the next section presents and discusses the results. Last section concludes.

2 Materials and Methods

The data about the fertility, mortality and migration for NUTS 3 regions for period 2002–2014 were obtained from Eurostat database (Eurostat, 2016). We consider the predominantly rural regions as defined by the Eurostat, i.e. the population which is living outside the urban areas is rural. Urban areas are determined as territories where the density of population is above 300 inhabitants/km^2 applied to grid cell and a minimum size threshold (5 000 inhabitants) applied to grouped grid cells above the density threshold. (DG Agri, JRC, Eurostat, 2012) There are 6 those regions in the CR and 4 in SR. We observed age-and-sex specific number of live births, deaths, net migration (statistically adjusted), total population on 1 January, and total population change in each region during the period 2002–2014 and described the development. Based on selected uncorrelated indicators, the regions were clustered into 3 groups according to observe the changes of their character.

The clustering procedure forms hierarchical groups of mutually exclusive subsets, each of which has members (regions) that are maximally similar with respect to demographic indicators. "Given n sets, this procedure permits their reduction to $n - 1$ mutually exclusive sets by considering the union of all possible $n(n - 1)/2$ pairs and selecting a union having a maximal value for the functional relation, or objective function, that reflects the criterion chosen by the investigator" (Ward, 1963). This procedure is repeated until there is only one group. We used Ward's method that merges

the clusters with minimal within-cluster sum of squared deviations from objects to centroids. The distances were measured by Squared Euclidean distance (*d*). The distance between two data points (X_i and Y_i) is calculated as the the sum of the squares of the differences between corresponding values (1).

$$d = \sum_{i=1}^{n} (X_i - Y_i)^2 \quad (1)$$

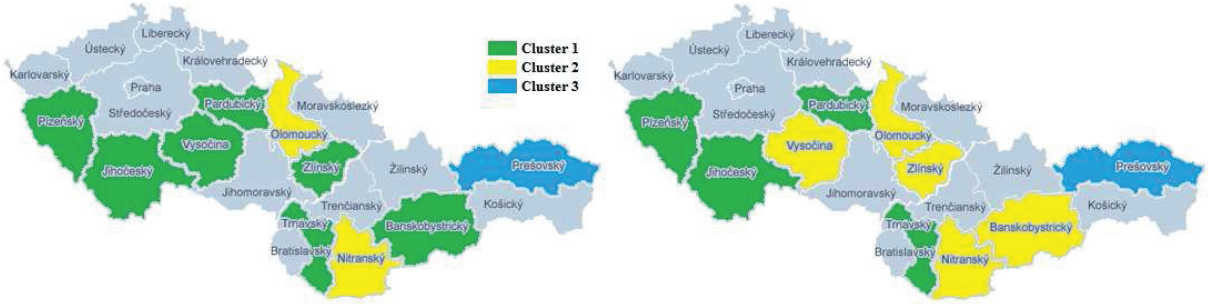
Ward's method tends to create relatively small clusters because of the squared differences, but with similar numbers of observations. Its disadvantage is that the distance between clusters calculated at one step of clustering is dependent on the distance calculated in previous step. It is also sensitive to outliers. The clustering process was stopped when three groups of regions were created. Calculations were done in Stata 11.2.

For the projection of future state of population a component method based on several assumptions is used. Cohort-component method which moves age-and-sex specific numbers of persons of particular sex (components) based on set scenarios (for detailed description of the modelling process see Šimpach (2015)). We set the scenarios of development of fertility rate, decrease of death rate and assume zero net migration balance in the calculations. Consequently were calculated the future age-and-sex specific death rates $m_{x,t}$, (where t denotes time and x age of the person) and number of live born persons N^v_t. Number of immigrants I_t and emigrants E_t might have changed, but the total net migration was supposed to be constant as there is a problem with its prediction. Data were known for time period from 2002 until 2014, and the prediction was done for five year intervals (until 2029). The projection horizon will cover two EU's programme and planning periods (2014–2020, 2021–2026). During calculating the regional population projections is in the case of the mortality scenario considered the coefficient of decline in the probability of death $q_{x,t}$. From the studies of e.g. Šimpach and Dotlačilová (2015) it is possible to overtake the coefficient in the range of 0.96 to 0.98 for every 5 years, differing by gender. Based on this assumption we gain the scenario of the development of age-and-sex-specific table numbers of survivors $l_{x,t}$, which serves not only to calculate other derived statistics of mortality tables ($d_{x,t}$, $L_{x,t}$, $T_{x,t}$ – see Šimpach and Dotlačilová, 2015), but also for prediction of future development of life expectancy at birth $e_{0,t}$, and also for the projection coefficients $P_{x,t}$. The level of total fertility rate (*tfr*) is considered for each analysed region to be the same as at the beginning of the projection, and is divided by the actual age-specific fertility rates of a particular region.

3 Results and Discussion

First, the correlation between chosen demographic indicators was assessed and only those where the paired correlation coefficient was lower than 0.8 were used – particularly: the number of live births, deaths, net migration and total population change. Hierarchical cluster analysis formed three clusters of regions with different characteristics. Almost each year the affiliation of particular region changed (i.e. it belonged to another group). Only exception is Jihočeský region, which stayed in the first cluster during whole period. Fig. 1 presents the map of the first and the last year of the observed period.

Fig 1. Clusters of rural regions in 2002 (left) and 2014 (right)



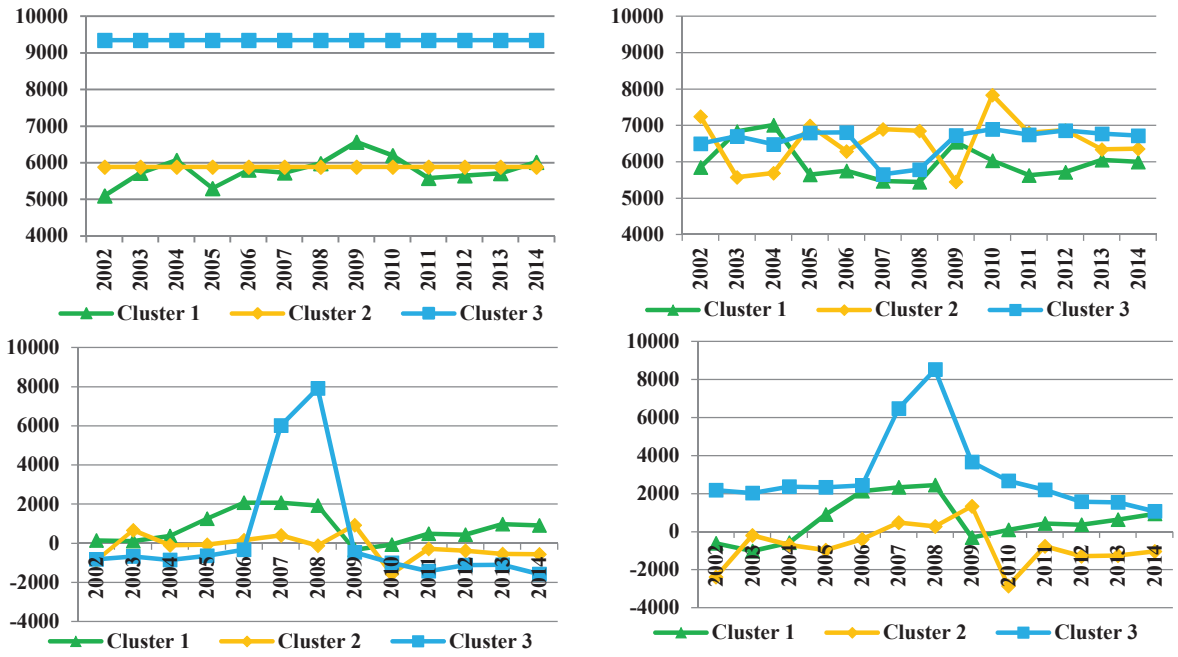
Source: own elaboration, 2016

Between those years Vysočina, Zlínský and Banskobystrický regions moved from first cluster to the second as their development situation slightly worsened. Prešovský region stay in the same group because it is characterized by the highest birth rate and the highest number of inhabitants from all regions. On the other hand, population in Vysočina was the lowest and further decreased which also contribute to the change of other determinants of its affiliation to particular cluster. Net migration and total population change was on average negative in 3 Czech and 2 Slovak regions. The highest positive population change was noted in Prešovský region and then in Jihočeský. On the other hand, total population change showed negative development in Nitranský, Zlínský, Banskobystrický, Olomoucký and Vysočina regions.

Clusters' characteristics are displayed in Fig 2. Cluster 3 is typical with the highest birth rate from all clusters of almost whole observed period. The number of deaths is lower than number of births which imply positive population change. It contains only Prešovský region with exception 2007, 2008 where there is Plzeňský region instead. The development of deaths is more similar to other groups, again with the exception of the years 2007 and 2008. Since 2012, the most people died in this cluster, but on the other hand, the most people were born. Net migration was highly positive in year 2007 and 2008 due to the presence of Plzeňský region instead of Prešovský. Plzeňský region belonged to the one with positive migration balance for almost whole period and hence it was included in the first cluster. Only in 2010 the number of outgoing was higher than incoming. Development in other clusters is various.

Second cluster might be characterized by higher death rate than birth rate (with exception of 2003, 2004 and 2009), by negative migration balance (number of emigrants was on average higher than number of immigrants for the whole period with exception of four years) and by negative total population change. This may negatively influence its development potential. In majority of years to this category belonged Olomoucký, Nitranský, and Zlínský regions. For those regions is typical very negative net population change and negative net migration. First cluster, contrary to the others, has the best potential of positive future development.

Fig. 2. Demographic characteristics of clusters: live births - total (top left), deaths – total (top right), net migration pus statistical adjustment (down left), total population change (down right)



Source: own elaboration, 2016

Number of deaths was lower than number of birth in period 2006 to 2010 and in year 2014. The number of live births was the most volatile from all clusters as same as the number of deaths. Total population change and migration balance were mainly positive or close to zero. It contained Jihočeský region for all periods, than Pardubický and Trnavský for 9 years.

Second, year 2014 was taken as a base for the projection of the state of population in next 15 years. Tabled numbers of surviving $l_{x,t}$ were estimated based on coefficient of decrease of probability of death $q_{x,t}$ (which was chosen to be the same for all ages, but various for gender in the range of 0.97–0.98 in order to ensure that during 20 years the life expectancy at birth would increase by approximately 1 year (it differs among gender and regions)). Life expectancy at birth divided according to the gender in each region is displayed in Table 1. It is higher for females than males. The most positive was for Vysočina region, where the females can live in 2029 up to 83.2 years and males to 77.2. Life expectancy of males was the highest in Plzeňský region in all periods. On the other hand, Nitranský and Banskobystrický regions report the lowest life expectancy in many periods.

Table 1. Projections of life expectancy at birth of males (M) and females (F) in rural regions of CR and SK

Year	2014		2019		2024		2029	
Region	M	F	M	F	M	F	M	F
Jihočeský	75.7	81.5	76.0	81.8	76.3	82.0	76.7	82.3
Plzeňský	76.3	81.4	76.6	81.7	76.9	81.9	77.3	82.1
Pardubický	74.0	80.0	74.4	80.3	74.7	80.6	75.1	80.8
Vysočina	76.2	82.5	76.5	82.8	76.9	83.0	77.2	83.2
Olomoucký	75.2	81.9	75.6	82.2	75.9	82.4	76.2	82.7
Zlínský	75.3	81.3	75.7	81.6	76.0	81.9	76.3	82.1
Trnavský	73.2	80.0	73.6	80.2	74.0	80.5	74.3	80.8
Nitranský	72.4	79.3	72.8	79.6	73.1	79.9	73.5	80.2
Banskobystrický	72.7	79.8	73.1	80.1	73.4	80.4	73.8	80.6
Prešovský	73.3	80.2	73.7	80.5	74.1	80.7	74.4	81.0

Source: own calculations, 2016

Total fertility rate is also assumed to stay the same for the whole projected period, but different for various regions. The number of children is decreasing as the fertility is lower than 2.08. Reproduction period of females was considered already from 10–14 years (because in Slovak regions these events occurred even in such a low age groups) to 50–54 years. It was set that there is 51.5% probability that a new born child is a boy and 48.5% that it is a girl. Numbers of live born boys and girls are presented in Table 2.

Table 2. Projections of number of live born boys (B) and girls (G) in rural regions of CR and SK

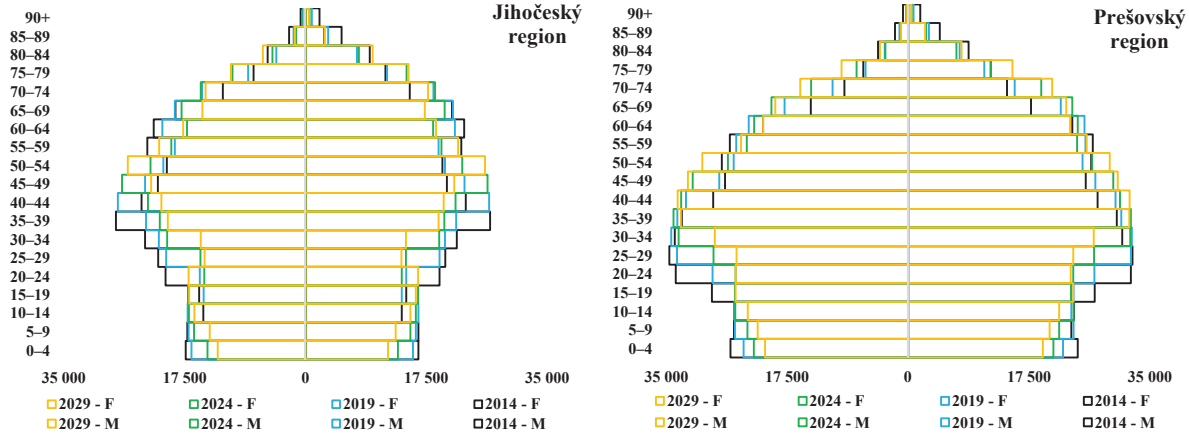
Year	2014–2018		2019–2023		2024–2028		2029–2034	
Region	B	G	B	G	B	G	B	G
Jihočeský	15 924	14 997	14 227	13 398	12 727	11 986	6 034	5 682
Plzeňský	13 866	13 058	12 079	11 375	10 636	10 016	5 821	5 482
Pardubický	16 354	15 401	14 687	13 832	13 107	12 344	6 198	5 837
Vysočina	13 445	12 662	12 311	11 594	10 997	10 357	5 175	4 874
Olomoucký	15 894	14 968	14 210	13 382	12 580	11 847	5 914	5 570
Zlínský	14 010	13 194	12 621	11 886	11 095	10 448	5 180	4 878
Trnavský	13 193	12 424	11 482	10 813	9 785	9 215	4 509	4 246
Nitranský	14 919	14 050	13 128	12 363	11 308	10 649	5 230	4 925
Banskobystrický	15 056	14 179	13 499	12 712	11 983	11 285	5 647	5 318
Prešovský	23 797	22 411	22 559	21 245	20 903	19 685	10 029	9 445

Source: own calculations, 2016

Finally, the population projections for years 2014, 2019, 2024 and 2029 are elaborated. From population trees at Fig. 3 it is possible to see that the population bases are downsizing due to generation ageing and increasing average age of economically productive persons in observed predominantly rural regions. In Jihočeský region in 2014, the most of males was in age group between 35–39 (27 thous.); in 2019 in category 40–44 (27 thous.); in 2024 in category 49–50 (27 thous.); and finally in 2029, the majority will continuously move to category 50–54 years (25 thous.). The same situation is with females. There are 26 thous. Of them in category 35–39

and almost the same number of them move to the group 50–54 in 2029. Number of young people will continuously decrease and therefore there will be also less people in the highest category of 90+. Similar situation is in Plzeňský region, but the number of males in 2014 starts at 25 thous. and ends at 24 thous. Number of females is also lower (25 thous. at the beginning and 24 thous. At the end). Also Pardubický, Vysočina, Olomoucký, and Zlínský regions follow the same pattern. Surprisingly the development of Slovakian Trnavský region was similar to Czech regions. However, other regions are more diverse. In Nitranský and Banskobystrický regions in 2029, the most males will be in the age from 40–49 as same as in 2024. Prešovský region is special as the population there is the youngest. The most males (35 thous.) in 2014 are in age 25–29. In 2029, the majority (33 thous.) continuously moves to category 40–44. Similar pattern is also for females, but there is less of them (32 thous.).

Fig. 3. Population projections in selected predominantly rural regions 2014, 2019, 2024, 2029



Source: own elaboration, 2016

Our results can be compared with those of other researchers or institutions. For example CZSO does extrapolation of population development and recalculates it every five years according to the current conditions and recent statistics on population development. From other institutions that prepare population projections including those of the Czech and Slovak Republic can be named the Eurostat or OECD. CZSO's and Eurostat's results are relatively similar. OECD studies tend to have divergent results because of different methodology. Our projections consider the population in regions as same as prediction of CZSO which project the population in regions up to 2050 (CZSO, 2014), but it is revised it only every four years.

4 Conclusion

The aim of the paper was to assess the population development in 6 Czech and 4 Slovakian predominantly rural regions and to project further trends based on the data for period 2002–2014 from Eurostat database. First, a hierarchical cluster analysis grouped the regions according to their similarity in each year based on demographic indicators: the number of live birth, deaths, net migration and total population change. Ward's method was used to create three clusters. Second, cohort-component method with zero migration balance assumption was used to forecast the population development in five-year intervals from 2014 to 2029.

Slovakian Prešovský region (alternated by Czech Plzeňský region) outstand in separate cluster number 3 due to its high number of inhabitants and birth rates. However, future development of this region is expected to adjust to the trends in other regions and the population will get older (majority of people will move from age category 25–29 to 35–39, 40–44). Cluster 2 had worse development conditions (number of deaths higher than live births, emigration higher

than emigration, and negative total population change during the whole observed period) than the Cluster 1 whose conditions were the best. The affiliation of regions to groups was changing during the period. The exception is Jihočeský region which stayed in Cluster 1 for all period. Population projection shows that the population bases are downsizing due to generation ageing and increasing average age of economically productive persons in all predominantly rural regions. Almost in all regions, the major share of population will move from age group 35–39 in 2014 to 50–54 in 2029. In future research it is planned to include other EU regions, project their development and cluster them for policy-making purposes.

Acknowledgements

The authors gratefully acknowledge to the Czech Science Foundation project No. P402/12/G097 DYME – "Dynamic Models in Economics" for supporting this paper.

References

Agarwal, S., and Errington, S. R. (2009) 'Measuring the determinants of relative economic performance of rural areas', *Journal of Rural Studies*, vol. 25, pp. 309 – 321, ISSN 0743-0167

Becker, S. O., Egger, P. H., and von Ehrlich, M. (2010) 'Going NUTS: The effect of EU Structural Funds on regional performance', *Journal of Public Economics*, vol. 94, pp. 578 – 590, ISSN 0047-2727

CZSO, Czech Statistical Office, Projekce obyvatelstva v krajích ČR - do roku 2050, Prague 2014, [on-line], Available: https://www.czso.cz/csu/czso/projekce-obyvatelstva-v-krajich-cr-do-roku-2050-ua08v25hx9 [Assessed 25 May 2016]

Dufek, J. and Minařík, B. (2010) 'Lidské zdroje jako významný faktor rozvojového potenciálu krajů České republiky', *Proceedings of RELIK 2010*, Prague, pp. 1 – 10

Eurostat, European Statistical Office, Population change - Demographic balance and crude rates at regional level (NUTS 3) [demo_r_gind3], Brussels, 2016 [Online], Available: http://appsso.eurostat.ec.europa.eu/nui/show.do, [Accessed: 23 Mar. 2016]

Eurostat, European Statistical Office, Urban-rural typology, Brussels, 2013, [Online], Available: http://ec.europa.eu/eurostat/statistics-explained/index.php/Urban-rural_typology, [Accessed: 20 Mar. 2016] ISSN 2443-8219

Leibert, T. and Wiest, K. (2016) 'The interplay of gender and migration in Europe's remote and economically weak rural regions', *Journal of Rural Studies*, vol. 43, pp. 261–266, ISSN 0743-0167 http://dx.doi.org/10.1016/j.jrurstud.2016.01.007

Pechrová, M. and Šimpach, O. (2013) 'The Development Potential of the Regions of the EU', *Region in the Development of Society*, Brno, pp. 322 – 335, ISBN 978-80-7375-884-4

Petrakos, G., Kallioras, D., and Anagnostou A. (2011) 'Regional convergence and growth in Europe: understanding patterns and determinants', *European Urban and Regional Studies*, vol. 19, pp. 375 – 391, ISSN 0969-7764

Šimpach, O. (2015) 'Experts in the field of education and other potential knowledge makers in 2051', *Efficiency and Responsibility in Education 2015 (ERiE)*. Prague, pp. 499 – 506, ISBN 978-80-213-2560-9

Šimpach, O. and Dotlačilová, P. (2015) 'The Impact of Smoothing of Age-specific Death Rates by the Gompertz-Makeham Function on the Results of Stochastic Projections by Lee-Carter Model', *Proceedings ITISE 2015*. Granada, pp. 363 – 373, ISBN 978-84-16292-20-2

Tvrdon, M. (2011) 'EU Rural Development Policy: the Case of the Czech Republic', *Proceedings of the 5th International Scientific Conference Rural Development*, Kauno, vol. 5, no. 1, pp. 254 – 260

Urbancová, H. and Hlavsa, T. (2014) 'Age Management Principles in Czech Agrarian Sector', *AGRIS on-line*, vol. VI, no. 3, pp. 93 – 102, ISSN 1804-1930

Ward, J. H. (1963) 'Hierarchical grouping to optimize an objective function', *Journal of the American Statistical Association*, vol. 58, pp. 236 – 244, ISSN 1537-274X

Westhoek, H. J., van den Berg M. and Bakkesa, J. A. (2006) 'Scenario development to explore the future of Europe's rural areas', *Agriculture, Ecosystems & Environment*, vol. 114, no. 1, pp. 7 – 20, ISSN 0167-8809, http://dx.doi.org/10.1016/j.agee.2005.11.005