Lett Spat Resour Sci (2012) 5:95–101
DOI 10.1007/s12076-011-0071-y

ORIGINAL PAPER

Comparison of regional inequality in unemployment among four Central European countries: an inferential approach

Josef Novotný · Vojtěch Nosek

Received: 14 February 2011 / Accepted: 24 October 2011 / Published online: 6 November 2011

Abstract The article focuses on the comparison of regional inequality indices between different economic systems. We have shown that the common measures of regional inequality are not independent of the specifics of considered spatial system such as size, number and size of its regions, and overall population variability. In order to control for the effects of these parameters, we have attempted to isolate the stochastic (or spatially contingent) component of regional inequality using the method of a repetitive random spatial resampling. Using the Theil coefficient and its decomposition, the standardized (adjusted) measures of regional inequality have been obtained by confronting observed (unadjusted) figures with what they would be were the considered characteristic randomly spatially distributed within the map of a given system. As an example showing the importance of the proposed standardization procedure, we have compared regional inequality in the unemployment rate among Czechia, Slovakia, Poland, and Austria.

Keywords Regional inequality · Spatial resampling · Theil index · Central Europe · Unemployment

JEL Classification R12 · O18 · J64 · C49

Abbreviations
T: overall Theil coefficient of inequality;
T_B: between-region component of Theil inequality;
T_W: within-region component of Theil inequality;

J. Novotný (✉) · V. Nosek
Department of Social Geography and Regional Development, Faculty of Science, Charles University in Prague, Albertov 6, 128 43 Prague, Czech Republic
e-mail: pepino@natur.cuni.cz

V. Nosek
e-mail: vojtaa.nosek@seznam.cz

n: population size (number of sub-regional units);
y: variable under analysis;
i: index for individuals;
j: index for regions;
k: number of regions;
r: size of average region;
σ: population variability

1 Introduction

Both academics and practitioners dealing with regional economic issues are often interested in comparing regional inequality between different spatial systems. However, different arrangements of these systems such as different size, different number of regions, or different underlying population variability often make meaningful comparisons of indices of regional inequality problematic. As such, one may wish to have these measures standardized. In this article, we will use the method of a repetitive random spatial resampling to control for the stochastic (spatially contingent) component of regional inequality determined by the above mentioned parameters. Using the Theil coefficient and its decomposition, the standardized (adjusted) measures of regional inequality are obtained by confronting observed (unadjusted) figures with what they would be were the considered variable randomly spatially distributed within the map of a given system. We will show that this procedure can be utilized for making inferences about the significance of regional inequality and, subsequently, for comparing standardized (adjusted) results among different spatial systems.

The remainder of this paper is organized into three sections. In the following Sect. 2, we will briefly describe the background of the proposed approach. In Sect. 3, we will demonstrate by theoretical simulations that regional inequality indices are indeed not independent of factors such as different size of the considered country, number of regions, or different underlying population variability. As an illustrative example showing importance of the proposed standardization procedure, in Sect. 4, we will compare observed and adjusted regional inequality in the unemployment rate among four Central European countries, including Czechia, Slovakia, Poland, and Austria.

2 Background

The quantitative assessment of regional inequality represents one of the most common tasks in regional economic analysis. Conventionally, inequality between regions is quantified by an inequality index such as the Theil coefficient that is used here. To enable judgments about whether observed regional unbalances deserve further attention, the measure can be compared with the same indicator observed in other countries (or with that for other variables within the same country). However, such a simple descriptive approach can only provide us with an incomplete idea about the actual significance of regional disparities.

From the perspective adopted here, a general problem resides in the fact that it is usually assumed that population (in a statistical sense) is analyzed when working with regional (regionally aggregated) data. This is usually because regional inequality index is computed from figures for all regions in a given country. Alternatively, let us now think about regions (regional subpopulations) as samples drawn from a parent population that corresponds to the population of individuals, households or to other sub-regional entities. The actual significance of regional inequality should then be assessed with regard to the distribution of the observed variable within this parent population. In such a case, the decomposition of the overall parent population inequality (T) by a subgroup consistent inequality index (such as the Theil coefficient which is applied below) into its between-region (T_B) and within-region (T_W) component becomes a helpful tool (Shorrocks and Wan 2005; Netrdová and Nosek 2009). Formally, the overall Theil coefficient decomposition can be expressed as:

$$T = \left(\sum_{j=1}^{k} \frac{n_j}{n} \frac{y_j}{y} \ln \frac{y_j}{y} \right) + \left(\sum_{j=1}^{k} \frac{n_j}{n} \frac{y_j}{y} \sum_{i=1}^{m} \frac{n_{ij}}{n_j} \frac{y_{ij}}{y_j} \ln \frac{y_{ij}}{y_j} \right) = T_B + T_W$$

where n stands for population size; y denotes variable under analysis (such as the unemployment rate in Sect. 4); and subscripts i and j are indexes for individuals (sub-regional units) and regions, respectively. The extent to which between-region inequality explains overall inequality can tell us the extent to which location actually matters. The share of the between-region component in overall inequality (T_B/T) can be thus used as a quantitative measure of the relative importance of regional inequality (Novotný 2007).

However, as we will demonstrate in Sect. 3, both T_B and T_B/T are not independent of specific characteristics of economic systems such as their size or various spatial arrangements. The comparison among different countries may therefore be misleading and further statistical inferences may be needed to control for the country specific characteristics.

The inferential framework for regional inequality analysis is nevertheless undeveloped (Rey and Janikas 2005, p. 163). This is again mainly because regional characteristics are commonly understood as pertaining to population (in the statistical sense) and so it is assumed that no further statistical tests for regional inequality indices are needed. In fact, there have been few recent attempts to incorporate statistical inference into the analysis of regional inequality. More frequently, this exercise is undertaken to test for the statistical significance of observed effects—mostly changes in regional inequality measures over time (e.g. Brülhart and Traeger 2005; Ezcurra et al. 2007). In this case, similar methods are applied as those for testing indices of inequality in a non-spatial context. Tests based on a bootstrap where the sampling distribution of the inequality measure considered is estimated by repetitive random resampling with replacement are regarded as the method of choice (Mills and Zandvakili 1997).

More interestingly, with regard to the topics discussed here, Rey (2004) proposed an inferential framework for regional inequality analysis based on repetitive random spatial permutations of the considered data for a given map pattern (also used in

Table 1 Sensitivity to country size and/or to number of regions

n	r	k	σ	T	T_B	T_B/T
10 000	100	100	1.2	3.50	0.52	14.8%
2500	100	25	1.2	3.41	0.45	13.2%
900	100	9	1.2	3.18	0.32	10.1%
400	100	4	1.2	2.93	0.22	7.4%
200	100	2	1.2	2.64	0.10	3.8%

Yildirim et al. 2009). Essentially, the method attempts to estimate the stochastic or spatially contingent component of observed regional inequality—i.e. what the regional inequality would be was the analyzed phenomenon randomly spatially distributed within a given economic system. In Sect. 4, we develop this idea to obtain adjusted (standardized) measures of regional inequality in order to compare regional inequality in the unemployment rate among four Central European countries. In other words, Sect. 4 provides a real data example of an attempt to control for the spatially contingent component studied and explained on simulated data in Sect. 3.

3 Sensitivity of regional inequality indices to country size, number and size of regions, and population variability

Consider a hypothetical country that consists of n sub-regional units that are randomly divided into k regions of size r. For each unit we generate a pseudorandom number drawn from a certain lognormal distribution derived from the respective normal distribution with standard deviation σ. From this random dataset, we calculate T, T_B, and T_B/T. Then we repeat this procedure 1000 times and compute average T, T_B, and a T_B/T. We undertake four tests, in each of them holding fixed one of the parameters n, k, r, or σ. Table 1 shows the results for decreasing n and k but constant r and σ. Table 2 displays an alternative situation with decreasing n and r but k and σ being fixed. The results in Table 3 are based on the situations when n and σ are held constant and r decreases with increasing k. Finally, Table 4 shows how the indices of regional inequality respond to changing population variability σ, when keeping the other three parameters fixed. Inspection of the values of regional indices in particular tables clearly confirms that neither T_B nor T_B/T are independent of the country size, number and size of regions, and underlying population variability. Note that in these cases of simulations using randomly distributed pseudorandom numbers no spatial autocorrelation is present (the values in nearby locations are independent of each other). The values of T_B and T_B/T in Tables 1–4 thus capture merely the spatially contingent component of regional inequality. In reality, however, the spatial dependency of a majority of processes and events tends to be pervasive with significant context-specific effects on observed regional inequality (e.g. Novotný and Nosek 2009). As such, one might obviously wish to control for the spatially contingent component when assessing the real significance of observed regional inequality or comparing the observed regional inequality among different spatial systems. The latter will be done in the following section.

Table 2 Sensitivity to country size and/or to size of regions

n	r	k	σ	T	T_B	T_B/T
10 000	100	100	1.2	3.50	0.52	14.8%
2500	25	100	1.2	3.08	0.80	30.0%
900	9	100	1.2	2.96	1.40	47.3%
400	4	100	1.2	2.79	1.62	58.1%
200	2	100	1.2	2.60	2.16	83.1%

Table 3 Sensitivity to number and size of regions

n	r	k	σ	T	T_B	T_B/T
10 000	10 000	1	1.2	3.50	0.00	0%
10 000	2500	4	1.2	3.50	0.03	0.9%
10 000	625	16	1.2	3.50	0.15	4.2%
10 000	100	100	1.2	3.50	0.52	14.8%
10 000	25	400	1.2	3.50	1.09	31.0%

Table 4 Sensitivity to variability in the underlying population

n	r	k	σ	T	T_B	T_B/T
10 000	100	100	1.2	3.50	0.52	14.83%
10 000	100	100	1.5	4.97	1.20	24.25%
10 000	100	100	2	6.31	2.03	32.17%
10 000	100	100	4	8.11	3.53	43.54%
10 000	100	100	8	8.79	4.18	47.61%

4 Comparison of observed and adjusted indices of regional inequality

In line with the remarks in the previous section, here we apply the method of repetitive spatial resampling to isolate the spatially contingent (stochastic) component of observed regional inequality. This is used to adjust regional inequality indices and compare the real significance of regional unemployment inequality among four Central European countries. The particular steps are as follows: First, we quantify *observed* regional unemployment inequality (T_B) and relative importance of the between-region component (T_B/T). Second, we run random spatial permutations of the overall population-level characteristics within given maps of particular countries. From these “randomized maps” we calculate again the Theil coefficient of regional inequality (T_B^P), repeat this procedure 1000 times, and compute the average *simulated* T_B^P. By subtracting T_B^P from *observed* T_B we obtain *adjusted* T_{B^*} and calculate *adjusted* T_{B^*}/T. Both the *adjusted* T_{B^*} and above all *adjusted* T_{B^*}/T (i.e. the relative importance of regional inequality controlled for the spatially contingent component) become useful for comparison of regional inequality among different spatial systems.

In our empirical example we compare regional inequality in unemployment among Czechia, Slovakia, Poland, and Austria. Note that unemployment is a binary variable at the individual level and it can only be considered as cardinal if the individuals

Table 5 Observed, simulated, and adjusted indices of regional unemployment inequality

	Indicator	Czechia	Slovakia	Poland	Austria
Observed regional inequality indices	T	0.111	0.089	0.042	0.107
	T_B	0.090	0.058	0.030	0.079
	T_B/T	82%	66%	72%	74%
Simulated spatially contingent component	T_B^P	0.042	0.014	0.013	0.057
Adjusted regional inequality indices	T_{B*}	0.048	0.045	0.018	0.022
	T_{B*}/T	43%	50%	42%	21%

are merged into groups. Here smaller territorial units represent such groups so that inequality between these units substitutes the overall population inequality.[1] The regions used for the analysis in particular countries correspond to larger administrative districts (209 in Czechia, 79 in Slovakia, 380 in Poland, and 99 in Austria). For the purposes of this empirical example, we use data on unemployment rates from the 2001 national censuses. More recent comparable data for the smaller territorial units are not available for all of the considered countries.

The observed, simulated, and adjusted indices of regional inequality for the four analyzed countries appear in Table 5. Let us first look at the observed (unadjusted) figures. These results suggest that Czechia and Austria accounted for the higher levels of observed regional disparities than Poland with Slovakia having the lowest levels. It holds both when considering T_B and T_B/T. The latter measure is especially significant in Czechia, where the observed between-region component accounted for 82% of inequality between smaller territorial units.

When we control for the simulated spatially contingent component and adjust the regional inequality indices accordingly, the results, however, change dramatically. Slovakia with the lowest observed (unadjusted) share of regional inequality (66%) has, on contrary, the highest adjusted share of regional inequality (50%). Similarly, the share of observed regional inequality in Austria is the second highest (74%), but the lowest after the adjustment (with the spatially contingent component being the highest, 53%, from all considered countries). These results clearly show that inferences about the significance of regional inequality drawn from the adjusted (i.e. standardized) figures differ completely from conclusions obtained on the basis of conventional descriptive comparison of unadjusted indices. This suggests the practical importance of the proposed standardization for empirical comparisons of regional inequality indices among different spatial systems.

5 Conclusion

The focus of this short article has been on the comparison of regional inequality indices between different economic systems. Firstly, we have shown by theoretical

[1] The considered smaller territorial units are: obce in Czechia (6248 units), obce in Slovakia (2920), gminy in Poland (2478) and gemeinde in Austria (2359).

simulations that regional inequality indices (the Theil coefficient and its decomposition have been applied) are not independent of the specifics of the considered economic system in terms of size, number and size of its regions, and overall population variability. As such, we have attempted to perform a standardization of regional inequality indices by the estimation of the stochastic, spatially contingent component of between-region variation using the method of repetitive random spatial permutations. In short, the standardized (adjusted) measures have been obtained by confronting observed (unadjusted) regional inequality figures with what they would be were the considered characteristic randomly spatially distributed. The empirical example of the comparison of regional inequality in the unemployment among four Central European countries have demonstrated that the procedure may become a helpful tool when making inferences about both the statistical and practical significance of regional inequality measures.

Acknowledgements The authors acknowledge support provided by the Ministry of Education, Youth and Sports of the Czech Republic, project no. MSM0021620831 "Geographic Systems and Risk Processes in the Context of Global Change and European Integration."

References

Brülhart, M., Traeger, R.: An account of geographic concentration patterns in Europe. Reg. Sci. Urban Econ. **35**, 597–624 (2005)
Ezcurra, R., Pascual, P., Rapún, M.: The spatial distribution of income inequality in the European Union. Environ. Plan. A **39**, 869–890 (2007)
Mills, J.A., Zandvakili, S.: Statistical inference via bootstrapping for measures of economic inequality. J. Appl. Econom. **12**, 133–150 (1997)
Netrdová, P., Nosek, V.: Přístupy k měření významu geografického rozměru společenských nerovnoměrností. Geografie **114**, 52–65 (2009)
Novotný, J.: On the measurement of regional inequality: does spatial dimension of income inequality matter? Ann. Reg. Sci. **41**, 563–580 (2007)
Novotný, J., Nosek, V.: Nomothetic geography revisited: statistical distributions, their underlying principles, and inequality measures. Geografie **114**, 282–297 (2009)
Rey, S.J.: Spatial analysis of regional income inequality. In: Goodchild, M., Janelle, D. (eds.) Spatially Integrated Social Science: Examples in Best Practice, pp. 280–290. Oxford University Press, Oxford (2004)
Rey, S.J., Janikas, M.V.: Regional convergence, inequality, and space. J. Econ. Geogr. **5**, 155–176 (2005)
Shorrocks, A., Wan, G.: Spatial decomposition of inequality. J. Econ. Geogr. **5**, 59–81 (2005)
Yildirim, J., Öcal, N., Özyildirim, S.: Income inequality and economic convergence in Turkey: a spatial effect analysis. Int. Reg. Sci. Rev. **32**, 221–254 (2009)