Charles University in Prague

Faculty of Science

Department of social geography and regional development

SPATIAL ASPECTS OF SOCIAL INEQUALITY: CZECHIA IN THE CONTEXT OF CENTRAL-EAST EUROPEAN COUNTRIES

Dissertation thesis summary

Vojtěch Nosek

Prague 2010

thesis supervised by: J. Novotný, Ph.D.

Main parts of the thesis:

Introduction

Theoretical context

Methods of quantification of inequalities and data base

“Theoretical results” and recapitulation of the methodological part of the thesis

Empirical results

Conclusions

Content of the thesis summary

1. Introduction, main goals and structure

Inequalities in society are one of the basic topics in social sciences. In Czechia, the interest in social inequality was enhanced by the change of societal system in 1989. Regional inequality research in Czechia as well as in other countries focuses mainly on assessing economic performance of regions and on related socio-economic variables. Systematic study of spatial distribution of other social variables with attention to their local variability which we try to analyze in this thesis (according to available data) is not very common at least among Czech authors. The most common are empirical analyses on administrative (or other regional) levels. However, analyses on these levels do not avoid an ecological error which is often seen as a pitfall of a regional analysis. Study of spatial inequalities at sub-regional levels in Czechia as a whole is not very common (see Blažek a Netrdová 2009 or Ouředníček et al. 2010).

Other possibilities open up thanks to a growing number of new methods and software as well as to still more accessible data which are integrated into geographical information systems (GIS). As an example we can mention spatial statistic methods and their integration into GIS environment (Anselin 1995, Rey a Janikas 2005). However, this trend brings not only opportunities but also difficulties. Seemingly very easily processed data and application of sophisticated (statistical) methods may lead to their inappropriate utilizations and even to wrong interpretations. Harvey (1969, p. 11) claims, that application of these methods may suppress the ability of speculation, intuition and geographical imagination. The solution is, in our opinion, not to ignore these methods but to strive to utilize them appropriately and interpret correctly. Such an interest in statistical methods is quite common among many researchers abroad (see Rey 2001 or Shorrocks and Wan 2005) but relatively unknown and for many maybe surprisingly undervalued in Czechia.

The majority of authors who try to quantify regional inequalities are satisfied with the mere value of the variability according to selected variables. Yet it is also interesting to study the relation of regional inequality to overall interpersonal inequality (or other sub-regional inequality) and estimate the share of a regional level on the overall social inequality. In other words, regional variables are not considered as a population but rather as a sample (though usually not independent) of a population of individuals (or other sub-regional entities). If we assess, for example, income inequality, we can estimate what part of the overall population income inequality may be explained by differences between regions – i.e. by the place of residence of individuals or households. Analysing these spatial aspects of social inequalities and attempt to their systematic research is in Czech geographical studies absent (see only Netrdová and Nosek 2009 or Novotný 2004). If there is a research studying spatial aspects of social inequality (such as Rey 2001), it has usually very partial character.

The goal of the methodological part of the thesis is to examine possibilities of chosen methods in the context of their utilization for studying spatial aspects of social inequality. These methods were already introduced in Spurná (2008) and Netrdová and Nosek (2009). We try to monitor the methods' behaviour under different conditions such as the number of studied units (regions), overall population variability etc. It is also highlighted that these methods may effectively complement each other. One chapter is also dedicated to testing of a statistical significance of results.

The main goal of the empirical part, and the central goal of the whole thesis, is to analyze the significance of spatial component of social inequality and examine its forms in detail. The analysis focuses on Czechia and some other central-east European countries (Slovakia, Poland, and Austria). We try to find regularities such as the relation between the complexity of studied variable and the level of observed inequality (and its spatial aspects). Thus we work with spatially very detailed data which are aggregated into several levels. A partial goal is also to compare the spatial dimension of social inequality with other dimensions (age, national, educational, sectoral etc.). There is a need for an individual data for this type of analysis. Thus it is limited only to Czech census 2001.

The dissertation thesis is structured into three main sections: theoretical, methodological, and empirical. After a short introduction, the chapter no. 2 named "Theoretical context" follows. In this chapter we discuss the very nature of geography as a scientific discipline and a position of quantitatively oriented research. In the next part, there are different approaches to geographical inequality and its measuring introduced. The concept of relative importance of spatial/regional inequality on the overall inequality and the concept of spatial autocorrelation which dominate in all analyses are highlighted there. We also explain the essential terminology. Subsequently we discuss some general regularities one can encounter when studying inequalities. The last part of the theoretical section (chapter 2.4) consists of literature summary dealing with inequalities in Czechia and other analyzed countries. On the basis of this literature we derive hypotheses which are tried to be verified during the empirical part of the thesis.

The third section is devoted to methodical and methodological issues. In chapter no. 3.1 and 3.2 are introduced methods suitable to fulfil the main goals. Two chosen methods, Theil index decomposition and spatial autocorrelation, are studied in detail other methods are mentioned very briefly. In chapter 3.3 we discuss possibilities of testing the statistical significance of results which is often being omitted from similar empirical studies. At the end of the third section (chapter 3.4) we introduce the data base and the whole region with which we work in the empirical parts of the thesis.

The chapter no. 4 focuses on detailed research of applied methods. We try to measure the influence of various conditions on the results of the Theil index decomposition and the spatial autocorrelation. Thanks to simulation experiments in chapter no. 4.1 we can subsequently more effectively interpret the empirical results. The most important findings from the theoretical and methodological section of the thesis are summarized in subchapter no. 4.2.

After the methodically and methodologically oriented section the presentation of empirical results follows. Firstly, in chapter 5.1, spatial aspects of social inequalities in Czechia are assessed – in detail using census data 1991 and 2001 and then through available more up-to-date data. In chapter no. 5.1.2 we try to evaluate the differences in results under alternative delimitation of regions and in chapter no. 5.1.4 we analyze individual data from census 2001 and compare spatial dimension with other (non-spatial) dimensions of social inequalities. In chapter no. 5.2 are results from Czechia compared with result from other countries. The main findings of the whole thesis are summarized in the concluding chapter no. 6.

2. Methodical notes

Inequalities are (in the context of this thesis) understood only as a statistical distribution of particular data, normative and other subjective approaches not being reflected. We deal only with social inequality (i.e. characteristics of population). Due to the fact that individual data are scarcely available individuals are often substituted by groups which may be constituted in various ways. Since the dissertation thesis has been written at the geographical department we focus predominantly on spatial aspects of social inequality. In this case the individuals are split into groups according to the geographical location of their residence, mainly into regions.

An approach to inequality measuring is also influenced by the phenomenon one observes. In this respect, we can distinguish between inequalities in opportunities and inequality in outcomes (see for example Nunez a Tartakowsky 2007). Inequalities in opportunities are often connected with different chances to a quality of life (such as differences in basic social, economic, and political rights and liberties, access to education or health care etc.) and also depending on race, sex, place of origin or the wealth of parents (Bourguignon et al. 2003). Although these approaches can never be fully distinguished we can say that in this thesis we consider only inequalities in outcomes.

Even if we disregard the difference between inequalities in opportunities and inequalities in outcomes, it is necessary to define inequality "among what" will be measured. The choice of appropriate variable depends on goals of a concrete research. When working only with social inequalities we can split variables into two groups: socio-demographic (such as marriage rate, age preference index, educational index etc.) and socio-economic (unemployment rate, entrepreneurial activity, employment in agriculture etc.). It is obvious that this division is very rough and none of analysed variables can be clearly assigned to one particular group. In the context of our analyses it is important to stress the difference in complexity of those two categories. In this respect, we derive differences between variables from the proportion of internal and external factors on which these variables depend. As more complex variables we consider (corresponding with Hampl 1998, p. 38) those variables which have relatively more important external factors (external environment influencing their differentiation) in comparison with internal factors. Socio-economic variables (such as unemployment rate) can be thus considered as more complex than socio-demographic ones (such as age preference index).

When quantifying spatial aspects of social inequality we applied predominantly two different methods. We tried not only to describe those methods but also to analyze the influence of different initial conditions (such as the number of units or regions under analysis, the overall variability etc.) on the results.

First possibility how to quantify spatial dimension of social inequality is by using the decomposition to between-group and within-group component of variability (groups may be defined by spatial units). Decomposition without a residuum can be achieved by measures from the group of generalized entropy measures. Inequality decomposition using generalized entropy measures is popular among many authors (see fro example Bourguignon 1979, Litchfield 1999, Shorrocks 1984). In this thesis we work only with the Theil index decomposition.

In order to assess spatial aspects of social inequality one can apply also methods of spatial autocorrelation. These methods were primarily developed for a different purpose. Nevertheless, they can be used as an appropriate tool in the context of the goals set in this thesis as well. The basic idea of

spatial autocorrelation is the similarity of spatial units which is examined from a perspective of their mutual distance and is based on a relative continuity of socio-geographic phenomena in space (Nezdařilová 1984). Spatial autocorrelation is measured by various autocorrelation statistics which describe an overall degree of spatial autocorrelation. The most often statistic is the Moran's I criterion (later denoted as MI, Cliff and Ord 1973).

Moran's I may be considered as a 'global statistic' because one resulting value indicates the level of spatial autocorrelation (or clustering) in the whole area under analysis. However, there are also local variants of spatial autocorrelation – the most common is probably the LISA analysis developed by Anselin (1995). Thanks to LISA analysis we can categorize observed variables into five groups. First two groups consist of above average/below average units surrounded by above average/below average units (type high-high and low-low). Other two types represent above average/below average units surrounded by below average/above average units (type high-low and low-high). The last type represents units with non-significant spatial autocorrelation. An important facet of LISA analysis is a possibility to depict this categorization in a map. This map can help with localization of formerly discovered spatial aspects of inequality.

An important part of the methodological section of the thesis deals with testing the statistical significance of results which is often being forgotten. We test the difference between empirical results and situation when all data are placed randomly. In the case of the international comparison the results are corrected by subtracting this random (stochastic) component of measured inequality.

3. Selected conclusions

There are two types of conclusions in the dissertation thesis. First, we were studying methods suitable for quantification of spatial aspects of social inequality, especially the Theil index decomposition and spatial autocorrelation. The importance of initial conditions on results was studied through experimental simulations. In the second part of the thesis we applied these methods and tried to quantify spatial aspects of social inequalities in Czechia and several neighbouring countries.

We can say that the Theil index decomposition and spatial autocorrelation were found to be suitable methods for measuring spatial aspects of social inequalities. In comparison with commonly used aggregate measures of inequality these methods quantify inequality in a different way and bring additional and practically valuable information. While standard error, coefficient of variation or the Gini coefficient can quantify only the aggregate value of regional (or other) inequality, the Theil index decomposition can express the relative share of respective regional levels (as well as respective regions or their groups) on the overall inequality. Spatial autocorrelation can enrich these decompositions with the value of 'clustering' in space. When compared with the decompositions, quantification of spatial autocorrelation does not depend on a delimitation of regions. These clusters can be then categorized and expressed in a map (LISA). As was hopefully clear from the theoretically-methodological part of the thesis, it is favourable to combine these methods and interpret them jointly. Variables can be categorized according to spatial and regional aspects of social inequalities.

We have also shown that when analyzing geographical inequalities it might be useful to distinguish between stochastic and contextual component of final results. Stochastic component can be expressed as the part of observed inequality which may be explained statistically – for example the inequality

one can expect when all data are randomly placed. In the context of this thesis, this situation was named as a null model. Null model was subsequently compared with the 'real' measured inequality in order to assess its (statistical) significance. Similar test was used for the Theil index decomposition but it might be also used for other aggregate measures of inequality (such as for the Gini coefficient).

Thanks to experimental simulations (as well as to some theoretical arguments) we can say that the values of overall inequality (T), regional inequality (T_B) and relative importance of regional inequality (T_B/T) are influenced by the number and size of regions of analyzed system. For a better illustration let's imagine two hypothetical countries which differ in their size (the country A is bigger than the country B). Simulations showed that: (i) in the bigger country A one can find higher overall inequality T (in other words the overall inequality positively depends on the size of studied system). (ii) If these countries are split into regions with same sizes (which means that in the country A will be more regions than in the country B), in the country with more regions one can observe higher regional inequality T_B and higher value of T_B/T (in other words, the measures of regional inequalities positively depend on the number of regions). (iii) However, if we split these countries into the same number of regions (which means that in bigger country A will be regions bigger when compared with country B), we will find lower values of both T_B and T_B/T (in other words, the measures of regional inequalities negatively depend on sizes of regions).

The values T_B and T_B/T also positively depend on the level of overall variability. It means that even when data are placed randomly in space higher levels of regional inequalities are to be found in societies that have higher overall inequality).

The values of Moran's I are, on the other hand, independent on mentioned factors. This finding might seem trivial in the case of the influence of the number and size of regions because the measurement of MI is not based on a delimitation of regions. Less intuitive is the Moran's I neutrality according to the size of studied system (number of units under analysis) and the level of the overall population variability.

Answering theoretically-methodological questions was a prerequisite to the main part of the thesis which was focused on analyses of the importance of spatial aspects of social inequalities in Czechia and other countries.

In Czechia we confirmed the expectation of increasing of regional inequalities (T_B) between years 1991 and 2001 in the case of socio-economic variables with rather stable period between 2001 and 2008. Socio-demographic variables had been expected to be relatively stable in the whole period (1991 – 2008). This hypothesis was confirmed. To some extent unexpected was the decline in regional inequality (T_B) between years 2001 and 2008 in the case of both age preference index and the unemployment rate. The values of MI showed very similar behaviour.

We could also confirm the expectation of the relation between the complexity of a variable (the share of external factors) and the importance of spatial aspects of inequality. We also showed the correlation between the complexity of a variable and the level to which can be attributed the most of the overall inequality. In this respect, the most distinct difference was found between socio-demographic variables (with inequality concentrating primarily on micro-scales) and the unemployment rate (inequality concentrating on macro-scales).

On the basis of individual data (census 2001) we also tried to compare the importance of spatial aspects with some other (non-spatial) dimensions of social differentiation. In this respect, it was found out that even though in case of every variable (such as the level of education, age, family status etc.) more important aspects can be found, spatial aspects are among the most important factors of social inequality in Czechia as well. It is clear that these results should be further verified, for example by applying a regression analysis which could control mutual relations among mentioned factors.

One of the expectations was also the effect of delimitation of regions on results. We compared sociogeographic regions according to Hampl with common administrative levels (ORP, LAU 2, NUTS 3) in year 2001. The analysis confirmed the importance of the choice of a regional level. This finding might be important when assessing various regional disparities or discussing public policies.

The results from the empirical analysis in Czechia were further compared with the situation in other countries (Slovakia, Poland, and Austria). Several similarities as well as differences were found.

The highest values of MI and, with few exceptions, also the values of T_B/T were found in Poland. The lowest T_B/T and MI in case of socio-demographic variables were found as excepted in Czechia and Slovakia which is, on the other hand, not true for socio-economic variables. These results might have been influenced by the spatial structure a differences in the number of units and regions in studied countries. In the case of the Theil index decomposition is thus appropriate to work with corrected values (subtract the null model). When using these corrected values $(T_B/T)^*$, the rank of studied countries according to spatial aspects of social inequalities is not that evident.

We also expected that historical borders in Poland and similarly in Czechia (Sudety) would be still visible in inequality patterns. This expectation was confirmed and not only in the case of socio-demographic variables but also for employment in agriculture and the unemployment rate. It is likely that this historical factor will be slowly losing its importance as results from year 2008 indicated.

The hypothesis about the dominant position (or significant difference) of capital regions in all countries (and other big cities in Poland due to its polycentric structure) was also confirmed. The difference between capitals and other regions is the best noticeable in the case of the share of university educated and the share of divorced in population.

The situation between years 2001 and 2008 was expected to be relatively stable according to spatial aspects of social inequalities, only in Poland was expected a slightly more dynamic change. Although in all studied countries the change between years 2001 a 2008 was not very dramatic, in Slovakia and Poland occurred rather significant changes. This period could not be satisfyingly assessed in Austria due to lack of proper data.

Analyses in other countries did not confirm the relation between the complexity of a variable and the relative importance of spatial dimension of social inequality especially due to prevailing macro-determinations in socio-demographic variables. The best example of this determination is the effect of Polish historical border on the structure of inequality in age preference index.

4. Selected bibliography

ANSELIN, L. (1995): Local indicators of spatial association – LISA. Geographical Analysis 27, no. 2, pp. 93-115.

BLAŽEK, J., NETRDOVÁ, P. (2009): Can development axes be indentified by socio-economic variables? The case of Czechia. Geografie, 114, no. 4, pp. 245-262.

BOURGUIGNON, F. (1979): Decomposable Income Inequality Measures. Econometrica, Econometric Society, 47, no. 4, pp. 901-920.

BOURGUIGNON, F., FERREIRA, F. H., MENENDEZ, M. (2003): Inequality of outcomes and inequality of opportunities in Brasil, DELTA Working Paper n. 2003-24.

CLIFF, A. D., ORD, J. K. (1973): Spatial autocorrelation. Pion, London, 178 p.

CUTRINI, E. (2008): Using entropy measures to disentangle regional from national localization patterns. Regional Science and Urban Economics, 39, pp. 243-250.

CZERNY, M. (2002), Introduction: Uneven Urban and Regional Development in Poland. In: European Urban and Regional Studies 9, no. 1, pp. 7-8.

GORZELAK, G. (2002): Polskie regiony w procesie integracji europejskiej. Studia Regionalne i Lokalne Nr. 2-3.

HAMPL, M. (1998): Realita, společnost a geografická organizace: hledání integrálního řádu. Praha, PřF UK, 110 p.

HAMPL, M. (2005): Geografická organizace společnosti v České republice: transformační procesy a jejich obecný kontext. Praha, Univerzita Karlova v Praze, Přírodovědecká fakulta, 147 p.

HAMPL, M. (2008): Nomotetická nebo idiografická geografie: alternativnost nebo komplementarita? Acta Universitatis Comeniae Bratislava, 50, pp. 19-31.

HARVEY, D. (1969) Explanation in Geography. London: Edward Arnold.

HURBÁNEK, P. (2008): Vývoj priestorovej polarizácie na regionálnej úrovni na Slovensku v rokoch 1996-2008. Geografia Cassoviensis, 2, no. 1, Košice, p. 53-58.

LIMPERT, E., STAHEL, W.E., ABBT, M. (2001): Log-normal distributions Gross the science: keys and clues. Bioscience, 51, no. 5, p. 341-352.

KORČÁK, J. (1941): Přírodní dualita statistického rozložení. Zvláštní otisk ze Statistického obzoru, sešit 5-6, předneseno v České statistické společnosti 20. 2. 1941.

KUKLIŃSKI, A. (1992): Societies and economies in transition. The case of Poland. Warsaw.

LITCHFIELD, J.A. (1999): Inequality: Methods and Tools. Text for World´s Bank Web Site on Inequality, Poverty, and Socio-economic Performance:

http://www.worldbank.org/poverty/iequal/index.htm.

MATLOVIČ, R. (2007): Hybridná idiograficko-nomotetická povaha geografie a koncept miesta s dôrazom na humánnu geografiu. Geografický časopis, 59, no. 1, pp. 3-23.

NETRDOVÁ, P., NOSEK, V. (2009): Přístupy k měření významu geografického rozměru nerovnoměrností. Geografie – Sborník ČGS, 114, no. 1, pp. 52-65.

NEZDAŘILOVÁ, E. (1984): Metody kvantitativní analýzy v geografii – se zaměřením na metody regrese a korelace. Diplomová práce. Katedra sociální geografie a regionálního rozvoje PřF UK, Praha, 172 p.

NOVOTNÝ, J. (2004): Decomposition of global and european socio-economic inequalities with attention to their regional dimension. Regional and Sectoral Economic Studies, 4, no. 2, pp. 53-72.

NOVOTNÝ, J., NOSEK, V. (2009): Nomothetic geography revisited: statistical distributions, basic generative mechanisms, and inequality measures. Geografie, 114, no. 4, pp. 282-298.

NUNEZ, J., TARTAKOWKSY, A. (2007): Inequality of Outcomes vs. Inequality of Opportunities in a Developing Country. An exploratory analysis for Chile. Estudios de Economica, 34, no. 2, pp. 185-202.

OUŘEDNÍČEK, M., TEMELOVÁ, J., POSPÍŠILOVÁ, L., eds. (2010) Atlas sociálně prostorové diferenciace České republiky. Karolinum, Praha. In print.

PALME, G. (1995): Divergenz regionaler Konvergenzklubs. Dynamische Wirtschaftsregionen in Österreich, WIFO Monatsberichte, 68, no. 12, pp. 769-781.

POON, J.P.H. (2003): Quantitative methods: producing quantitative methods narratives. Progress in Human Geography 27, no. 6, pp. 753-762.

REY, J., S. (2001): Spatial Analysis of Regional Income Inequality. Regional Economics Applications Laboratory, University of Illinois, 22 p.

REY, J., S., JANIKAS, M., V. (2005): Regional convergence, inequality, and space. Journal of Economic Geography, 5, no. 2, p. 155-176.

SHORROCKS, A.F. (1984): Inequality Decomposition by Population Subgroups. Econometrica, 52, pp. 1369-1385.

SHORROCKS, A.F., WAN, G. (2005): Spatial decomposition of inequality. Journal of Economic Geography, 5, no. 2, pp. 59-81.

SLOBODA, D. (2006): Slovensko a regionálne rozdiely: teórie, regióny, indikátory, metódy. Konzervatívny Inštitút M. R. Štefánika, Bratislava, 49 p.

SPURNÁ, P. (2008): Geograficky vážená regrese: metoda analýzy prostorové nestacionarity geografických jevů. Geografie – Sborník ČGS, 113, no. 2, pp. 21-35.

SUBRAMANIAN, S. (2004): Indicators of Inequality and Poverty. WIDER research paper no. 2004/25.