Journal of Development Economics 106 (2014) 35–51

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Spatial inequality and development — Is there an inverted-U relationship?☆

Christian Lessmann *

Technische Universität Dresden, Germany
CESifo, Germany

ARTICLE INFO

Article history:
Received 20 October 2011
Received in revised form 16 August 2013
Accepted 25 August 2013

JEL classification:
D31
R1
O15

Keywords:
Regional inequality
Kuznets curve
Panel data
Semiparametric estimates

ABSTRACT

This paper studies the hypothesis of an inverted-U-shaped relationship between spatial inequality and economic development. The theory of Kuznets (1955) and Williamson (1965) suggests that (spatial) inequality first increases in the process of development, and then decreases. To test this hypothesis I have used a unique panel data set of spatial inequalities in 56 countries at different stages of economic development, covering the period 1980–2009. Parametric and semiparametric regressions are carried out using cross-section and (unbalanced) panel data. The results provide strong support for the existence of an inverted-U. I also find some evidence that spatial inequalities increase again at very high levels of economic development.

1. Introduction

The growing spatial income inequalities around the world have begun to attract considerable interest among academics and politicians. Spatial inequalities are important for at least two reasons. The first reason is that interregional inequality often goes along with undesirable interpersonal (or overall) inequality. The second reason is that interregional inequality often goes hand in hand with political and ethnic tensions, which undermine social cohesion and political stability Kanbur and Venables (2005a). In very extreme cases, this may increase the risk of internal conflicts and civil wars [see, e.g., Deiwiks et al. (2012), Buhaug et al. (2012), Lessmann (2013b)].

Despite the importance of spatial inequality for policy concerns, little is known about its determinants.[1] One of the most important theories on the determinants of spatial inequality dates back to Kuznets (1955) and Williamson (1965). In his seminal paper, Kuznets shows that as countries develop from farm-based economies to industrial economies, income inequality first increases, then peaks, and then decreases. Thus, the trajectory of this relationship is inverted-U-shaped — what we call the *Kuznets curve* today. The reason is that in an early stage of development, very few people benefit from the increasing investment in physical capital, and income inequality increases. At a later stage of development, more and more workers shift from the agricultural sector to the industrial sector so that inequality falls. Williamson adopted this idea for the case of spatial inequality. He argues that the industrialization was driven by the discovery and utilization of natural resources such as coal and iron. Those natural resources are often not equally distributed within countries (think of the western regions of France or the Ruhr area in Germany). The economic prosperity in the industrialization process is therefore also unequally distributed within countries, so that spatial inequalities rise in this process. At a later stage of development, the more attractive employment opportunities in the booming regions attract workers from abroad, depressing wages in destination regions but increasing wages in home regions. Thus, a natural convergence process starts, possibly encouraged by government policies; therefore, regional inequality falls, creating again an inverted-U-shaped relationship.

Surprisingly, only a few empirical studies – such as Williamson (1965), Amos (1988), Ezcurra and Rapun (2006), or Barrios and Strobl (2009) – have tried to provide evidence for the Kuznets curve in spatial inequality. A major reason for the scarcity of research in this field is the poor availability of regional economic accounts, which are necessary to

☆ I would like to thank Susan Kircheis for helping me with the data collection and preparation. I would also like to thank Nicola Gennaioli, Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer for sharing their regional data with me. Vincenzo Verardi gratefully advised me on semiparametric estimation techniques in Stata. Finally, I would like to thank David Weil and two anonymous referees for their helpful comments.

* Faculty of Business and Economics, D-01062 Dresden, Germany. Tel.: +49 351 463 32895; fax: +49 351 463 37052.
E-mail address: christian.lessmann@tu-dresden.de.

[1] See Barrios and Strobl (2009) for an overview of theoretical studies concerning the relationship between spatial inequality and development.

0304-3878/$ – see front matter
http://dx.doi.org/10.1016/j.jdeveco.2013.08.011

calculate inequality measures. In the case of OECD countries, data collection is quite easy, since the OECD Regional Statistics, EUROSTAT, and Cambridge Econometrics (CAMECON) provide many useful regional data. Unfortunately, data collection is more difficult for other countries than OECD member states, since no international database contains the relevant information. The major problem for this kind of research is that it is essential either to have historical data for single countries or to include poorer countries in a cross-country data set, since the theories of Kuznets and Williamson point at the deep structural changes associated with the industrialization process. In this paper, I use a unique panel data set on spatial inequality, which covers 56 countries at very different levels of economic development for the period 1980–2009 to investigate this hypothesis. I have collected a new data set on spatial inequality around the world, where much of the regional data was provided by national statistical offices or central banks on individual request. I show, based on cross-country as well as panel data, that the relationship between spatial inequality and economic development has a nonlinear, inverted-U-shaped trajectory implying that economic growth first increases spatial inequality, and later – at higher stages of economic development – inequalities fall. This result is obtained in standard parametric regressions using polynomial functions of the income variable as well as in regressions which use a more flexible semiparametric approach. There is also some evidence that regional inequalities increase again at very high levels of economic development, which may be related to tertiarization, i.e., a shift from manufacturing industries to modern service sectors. However, the finding of increasing spatial inequalities at very high levels of economic development is sensitive to the estimation procedure and the logarithmic transformation of income data. The general finding of an inverted-U is robust to the inclusion of a number of covariates the literature has shown to affect spatial inequality. Moreover, I have sectoral data, which suggests that the shift from agriculture to the industry and service sectors causes the inverted-U. The empirical evidence therefore supports the theory of Kuznets (1955) and Williamson (1965).

The remainder of the paper is organized as follows: Section 2 surveys the existing theoretical and empirical literature on the link between spatial inequality and economic development. Section 3 presents the unique data set on regional inequality and discusses measurement issues. At this, I also relate the measure of spatial inequality to a measure of personal income inequality. Section 4 provides the main econometric analysis. Finally, Section 5 sums up the results and concludes.

2. Spatial inequality and development: Existing theory and evidence

2.1. Theory

Williamson (1965) was the first who suggested an inverted-U-shaped relationship between spatial inequality and economic development. Based on the ideas of Kuznets (1955), who studied personal inequality, he argues that spatial inequalities are affected in quite a similar manner. The spatial concentration of wealth- and income-generating resources results in increasing regional inequalities in early stages of economic development, followed by a more widespread dispersion of income in later stages. Following Williamson, four reasons are decisive for the evolution of spatial inequalities: natural resources, migration, capital mobility, and government policies. He argues that most natural resources are point resources and, thus, are unequally distributed among the different regions of a country. A discovery of new resources will then increase unbalanced development of regions, and a selective influx of labor and capital, perhaps encouraged by government policies, will lead to a further increase in spatial inequality. At later stages of economic development, new resources will be discovered in less developed regions (or the demand for existing resources will increase), and government policies will concentrate on lagging regions, so that the process is reversed. Based on these ideas, he formulates the hypothesis

> "that the early stages of national development generate increasingly large North–South income differentials. Somewhere during the course of development, some or all of the disequilibrating tendencies diminish, causing a reversal in the pattern of interregional inequality. Instead of divergence in interregional levels of development, convergence becomes the rule, with the backward regions closing the development gap between themselves and the already industrialized areas."
>
> [Williamson, 1965, p. 9]

Thus, the relationship between spatial inequality and economic development is expected to be inverted-U-shaped. Williamson was able to find support for this hypothesis in cross-country data.

A more formal foundation of the inverted-U hypothesis is provided by Barrios and Strobl (2009) based on Lucas (2000). Their model analyzes the dynamics of regional growth after a technological shock (innovation) takes place, which initially benefits one region. Growth is accelerated in this leading region, implying that regional inequalities increase. The other regions will follow the leading region with a lag, whose magnitude depends on differing technological capabilities. Those lagging regions, which adopt the new technology, will grow at the rate of the leading region plus an additional growth effect determined by the natural rate of convergence. Thus, regional inequalities increase, peak, and decrease. It is important to note that while the argument of Kuznets and Williamson applies to long-lasting structural changes, this framework suggests an inverted-U-shaped relationship even in a shorter time horizon.

Amos (1988) discusses the inverted-U hypothesis for U.S. counties (intra-state inequalities). Since the U.S. has already reached a very high level of economic development, he argues that the inverted-U pattern in the scheme of Kuznets and Williamson must have been completed. Thus, he is primarily interested in what happens after the inverted-U: stabilization or increase of spatial inequalities (a further decrease would just imply that the inverted-U pattern has not been completed yet). The neoclassical growth theory suggests stabilization. Amos argues that neoclassical factor market adjustment mechanisms had more than 100 years to compensate for the disequilibrating technological shocks caused by the industrial revolution, and therefore regional inequality should have stabilized. In contrast, increasing inequalities may reflect other aspects of regional development not covered by the neoclassical theory: urban decay, suburbanization, rural decline, etc. Increasing inequalities may, however, also be the result of disequilibrating shocks and the beginning of a new inverted-U process which follows the initial one. The empirical findings of Amos point at increasing inequalities within U.S. states. Interestingly, this finding is consistent with studies of personal inequality such as List and Gallet (1999).

2.2. Existing studies

As mentioned in the introduction, existing empirical evidence on the relationship between spatial inequality and economic development is scarce. The highly influential study by Williamson (1965) was the first to explore a possible inverted-U-shaped relationship between spatial inequality and development. Williamson analyzes cross-country and time series data of 24 countries, including a number of developing countries such as Indonesia, India, and several South American countries. The evidence supports the hypothesis of an inverted-U. A more recent cross-country study is provided by Ezcurra and Rapun (2006), who consider 14 Western European countries for the period 1980–2002. Using a semiparametric approach, the authors do not find an inverted-U-shaped relationship between spatial inequality and development. As the authors state, this is not surprising, since all considered countries have reached a high level of development. But there is some evidence that increases in GDP coincide with a decrease in spatial inequality at the beginning of the observation period, indicating that the inverted-U pattern had not been completed at that time. At later stages of economic development, spatial inequalities tend to stabilize. A related study is Barrios and Strobl

(2009), who focus on 12 EU countries for the period 1975–2000. Although only highly developed countries are considered, they find evidence of an inverted-U, based on a semiparametric regression approach.

Besides the cross-country studies, there also exist some studies on single countries. As mentioned above, Amos (1988) analyzes spatial inequality within U.S. states, finding that inequalities increased with development. The major finding is supported by Fan and Casetti (1994). Another case study is provided by Terrasi (1999) for Italian regions (1953–1993). Similarly to the U.S. case, her parametric estimates point at a U-shaped relationship between spatial inequality and development. Terrasi's interpretation of the empirical finding is that a "new era of divergence has begun in connection with the emergence of high technology industries and producer services as the new leading sectors" [Terrasi (1999), p. 508]. Altogether, the studies of highly developed countries point at increasing spatial inequalities at very high levels of economic development.

The discussion of the existing literature shows that no study has been carried out since Williamson (1965) looking at countries at different levels of economic development. This is surprising as the original theory of Kuznets and Williamson aims to explain the effect of deep structural changes, which are difficult to isolate in high-income economies such as those of western Europe or the U.S. without having historical regional accounts. The aim of this paper is to fill this gap in the literature. For this purpose, a unique data set on spatial inequality was collected, as described in the following section in detail. My reexamination of Williamson's work is also interesting in that several studies point at increasing spatial inequalities after the inverted-U pattern has completed. Based on the new cross-country data, this hypothesis can be tested for a wide range of countries. Of course, today's available econometric methodologies have several advantages over those of the 1960s, so that one might have more trust in the new findings. Nevertheless, to make the results comparable to the initial study of Williamson, I conduct parametric regressions as well as semiparametric regressions, which are commonly used in the recent literature.

3. Spatial inequality around the world — a new data set

Spatial inequality matters because it might be a consequence of ethnic discrimination or market failures such as excessive migration [Mills and Ferranti (1971), Boadway and Flatters (1982)]. But spatial inequality also matters because it is a component of overall inequality [Kanbur et al. (2005)]. Large horizontal inequality between the regions of a country might cause large vertical inequality between all individuals [Stewart (2000, 2002)]. Therefore, spatial inequality is becoming a more and more prominent topic in the development economics literature. This section discusses the measurement concept of spatial inequality, presents the new data set, and clarifies the relationship between interregional (spatial) and interpersonal inequality.

3.1. Measuring spatial inequality

The discussion in the introductory section mentions the problems related to data availability, but even if one has access to suitable regional data, the selection of an appropriate inequality measure is difficult. I resort to the weighted coefficient of variation (*WCV*) of regional GDP per capita (p.c.), which is widely used in the literature on spatial inequality [see, e.g., Williamson (1965), Ezcurra and Rapun (2006), Rodríguez-Pose and Ezcurra (2010)][2]:

$$WCV := \frac{1}{\bar{y}}\left[\sum_{i=1}^{n} p_i(\bar{y}-y_i)^2\right]^{1/2}, \tag{1}$$

where $\bar{y}$ is the country's average GDP p.c., y_i is the GDP p.c. of region i, p_i is the share of the country's total population in region i, and n is the number of spatial units.[3] The advantages of this measure are that it is mean-independent, independent of the sizes and the number of spatial units, and robust against single extreme observations, and that it satisfies the Pigou–Dalton transfer principle [Dalton (1920), Pigou (1912)], which states that a transfer from poor to rich regions should unambiguously increase the inequality measure [see Sen (1973) and Mehran (1976) for details]. Other inequality measures such as the (ln of the) standard deviation of regional GDP p.c., which are often used in the literature on growth and convergence [see, e.g., Barro and Sala-i-Martin (1992), Sala-i-Martin (1996)], are less appropriate in our context, since they cannot account for the heterogeneity of regions with respect to (population) size. This is a very important issue here due to the lack of a uniform territorial classification for all countries in the data set. In countries with large economic differences and a very unequally distributed population, an unweighted inequality measure might be difficult to interpret. An example should illustrate the problem. The northern Canadian Territories are much poorer than the provinces to the south, so that an inequality measure might indicate large economic differences, although very few people are actually poor (note that the Territories are inhabited by only 100,000 people in total). Therefore, it is necessary to calculate a concentration measure such as the *WCV*, which incorporates the different sizes of spatial units within a country.

For explaining spatial inequalities, which is at the heart of this study, it would be favorable to have information on economic activities in gridded data, since this implies perfect comparability of individual cells. An example is the data set provided by Nordhaus et al. (2006), which contains gridded output for the year 1990. This data is, however, not suitable for panel data analysis since there is no variation over time, and the quality is problematic in developing countries. Therefore, I have tried to use a territorial classification, which provides a high degree of homogeneity of sub-national units. I refer to regional data based on OECD TL2 or NUTS2 level for OECD member countries. If NUTS2 data is not provided, I use the NUTS3 level data instead.[4] For non-OECD countries, I have usually no choice concerning the territorial level, since the statistical offices provide mostly data on state or province level. Here, weighing of the inequality measure becomes increasingly important. The main results of this study are, however, not sensitive to the choice of the territorial level. In the robustness section, I provide estimation results using NUTS1 instead of NUTS2 level data for EU countries. Moreover, following Gennaioli et al. (2013a), I have also considered the territorial level of a country, that has the highest political authority. Again, the main findings of the analysis remain robust. Some descriptive statistics on the territorial characteristics of the countries considered in the analysis are provided by Table A.2 in the Appendix.

3.2. Spatial inequality: results

I have calculated the weighted coefficient of variation based on the regional GDP p.c. for 56 countries covering the period 1980–2009. Note that the frequency of the data varies by country: for the OECD countries the underlying panel is almost balanced, but for developing countries there are quite large gaps in the data. Table 1 presents the means of these calculations for the most recent years, 2000–2009, subdivided by the different regions of the world following the standard World Bank classification.

[2] See Bendel et al. (1989) for a comparison of standard inequality measures.

[3] Note that the Theil index is not applicable for cross-section analysis with large variations in the number of sub-national units of the countries considered, since its values range from 0 to ln n [Hale (2003)].

[4] A complete list of countries, territorial levels, periods covered, and sources is provided in Table A.1 in the Appendix. NUTS refers to *Nomenclature of Territorial Units for Statistics*. OECD territorial level 2 refers to macro-regions.

Table 1
Spatial inequality around the world.

Country	WCV	Country	WCV
Europe & Central Asia		*North America*	
Austria	0.20	Canada	0.16
Belgium	0.35	United States of America	0.17
Bulgaria	0.29	Average	0.17
Croatia	0.21	*Latin America & Caribbean*	
Czech Republic	0.39	Bolivia	0.29
Denmark	0.11	Brazil	0.48
Estonia	0.42	Chile	0.35
Finland	0.17	Colombia	0.46
France	0.29	Mexico	0.59
Georgia	0.19	Panama	0.46
Germany	0.20	Peru	0.42
Greece	0.13	Average	0.44
Hungary	0.40	*East Asia & Pacific*	
Ireland	0.17	Australia	0.09
Italy	0.27	China	0.51
Kazakhstan	0.75	Indonesia	0.89
Latvia	0.53	Japan	0.13
Lithuania	0.30	Korea, Rep. (South)	0.06
Netherlands	0.14	Mongolia	0.67
Norway	0.32	New Zealand	0.07
Poland	0.25	Philippines	0.62
Portugal	0.25	Thailand	0.88
Romania	0.39	Average	0.44
Russian Federation	0.37	*South Asia*	
Slovak Republic	0.46	India	0.42
Slovenia	0.18	*Sub-Sahara Africa*	
Spain	0.21	South Africa	0.41
Sweden	0.21	Tanzania	0.37
Switzerland	0.20	Average	0.39
Turkey	0.43	*Middle East & North Africa*	
Ukraine	0.58	Iran, Islamic Rep.	0.56
United Kingdom	0.37	Malta	0.07
Uzbekistan	0.51	Average	0.31
Average	0.31		

A first observation from the summary statistics is an obvious link between spatial inequality and development. High-income countries in the core of Europe, Scandinavia, and North America have much lower spatial inequalities than low- and middle-income countries in South America and Asia. But there are also interesting variations within the different country groups; for example, among the European countries the United Kingdom and Belgium have quite high spatial inequalities, while Denmark and the Netherlands are much more homogeneous.

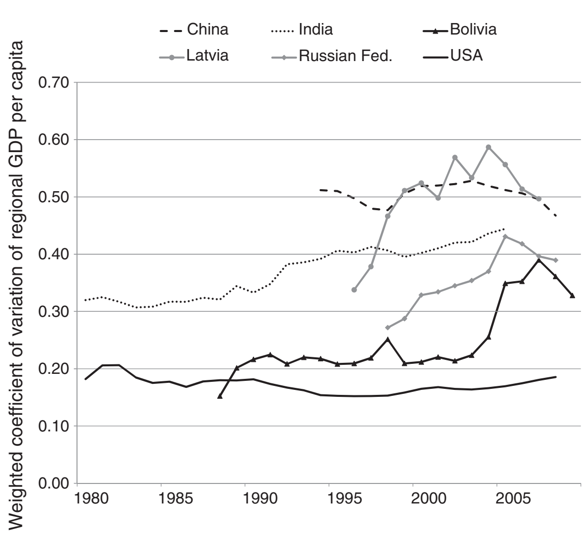

Fig. 1. Trends in regional inequality 1980–2009.

Inequality levels are relevant not only for this kind of analysis, but also for the development over time. Fig. 1 shows the weighted coefficient of variation for China, India, Bolivia, Latvia, Russia, and the United States. The figure illustrates that spatial inequalities vary quite a lot within countries over time, which is important for the panel data analysis conducted below. Concerning the individual time series, an interesting case is China, where the development of spatial inequality since the end of the 1990s has some resemblance to an inverted-U-shaped curve. Jian et al. (1996) find, based on long time series data on Chinese provinces, that regions converged before the Cultural Revolution, diverged during it, and subsequently converged again. China has experienced rapidly increasing spatial inequalities since the opening of the country to the world market in the 1990s, which was accompanied by fast economic growth [see Chen and Fleisher (1996), Wei et al. (2009) and Kanbur and Zhang (2005)]. Since then, spatial inequality has been on the decrease again, perhaps caused by the Western Development Program, which aims to restore a more balanced regional development [see Lessmann (2013a)].

We can also learn from the experience of Bolivia, which faced rapidly increasing spatial inequalities because the regions of the country have benefited differently from increasing resource revenues. But since the central government gained control of the natural gas resource revenues in 2006, spatial inequality has started to decrease again. Also, the data of Latvia and the Russian Federation, which both faced a rapid structural change after the breakup of the Soviet Union, support the hypothesis of a nonlinear relationship between spatial inequality and development. The case of India shows that the strong growth period, which started at the beginning of the 1990s, has increased spatial inequality significantly, and no turning point of this trend has been reached yet. The examples show that there is quite a lot of variation in my measure of spatial inequality. However, as the example of the United States illustrates, this does not apply to all countries in the data set. This is one of the reasons why I explore cross-section data, which focuses on the variation of spatial inequality *between* countries, as well as panel data, which focuses on the variation *within* countries [see Wooldridge (2002) for the pros and cons of cross-section versus panel data].

A remark is necessary concerning the quality of the regional economic accounts used to calculate the inequality measures. One might reasonably argue that the quality of the data is low in less developed countries due to the limited resources of national statistical offices. I have to admit, that this might be a source of bias in my analysis. However, omitting the poorest countries is no solution, since this information is needed to estimate the upward sloping part of the Kuznets curve. In order to reduce a potential bias, I have used only these data of developing countries, that are officially published by the national statistical offices or central banks. I have refrained from using data of other external sources, such the United Nations Human Development Reports (HDR). In several cases, the aggregation of regional data reported in the UN HDRs does not fit to the country level income data published in the World Development Indicators, which is one of my major data sources. This implies that the quality of regional data is very low here, therefore I concentrate on data, where the regional economic accounts do not deviate from national accounts. Note that also in high developed countries, regional economic accounts are often breakdowns from national accounts, which require many critical assumptions on the distribution of economic activities among the local economies. Therefore, one should keep in mind that the quality of the regional data used in this sort of studies is not comparable to national accounts.

3.3. Sub-national price disparities

An important issue for the measurement of spatial inequality is the possible differences in regional price levels within a country. The data used to calculate inequality measures at the country level is not adjusted for regional price disparities, since the necessary information is not available for the majority of countries [see also Gennaioli et al. (2013a), pp. 111–112]. But prices in non-tradable goods such as housing, prepared food, etc. differ across sub-national regions [Engel and Rogers (1996)]. This may bias the inequality measure, since a part of the nominal differences in income between regions is compensated by price adjustments in non-tradable goods. Thus, ordinary inequality measures overestimate inequality in real terms. The problem can be illustrated using the examples of Canada and the United States of America.

Statistics Canada is one of the few statistical offices, which provides regional consumer price indices (CPI) for a longer period (1983–2008; base year 2002).[5] The CPI measures changes in the price level of a market basket of consumer goods and services purchased by households. All relevant product groups such as food, communication, housing, etc. are considered in the provincial price indices. Importantly, the regional price data published by Statistics Canada is concerned only about changes in prices, i.e. there is no information available on regional differences in price levels. Therefore, I can only investigate whether the consideration of regional prices yields different conclusions for the development of spatial inequality over time compared to national prices. This is relevant for the panel data analysis presented in Sections 4.2.2 and 4.3.2. However, no conclusions about level effects are possible, which are at the heart of cross-section regressions.

In June 2013, the Bureau of Economic Affairs in the U.S. Department of Commerce released experimental real, or inflation-adjusted, estimates of personal income for states and metropolitan areas. The inflation-adjustments are based on regional price parities that provide a measure of differences in price levels across each state and region relative to the national price level (and not relative to a common base year by state as in the case of the Canadian data). Therefore, the U.S. data can be used to compare price levels between sub-national regions, which is relevant for the cross-section regressions presented in Sections 4.2.1 and 4.3.1. The data also allows one to derive conclusions concerning changes in the distribution of nominal and real incomes respectively, but the period (2007–2011) is very short.[6] That is the major reason why I refer to the two different data sources: Canadian and U.S. data.

Fig. 2 compares the *WCV* based on nominal and real regional prices for Canada and the U.S. The solid line reflects the *WCV* based on the nominal regional GDP, which is equal to the *WCV* based on real national prices due to the property of mean independence of the inequality measure. The dotted lines reflect the *WCV* considering differences in regional price changes (Canada) and/or differences in price levels (U.S.). Recall that I cannot interpret differences in the level of the *WCV* in the case of Canada for the reasons discussed above. I first discuss the Canadian data. In the base year 2002, the inequality measures based on nominal and real prices are equal by construction. As expected, the variation in spatial inequality is lower if changes in real provincial prices are considered. The inequality measure varies between [0.123; 0.261] using national prices [$\mu = 0.161$; $\sigma = 0.035$], and [0.136; 0.231] including price adjustments at the province level [$\mu = 0.166$; $\sigma = 0.028$]. Thus, regional prices adjustments compensate for some part in spatial inequality in nominal figures. Moreover, the calculations show that the development of spatial inequality over time is very similar for nominal and regionally adjusted prices. The correlation between the two time series is very high with a correlation coefficient of 0.955. This implies that – at least in case of Canadian provinces – the choice of the particular measurement does not make a serious difference for panel data regressions, which focus on the within country variation over time.

[5] The provincial CPIs are taken from CANSIM Table 326-0021.

[6] Regional price parities have been calculated also for 2005 and 2006, but there has been a change in the methodology, therefore the 2007–2011 data is the most recent, consistent data set.

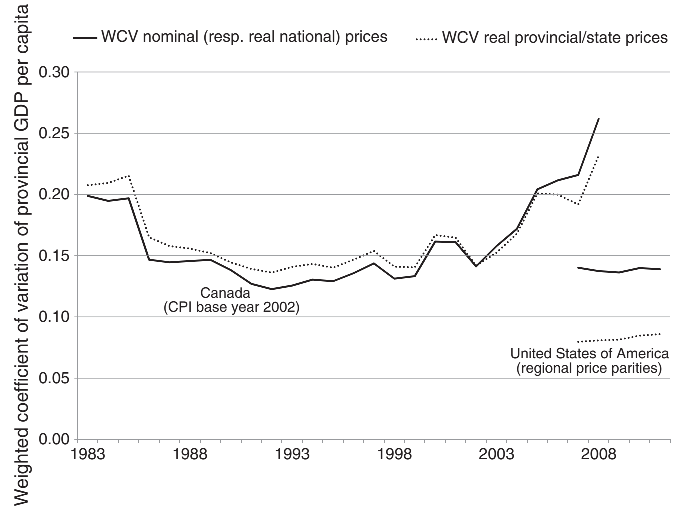

Fig. 2. Measures of spatial inequality based on real sub-national and national prices.

Let me now turn to the U.S. data that allows for an interpretation of level effects. Fig. 2 compares the *WCV* based on the nominal regional GDP p.c. and the GDP p.c. based on regional price parities, respectively. We see that the price adjustments in nontradables at the state level compensate for a quite large share of nominal income differentials. The *WCV* is about one third smaller using regional price parities compared to the case where I use the national price level. Importantly, spatial inequality does not vanish due to the price adjustments, which would suggest that real spatial inequality is not an issue at all. But one should keep in mind that the inequality measures based on national prices overstate real spatial inequality to some extent. Whether these price effects are comparable for all countries in my data set is not clear. If there is a large distance between regional markets, the lower degree of regional integration may result in comparatively large price adjustments for nontradables at the regional level. The U.S. forms a very large country in terms of area therefore price levels differ significantly. In contrast, small countries with highly integrated markets might have lower regional price disparities implying that the bias in the degree of spatial inequality based on nominal values is relatively small. This problem in the measurement of spatial inequality cannot be solved due to data limitations. In the forthcoming regression analysis, I consider different spatial control variables in order to reduce this bias. Note that this issue is only a problem in cross-section regressions. The panel regressions consider country fixed effects which control for a time invariant, unobservable heterogeneity in the inequality measure.

3.4. Spatial inequality vs. personal inequality

Kanbur et al. (2005) argue that interregional (spatial) inequality contributes to interpersonal inequality. Estimations by Yemtsov (2005) for Russia and Elbers et al. (2005) for Ecuador, Madagascar, and Mozambique show that spatial inequality explains about one-third of interpersonal income inequality. Hence, the question arises how strong this relationship is in a large cross section of countries like the one used in this paper.

To shed some light on this research question, I relate my measure of spatial inequality to the Gini coefficient of household income inequality. Gini coefficients are taken from the World Bank project "All the Ginis" [Milanovic (2013)]. This dataset includes combined and standardized Gini data from eight original sources (Luxembourg Income Study dataset, Socio-Economic Database for Latin America and the Caribbean, etc.). The quality of the datasets determines which values are included in the

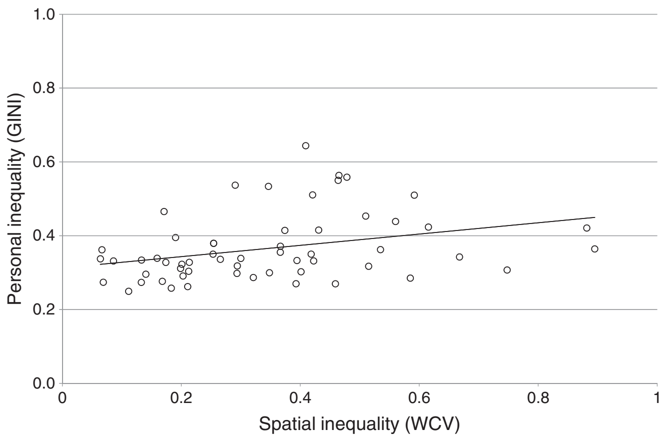

Fig. 3. Spatial inequality and personal inequality.

standardized Gini variable. I take averages of the period 2000–2009 which is similar to the data on spatial inequality presented above.[7]

The main finding is illustrated in Fig. 3 The abscissa denotes the degree of spatial inequality measured by the weighted coefficient of variation of regional GDP per capita. The ordinate denotes the degree of personal inequality measured by the Gini coefficient. The trendline for the relationship between both variables is also included in the graph. A simple linear regression of personal inequality on spatial inequality yields $GINI_i = 0.313 + 0.152 \cdot WCV_i$, with the coefficient of WCV_i being statistically significant (t-value: 2.97). The correlation coefficient between both variables is 0.324 suggesting that spatial inequality explains a significant proportion of personal inequality, which is in line with the results of Yemtsov (2005) and Elbers et al. (2005). But still spatial inequality is something different. A very high correlation between both variables would imply that it does not make a difference which dimension is under study: horizontal (spatial) inequality or vertical (personal) inequality. This is obviously not the case. Take for example the U.S.: personal inequality is fairly high ($GINI = 0.47$), while regional inequality is relatively low ($WCV = 0.17$).[8] The opposite is true for other countries. In Slovenia, for example, personal inequality is very low ($GINI = 0.27$), while spatial inequality is relatively high ($WCV = 0.58$). In light of this, it is interesting to learn more about the determinants of spatial inequality, in particular the relationship between spatial inequality and development, which is at the heart of the rest of the paper.

4. Econometric analysis

4.1. Methodology

This study uses two different approaches to test for the pattern of an inverted-U-shaped relationship between spatial inequality and development. First, I examine cross-section data as presented in Section 3 using a period average over 2000–2009. The theory of Kuznets and Williamson suggests that less-developed (more-developed) countries tend to fall along the positively (negatively) sloped region of the Kuznets curve, which can be tested in a cross-country framework focusing on between-country differences. Here I follow Williamson (1965). Second, I analyze the (unbalanced) panel data set covering the period 1980–2009. Using panel data has the advantage that I can eliminate unobserved heterogeneity between countries by including country fixed effects [Baltagi (2005)]. Since there exist numerous unobservable factors driving spatial inequality, which might cause an omitted variable bias, this is important for maintaining the quality of the analysis. Examples include geographic factors such as fragmentation, mountains, coasts, deserts, etc., which are determinants of spatial inequality, but difficult to consider in an econometric analysis, which focuses on country-level data. The country dummies capture all of these country-specific determinants of spatial inequality. In contrast to the cross-section estimations, panel regressions concentrate on within-country variations, which are important here, because they consider the dynamics of structural changes. Besides these advantages, the potential gain from panel data is limited in my study. The theoretical motivation for this study is based on structural-changes in the economy, which cause the process of economic development and affects regional inequalities [see Kuznets (1955) and Barrios and Strobl (2009)]. These structural changes may take decades to be observable in the data, therefore it is desirable to have long time series data. As my data set covers the period 1980–2009, my analysis takes a medium term perspective of only 30 years. This data limitation is basically the reasons why I use the two different approaches: cross country and panel data.

Concerning the estimation procedure, I consider two different econometric methods: a parametric ordinary least squares approach and a semiparametric partially linear model. The econometric representation of the Kuznets curve in the parametric regression approach is given by

$$WCV_{i,t} = \alpha_i + \sum_{j=1}^{k} \beta_j Y_{i,t}^{j} + \sum_{m=1}^{q} \gamma_m X_{m,i,t}, + \mu_t + \epsilon_{i,t}, \qquad (2)$$
$$i = 1, 2, \ldots, N, \quad t = 1, 2, \ldots, T$$

where $WCV_{i,t}$ is the weighted coefficient of variation of the regional GDP per capita for country i at time t; $Y_{i,t}$ is the ln of the GDP p.c. at the country level, which enters the regression in a polynomial form of degree j; $X_{m,i,t}$ represents q different control variables; α_i are the estimated country fixed effects; μ_t are time fixed effects; and $\epsilon_{i,t}$ is a random error term. The coefficients of interest are the β_j. For $k = 2$ the polynomial function is quadratic, and I expect β_1 to be positive and β_2 to be negative, implying an inverted-U-shaped relationship between spatial inequality and development. But as Amos (1988) shows, spatial inequality might increase at high levels of development after the inverted-U pattern has been completed; therefore I also consider polynomials of higher degree. The estimation equation of the cross-sectional model looks similar to the panel data model represented by Eq. (2), but it has no time dimension t and no fixed effects.

The parametric regression model described above has the advantage that it directly tests for the existence of an inverted-U as suggested by the theory. The functional form of the effect of economic development Y on spatial inequality WCV is given by the polynomial of degree j. However, by simply using higher order terms to estimate a possibly nonlinear relationship we place a fairly strong restriction on the possible link between the variables of interest that may not reflect the true underlying relationship. As suggested by Durlauf (2001), semiparametric methods are the more appropriate approach for studies of growth and convergence because of parameter heterogeneity. That means that the effect of one variable on another cannot be captured by a constant coefficient. The marginal effect varies by country or with other variables. By using a semiparametric approach one can investigate the possible nonlinear effect of economic development on spatial inequality in a flexible way, while simultaneously allowing for linear effects of other conditioning variables.[9] The equation to be estimated has the following form (omitting subscripts for reasons of clarity):

$$WCV = \alpha + f(Y) + \gamma X + \epsilon, \qquad (3)$$

where X is a set of explanatory variables that are assumed to have a linear effect on WCV, $f(\cdot)$ is an unknown smooth function of Y, which I

[7] The "All the Ginis" data set does not cover all countries, therefore I make use of additional other resources. No inequality data is provided for Malta, therefore I refer to the data reported by the CIA World Factbook. In case of New Zealand the most recent data is provided for 1997, which I take here instead of the period average.

[8] Sample means: $\overline{GINI} = 0.37, \overline{WCV} = 0.35$.

[9] DiNardo and Tobias (2001) provide a very helpful discussion of semiparametric methods.

expect to be nonlinear, and is a random error term. Thus, γX represents the parametric and $f(Y)$ the nonparametric part of the model. In the examination of cross-section data, I refer to the estimator proposed by Robinson (1988). The intuition for this estimator is the following: in a first step, an estimate of $\hat{\gamma}$ is obtained using a procedure that is similar to the way in which variables can be partialed out of an OLS regression (but using nonparametric regressions); in a second step, a kernel regression of $WCV-\hat{\gamma}X$ on Y is performed. In all stages, a Gaussian kernel weighted local polynomial fit is used for kernel regressions. This estimation procedure has been implemented in *Stata* by Verardi and Debarsy (2012). The panel data estimations are based on the estimator proposed by Baltagi and Li (2002), which has been developed for fixed effects models. To remove the unobservable fixed effects, the estimator takes first differences of Eq. (3). Then, OLS regressions are carried considering spline functions for the non-linear part of the model. The fixed effects are calculated based on the estimated parameters, and finally, $f(\cdot)$ is estimated with kernel or spline regression using an error component model. I use aB-spline regression model of order $k = 4$. Libois and Verardi (2012) make this procedure available for *Stata* users.[10]

The estimations consider different control variables, which are selected based on the existing empirical literature on spatial inequalities. In order to reduce a potential bias from heterogeneity in the territorial level between countries, I consider as control variables: (1) the ln of the number of sub-national units, which has been used to calculate the inequality measures, (2) the ln of country size in square kilometers, and (3) the ratio of both variables reflecting average sub-national unit size. These controls should reduce potential measurement errors in my measure of spatial inequality. Note that I also provide regression results using a higher territorial level (NUTS1 instead of NUTS2) in the robustness section. In addition to this, I also consider variables which are not directly related to geography that have been shown to impact regional inequalities in the literature. An important determinant of spatial inequality may be the heterogeneity of the population living in the different parts of a country, since the different regions are often inhabited by different ethnic groups. Think of Belgium, with the Dutch-speaking Flemings living in the north and the French-speaking Walloons in the south, or India, with the Indo-Aryans in the north and Dravidians in the south. As discussed in Kanbur and Venables (2005b), ethnic fractionalization may result in ethnic discrimination or conflict, promoting the divergence of regions. Thus, I control for the degree of ethnolinguistic fractionalization as calculated by Alesina et al. (2003). Theoretical studies such as Krugman and Elizondo (1996) and Zeng and Zhao (2010) as well as empirical studies such as Rodriguez-Pose and Gill (2006) and Rodríguez-Pose (2012) suggest that trade openness affects spatial inequality. Therefore, I control for the sum of imports and exports as a share of the GDP. To capture agglomeration effects, I control for the share of the population living in urban areas [see Venables (2005)].[11] Last but not least, I draw on the literature on decentralization and spatial inequality and include a dummy variable for federal countries [see, e.g., Lessmann (2009, 2012) and Rodríguez-Pose and Ezcurra (2010) for recent contributions]. Note that I have also considered measures of institutional quality and continental dummies, which did not turn out to have significant effects.[12] Summary statistics of all variables considered in the analysis provides Table A.3 in the Appendix.

Table 2
Cross-section data: parametric estimates.

Dependent variable: Weighted coefficient of variation of regional GDP p.c. (*WCV*), mean 2000–2009	(1)	(2)	(3)	(4)	(5)
ln(GDP p.c.)	−0.098***	0.153	0.293	0.330*	3.864**
	(−5.77)	(0.74)	(1.51)	(1.68)	(2.60)
(ln(GDP p.c.))^2		−0.015	−0.022*	−0.021*	−0.451**
		(−1.25)	(−1.98)	(−1.83)	(−2.44)
(ln(GDP p.c.))^3					0.017**
					(2.29)
ln(units)			0.056	0.103**	0.117**
			(1.00)	(2.08)	(2.36)
ln(area)			0.006	0.039**	0.029*
			(0.43)	(2.55)	(1.92)
ln(area)/ln(units)			−0.003	−0.001	0.003
			(−0.23)	(−0.17)	(0.28)
Ethnic				0.180*	0.162*
				(1.85)	(1.82)
Trade/GDP				0.003***	0.003***
				(4.98)	(4.75)
Urbanization				−0.003*	−0.003**
				(−1.82)	(−2.13)
Federal dummy				−0.092**	−0.071**
				(−2.38)	(−2.05)
Constant	1.200	0.157	−0.685	−1.746*	−11.14***
	(7.54)	(0.18)	(−0.79)	(−1.95)	(−2.90)
Obs. (N)	56	56	56	56	56
Adj. R-sq	0.428	0.433	0.486	0.663	0.687

Note: *t*-values are reported in parentheses; standard errors are calculated using White correction; ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

Table 3
Panel data: parametric estimates.

Dependent variable: Weighted coefficient of variation of regional GDP p.c.	Annual data			5-year period averages		
	(1)	(2)	(3)	(4)	(5)	(6)
(ln(GDP p.c.))	0.051	0.345**	−0.170	0.053	0.257**	−0.779
	(1.16)	(2.42)	(−0.18)	(1.22)	(2.11)	(−0.97)
(ln(GDP p.c.))^2		−0.018**	0.047		−0.014*	0.119
		(−2.07)	(0.38)		(−1.75)	(1.11)
(ln(GDP p.c.))^3			−0.003			−0.006
			(−0.50)			(−1.18)
Trade/GDP	0.001	0.001	0.001	0.001*	0.001*	0.001*
	(1.51)	(1.57)	(1.59)	(1.84)	(1.93)	(1.90)
Urbanization	−0.004	−0.005	−0.005	−0.003	−0.004	−0.004
	(−1.34)	(−1.61)	(−1.59)	(−1.27)	(−1.66)	(−1.66)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Obs	915	915	915	207	207	207
N	56	56	56	56	56	56
Adj. R-sq	0.333	0.363	0.365	0.413	0.434	0.444

Note: *t*-values are reported in parentheses; standard errors are calculated using White correction; ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

4.2. Parametric regressions

4.2.1. Cross-section data

In this section, I focus on the parametric regression analysis using cross-section data. The results are presented in Table 2. At this, I use averages of all variables in OLS regressions using the period 2000–2009. I have refrained from using a longer period for averaging because the data of middle- and low-income countries is largely confined to recent years, so that this sample is the most complete one. I present the results of five different specifications of Eq. (2). The first specification is a bivariate regression where spatial inequalities are regressed on the development indicator. The second specification considers an additional quadratic term of the development indicator. The third specification adds the territorial control variables. The forth specification considers the full set of control variables. The fifth specification considers an additional cubic term.

[10] See Desbordes and Verardi (2012) for further details on the estimation procedure, and an application in a context of interpersonal income inequality and development.

[11] Following Lessmann (2009), I have also used a concentration measure of the population within countries, which does not turn out to affect the regional inequality significantly.

[12] The results are available from the author upon request.

Table 4
Cross-section data: semiparametric estimates (linear part of the model).

Dependent variable: Weighted coefficient of variation of regional GDP p.c. (mean 2000–2009)	
ln(spatial units)	0.122***
	(2.69)
ln(area)	0.029*
	(1.96)
ln(area)/ln(spatial units)	0.004
	(0.39)
Ethnic fractionalization	0.165*
	(1.95)
Trade/GDP ratio	0.003***
	(4.35)
Urbanization	−0.003**
	(−2.13)
Federal dummy	−0.076*
	(−1.72)
Obs. (N)	56
Adj. R-sq	0.367

Notes: *t*-values are reported in parentheses; ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

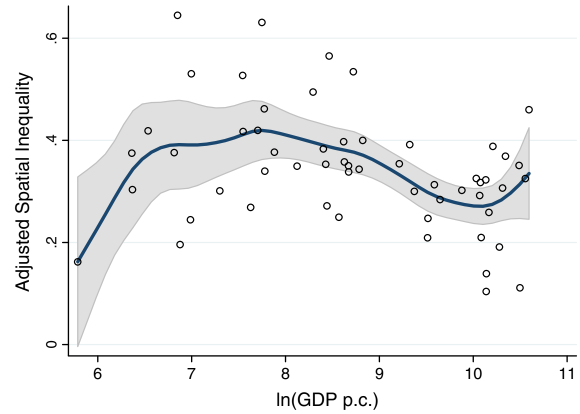

Fig. 4. Partial fit of the relationship between spatial inequality and economic development in cross-section data. Note: Plot of the estimated partial-regression functions. The inequality measure has been adjusted for the effects of the linear explanatory variables in the model [see Eq. (3)]. Shaded areas correspond to 90% confidence bands.

Column (1) reports results of a bivariate model without any control variables, showing that more highly developed countries (in terms of the natural logarithm of the GDP p.c.) have lower spatial inequalities. In column (2) the GDP enters in a quadratic form [$k = 2$ in Eq. (2)]. The coefficients have the expected signs – β_1 is positive and β_2 is negative – but neither of these coefficients is statistically significant, nor is their joint effect [see Brambor et al. (2006) on how to calculate marginal effects in interaction models]. In column (3) the three control variables, which control for spatial effects (ln of the number of spatial units, ln of the total area in square kilometers, and average unit size), are added as explanatory variables. Thereby, the coefficients β_1 and β_2 reach significance or are close to it. In column (4) all discussed control variables enter the regression. Both main coefficients of interest are significant and also the control variables show the expected signs. This specification of the model supports the hypothesis of an inverted-U-shaped relationship between spatial inequality and economic development. However, as the discussion in Section 2 has shown, there might be increasing spatial inequalities at higher levels of economic development. I thus consider a third order polynomial in the estimations reported in column (5). All main coefficients of interest are statistically significant in these regressions. The signs of the coefficients β_1 and β_2 remain unchanged, and the sign of β_3 is positive, implying that spatial inequality increases after the inverted-U pattern has been completed, and thus supporting earlier findings of Amos (1988) and others.

Based on the estimated coefficients reported in Table 2, I can also quantify the effects. The estimation equation can be simplified to $WCV = \alpha + \beta_1 Y + \beta_2 Y^2 + \beta 3Y^3 + \gamma X + \epsilon$ with X as vector of control variables. The local maxima and minima can be calculated through: $\partial\, WCV/\partial\, Y = \beta_1 + 2\beta_2 Y + 3\beta_3 Y^2 = 0$. Based on the estimations reported in column (5), the calculations indicate that spatial inequalities increase up to an income of ln(GDP p.c.) = 7.5, which corresponds to a GDP p.c. of approx. 2000 US$, and decrease until the threshold of ln(GDP p.c.) = 10.1 is reached, which corresponds to a GDP p.c. of approx. 24,000 US$. Beyond this threshold, spatial inequalities increase again.

My interpretation of the results is the following: The process of industrialization leads first to greater regional inequalities; yet beyond a certain critical GDP level, further increases in economic development lead to less inequality. The analysis also shows that regional inequalities start to increase again at later stages of economic development. My result is consistent with previous studies of (income) inequality, which also tends to increase at very high levels of economic development [e.g., Amos (1988), Ram (1991), or List and Gallet (1999)]. As discussed by List and Gallet (1999), one explanation for the renewed positive relationship between inequality and development may be the process of tertiarization, in which the economies shift from a manufacturing base towards a modern service base.

4.2.2. Panel data

The cross-section evidence reported in the previous section supports the inverted-U hypothesis, but – in addition – suggests that regional inequalities rise again at very high levels of economic development. However, the empirical approach has focused on between-country variations, while the theory of Kuznets and Williamson is ultimately related to the development of spatial inequalities within a country in time. This effect can be estimated in a panel of countries. Moreover, a panel data analysis allows us to consider country fixed effects, thereby eliminating unobserved heterogeneity between countries.

In the panel regressions, I consider annual data in the original frequency as well as 5-year period averages. Using period averages is particularly important, since the results of the regressions using annual data may be biased by business cycle effects. The following example should illustrate the problem. Consider a country consisting of two regions, a rich one and a poor one. Suppose further that the rich region has an industry which depends to a large extent on exports. Therefore, the income in the rich region is highly volatile over the business cycle, while the poor region has a less productive and also less volatile economy. In such a situation, one might expect increasing regional inequalities in times of expansion, and decreasing inequalities in times of recession. This is for example the case in Germany, where the eastern part of the country has much less volatile growth rates than the western states, which are richer at the cost of higher volatility. The estimation results are provided in Table 3.

Columns (1)–(3) report the results using annual data; columns (4)–(6) report the results using 5-year period averages. Each regression considers country and time fixed effects. Note that only the time-varying control variables (trade/GDP and urbanization) can be considered in the estimation. In columns (1) and (4) I consider a linear relationship between income and inequality, in columns (2) and (5) I add a quadratic term of the income variable, and in columns (3) and (6) I consider a cubic relationship. The panel regressions support the inverted-U hypothesis. The coefficient of the income variable is positive and significant in the results reported in columns (2) and (4), while the coefficient of the quadratic term is negative and significant as well. Note that the cubic term is highly insignificant. Concerning the size of the effects, there are no noteworthy differences between the annual panel and the period averages. Interestingly, the adjusted R square is slightly larger using 5-year averages suggesting that the model fits better to medium

Table 5
Panel data: semiparametric estimates (linear part of the model).

Dependent variable: Weighted coefficient of variation of regional GDP p.c. (*WCV*)	Annual data	5-year averages
	(1)	(2)
Trade/GDP ratio	0.001	0.001*
	(0.13)	(1.93)
Urbanization	−0.007 ***	−0.006***
	(−3.20)	(−3.17)
Country fixed effects	yes	yes
Time fixed effects	yes	yes
N	858	151
Adj. R-sq	0.054	0.396

Note: *t*-values are reported in parentheses; ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

term instead of short term relationships. To sum up, the panel estimations which focus on the within-country variation in the data support the inverted-U hypothesis, but inequalities do not seem to increase again at higher levels of economic development. Note, however, that this finding depends on the logarithmic transformation of income data [see Section 4.4.3 for details].

4.3. Semiparametric regressions

4.3.1. Cross-section data

This section presents estimation results using a semiparametric regression procedure. To stick to the structure of the previous section, I first focus on a cross-section of countries, and subsequently present the panel estimates. As discussed in Section 1, I estimate the semiparametric Eq. (3) using the procedure proposed by Robinson (1988). The estimation output consists of two parts: a table which reports the regression coefficients of the linear part of the model, and a graph which illustrates the functional form of the nonlinear part, that is, the relationship between spatial inequality and development. Table 4 reports the corresponding results.

The size and significance of the control variables, which enter the regression linearly, are comparable to the parametric estimates. Interestingly, the territorial control variables also show significant effects implying that there is some heterogeneity in the territorial classification between countries, which has to be considered in the analysis. The nonparametric part of the estimation is illustrated in Fig. 4 including confidence bands for the 90% confidence level.

The graph supports the main hypothesis of an inverted-U-shaped relationship between spatial inequality and development, and it also shows increasing inequalities at high levels of economic development. The thresholds of income, where the relationship between spatial inequality and development reverses, are similar to those calculated in the parametric regression. All in all, the semiparametric regression based on the cross-section data set support the parametric results.

4.3.2. Panel data

The last step of this analysis is to consider the full panel data set in the semiparametric approach, which uses the estimator proposed by Baltagi and Li (2002). The results of the linear part of the model are presented in Table 5. Again, I present two different specifications: the first specification uses annual data, while the second specification uses a 5-year period averages. In both specifications, I linearly control for the time-varying control variables and for country and time fixed effects.

The signs of the coefficients of the control variables are similar to those in the parametric panel regressions as reported in Table 3. But in contrast to the results of the parametric regressions now also the degree of urbanization shows a significant negative effect. The functional form

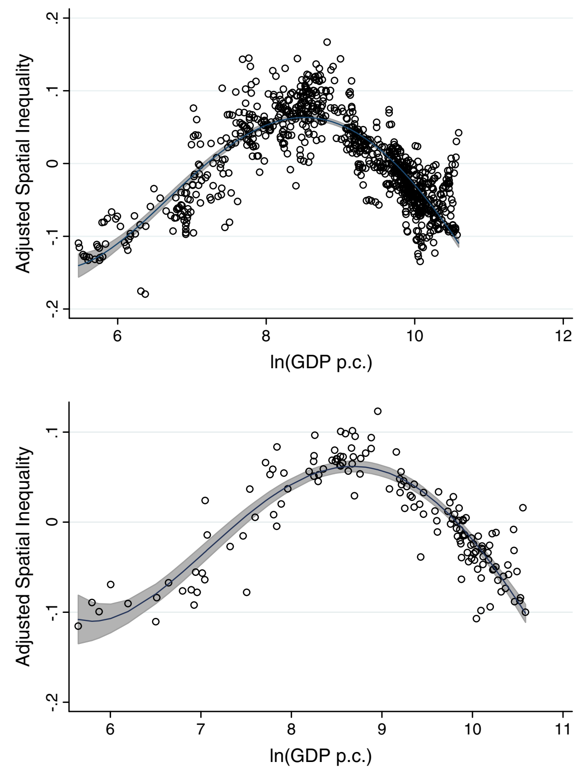

Fig. 5. Partial fits of the relationship between spatial inequality and economic development in panel data. Note: Plots of the estimated partial-regression functions. The points in each graph are partial residuals for spatial inequality; the inequality measure has been adjusted for the effects of the other explanatory variables in the model [see Eq. (**3**)] and the partial residuals are centered around the mean. Shaded areas correspond to 90% confidence bands.

of the relationship between spatial inequality and development is shown by Fig. 5.

The graph reveals an inverted-U-shaped relationship. There are no significant differences between the specification which uses annual data (upper part) compared to the specification which uses period averages (lower part). Importantly, the estimations do not indicate that regional inequalities increase again at high levels of economic development.

4.4. Robustness tests

4.4.1. Capital cities and regions

I provide a number of robustness tests in this section. A first test is concerned with the effect of capital cities or regions on spatial inequality. In their study on growth in regions, Gennaioli et al. (2013b) find a strong result that the regions that include a national capital grow much faster compared to the rest of the country. Moreover, the results of Ades and Glaeser (1995) suggest that the effect of national capitals on agglomeration and development might be even larger in developing countries. They show that political leaders in countries with a low level of democracy and high political instability tend to concentrate resources in single, large urban areas, which than dominate the structure and prospects of the whole economy. A good example for this issue is Argentina, where about one third of the population is concentrated in the metropolitan region of Buenos Aires which is three times as rich as the country average.

In order to check whether my results concerning the inverted-U hypothesis are sensitive to the dominating role of capital cities or regions, I calculate the inequality measures excluding these particular regions. Let me give some examples: in Great Britain I exclude the two NUTS2 regions Inner and Outer London, in Japan I exclude Southern-Kanto where Tokyo is located, etc. Fig. 6 compares the *WCV* using all

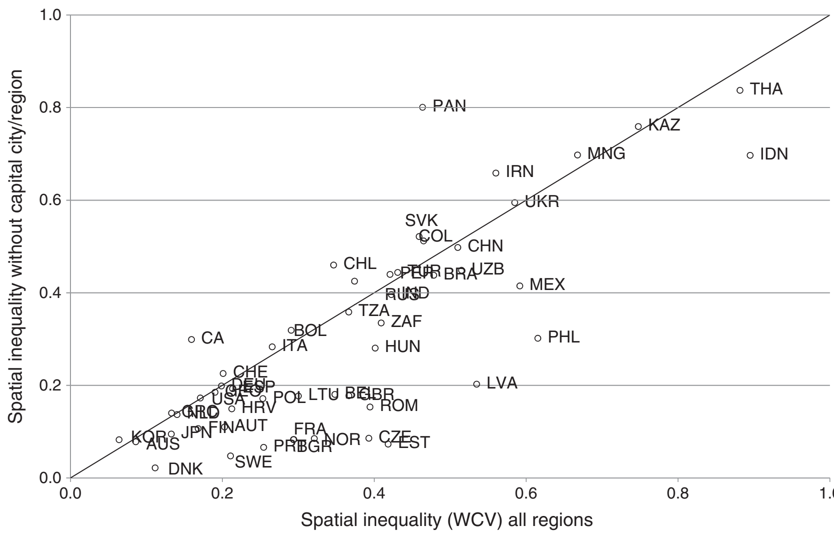

Fig. 6. Spatial inequality with and without capital cities or regions.

regions (abscissa) with the *WCV* excluding capital cities or regions (ordinate). I also include the bisecting line where both inequality measures are identical.

The calculations show some interesting results, which are in line with the findings of Gennaioli et al. (2013b). In about two third of the countries considered, spatial inequality is higher if I include the capital region. Capital regions are often significantly richer than the country average, therefore excluding them usually reduces the inequality measure. Some extreme cases are the transition economies in Eastern Europe, France or the Philippines. For example in France the *WCV* is 0.29 including Paris, which is a value significantly above the OECD average. Concentrating on spatial inequality in the rest of the country reveals that the remaining regions are highly homogeneous with a *WCV* of just 0.08. By contrast, in the United States the effect of Washington D.C. on the measure of spatial inequality is negligible. Although the examples point at strong effects of capitals on spatial inequality in individual cases, the correlation between the inequality measures is very high with a correlation coefficient of 0.81.

In my regression analysis, a part of the effect of capital cities on spatial inequality is captured by the spatial controls and the controls for agglomeration. If I regress the absolute value of the difference of the two inequality measures on these controls, I find that the deviation in inequality measures is smaller in large countries, larger in highly urbanized countries, and smaller in federal countries. These suggestive results fit well to the results of Ades and Glaeser (1995). Developing countries with weak institutions are usually highly centralized. Centralized countries tend to concentrate the economic activity in the capital region, while decentralized economies are less depending on single agglomeration centers. The problem is that income is also correlated with these variables. Therefore, it remains to study whether the inverted-U hypothesis still holds if I exclude the capital cities or regions. The results of cross country regressions similar to those presented in Section 1 are reported in Table A.5 in the Appendix.[13] Again, there is a robust evidence of an inverted-U relationship between spatial inequality and development. The only noteworthy change concerns the last specification which tests for a cubic relationship. Disregarding capital cities shows that spatial inequality does not increase at higher levels of development.

4.4.2. Sample adjustments and additional controls

The next set of robustness tests is concerned with the particular measurement of spatial inequality, sensitivity to the country sample, and sensitivity to the inclusion of additional control variables. Table A.6 in the Appendix reports the results. The regression results shown in the upper part of the table (columns 1–18) deal with the heterogeneity of territorial levels and alternative measures of spatial inequality, the results reported in the lower part (columns 19–35) concentrate on sample adjustments an additional control variables. Each robustness test considers three specifications of the cross-section model. The first regression tests of an inverted-U including the territorial control variables, the second adds the other control variables, and the third considers a cubic function.

4.4.2.1. Upper part of Table A.6. Columns (1)–(3) report results based on estimations without logarithms of the territorial control variables. Columns (4)–(6) report results using NUTS1 instead of NUTS2 level data for European countries as basis for the calculation of spatial inequalities. This might be seen as an important robustness test, since the different territorial classification might have an effect on the results, in particular if the political authorities of sub-national governments are more relevant on higher aggregation levels than on lower ones. Note that the correlation between the weighted coefficient of variation based on NUTS1 and NUTS2 level data is 0.95 for EU countries.[14] Consequently, the main result concerning the relationship between spatial inequalities and development is independent from the territorial classification. Columns (7)–(9) report results using only those countries, which have at least five sub-national units. Thereby, very small countries such as Malta or Ireland are dropped from the sample. Columns (10)–(12) report results for countries with less than 30 sub-national units. The last two robustness tests concern the inequality measure. All presented estimations up to this point of the paper have used the population-weighted coefficient of variation as discussed in Section 3. This measure is commonly used in the economic-geography literature, but studies on growth and convergence such as, e.g., Barro and Sala-i-Martin (1992) concentrate on unweighted measures. One might argue that weighting of observations by the (population-)size of the regions

[13] Semi-parametric estimations and regressions with panel data yield similar results.

[14] I have also considered measures of spatial inequality calculated based on the regional data, which is used in a study of human capital and regional development by Gennaioli et al. (2013a). At this, I want to thank the authors for sharing their data set with me. Gennaioli et al. (2013a) use the territorial classification at that sub-national level, which has the greatest political power. For my sample of countries, the main finding remains unchanged. The results are available from the author upon request.

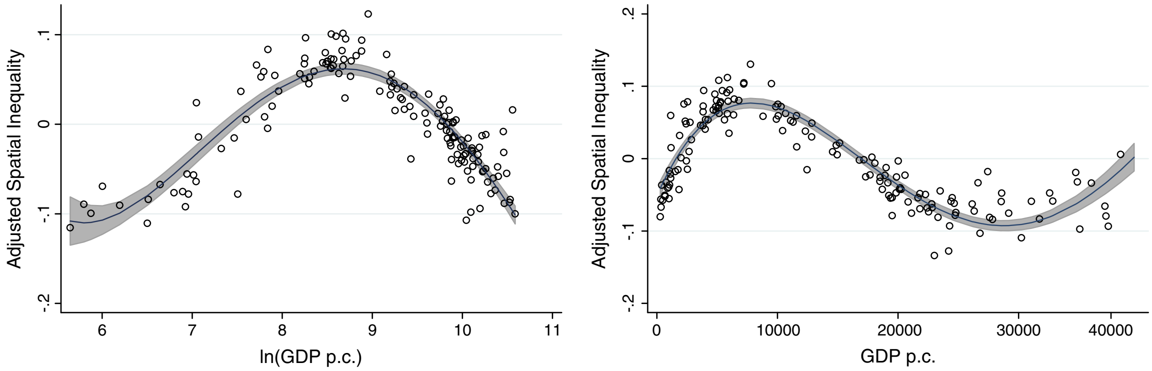

Fig. 7. Robustness test: panel data with 5-year averages. The estimation reported in the left panel considers the GDP p.c. in logarithmic transformation, the estimation reported in the right panel considers the GDP p.c.

distorts the inequality measure. Small regions, which may be extremely rich (e.g., capital regions) or poor (e.g., special zones of ethnic minorities), have only a minor effect on the overall indicator, although the deviations from the respective countries' means are very important in light of the risk of conflict and secession. To allow for this argument and to make my results comparable to the convergence literature, I also calculated the (unweighted) coefficient of variation (*CV*) of the regional GDP p.c. for each country as well as the Gini coefficient. The results are reported in columns (14)–(15) and (16)–(18), respectively. The main finding concerning the inverted-U remains robust.

4.4.2.2. Lower part of Table A.6. Columns (19)–(21) report results excluding the poorest countries (GDP p.c. < 1000 US$), columns (22)–(23) reports results excluding the richest countries (GDP p.c. > 30,000 US$). In particular, excluding of the poorest countries is an important test, since the quality of regional data is questionable. It turns out, that the effect of development on inequality has a cubic functional form, while the inverted-U disappears. This is not surprising, since the upswinging part of the Kuznets curve is based on the data of the excluded countries. Columns (25)–(27) report results excluding transition economies as defined by the IMF [see IMF (2000)]. Columns (28)–(30) report results excluding former Soviet countries. The reason for this robustness check is that those countries have faced shard increases in spatial inequalities, which are caused by the rapid structural changes in the economies. It is unlikely, that these changes in spatial inequalities are related to the sectoral change discussed by Kuznets (1955) and Williamson (1965). Columns (31)–(33) and (34)–(36) report results including dummies for EU membership (EU27 and EU15). The European Union pays a significant amount of transfers to lagging regions in order to promote regional convergence. If these regional policies are successful, spatial inequalities can be expected to be lower within EU countries. The coefficient of the EU27 dummy is indeed negative, but not statistically significant. The main finding concerning the relationship between spatial inequalities and development remains stable.

Table 6
Cross section: parametric estimates using sectoral data.

Dependent variable: Weighted coefficient of variation of regional GDP p.c.					
	(1)	(2)	(3)	(4)	(5)
Non-agricultural GVA/GDP	−0.009***	0.033	0.042**	0.035**	0.154
	(−4.00)	(1.66)	(2.13)	(2.23)	(1.14)
(Non-agricultural GVA/GDP)^2		−0.001**	−0.001**	−0.001**	−0.002
		(−2.11)	(−2.51)	(−2.23)	(−1.02)
(Non-agricultural GVA/GDP)^3					0.001
					(0.89)
ln(units)			0.071	0.113**	0.104*
			(1.12)	(2.18)	(1.92)
ln(area)			−0.013	0.037*	0.039*
			(−0.70)	(1.76)	(1.85)
ln(area)/ln(units)			−0.005	−0.002	−0.003
			(−0.47)	(−0.22)	(−0.27)
Ethnic				0.219**	0.228**
				(2.38)	(2.41)
Trade/GDP				0.003***	0.003***
				(5.30)	(5.33)
Urbanization				−0.003**	−0.003**
				(−2.31)	(−2.34)
Federal dummy				−0.105***	−0.109**
				(−2.81)	(−2.69)
Constant	1.085***	−0.400	−0.727	−1.671***	−4.307
	(5.68)	(−0.57)	(−1.07)	(−2.89)	(−1.44)
Obs. (N)	55	55	55	55	55
Adj. R-sq	0.239	0.276	0.332	0.642	0.640

Note: *t*-values are reported in parentheses; standard errors are calculated using White correction; ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

4.4.3. ln(income) as development indicator

In explaining the determinants of spatial inequalities, this paper focuses on the natural logarithm of the GDP p.c. as development indicator. This is in line with highly influential studies on interpersonal inequality and development such as Barro (2000), Ravallion (2001), or Easterly (2001), and recent contributions such as Desbordes and Verardi (2012). However, one might reasonably argue that using a logarithmic transformation of the income data might have an effect on the functional form of the relationship between spatial inequality and development. Moreover, the theoretical considerations of Kuznets (1955) and Williamson (1965) suggest that structural changes cause the inverted-U, rather than income growth per se. Therefore, I would like to provide some evidence, that the finding of an inverted-U in spatial inequalities is robust to: (1) the logarithmic transformation, and (2) the use of data on the sectoral structure of the economy.

The logarithmic transformation is commonly used in macroeconomic studies. This can be justified by the much higher correlation between the logarithm of income and other potential indicators of development (e.g., the infant mortality rate, the primary school enrollment rate, etc.), as if the GDP p.c. is taken into consideration.[15] In my analysis the logarithmic transformation implies a downward compression of the data of the richest countries, which may affect the result. To illustrate this effect, I have estimated my semi-parametric model using the 5-year period averaged panel data using the ln(GDP p.c.) as well as the GDP p.c. without this transformation. The result is presented in Fig. 7. In both graphs, we observe the inverted-U pattern. However, there is no sign of increasing inequalities at high levels of economic development if logarithms are considered (left panel), while the use of original GDP data shows that inequalities increase slightly at high levels of economic development. Therefore, the results concerning the inverted-U hypothesis is

[15] For example, the correlation between the ln(GDP p.c.) an the infant mortality rate is 0.77 based on the cross-section data, while the correlation between the GDP p.c. and the infant mortality rate is just 0.42.

not affected from the transformation of income data, while the increase at high levels of economic development depends on these measurement issues.

The theory of Kuznets (1955) and Williamson (1965) is based on structural changes in the economy, i.e. the shift from agriculture to industry. Structural changes are usually accompanied by an increase in income, since a new technology would not be adopted, if income in the new industrial sector (or the modern service sector) is similar to the old agricultural sector. Therefore, studies like this use income as a proxy for sectoral change. But is it really the structural change that causes the inverted-U observed in the data? To test this, I have also used data on the sectoral structure of the economies. Thereby, I focus on the gross value added (GVA) by modern sectors (manufacturing + service) as a share of the GDP. I aggregate the modern sectors, since some high income countries specialize in the industrial sector (e.g. Germany), while others specialize in a modern service sector (e.g. Great Britain). The results of parametric regressions using cross-country data are presented in Table 6.[16]

Interestingly, the regressions show that spatial inequalities increase as the relative size of the modern sectors increases, than spatial inequalities decrease again.[17] This supports the transmission channel suggested by Kuznets (1955) and Williamson (1965).

5. Summary and conclusion

This paper studies the hypothesis of an inverted-U-shaped relationship between spatial inequality and economic development. The theory of Kuznets (1955) and Williamson (1965) suggests that (spatial) inequality first increases \in the process of development, peaks, and then decreases. With the exception of the initial study of Williamson (1965) himself, empirical evidence for this hypothesis, which covers developing and developed economies, does not exist. This gives reason to reexamine the original work using a broader data set as well as recent econometric techniques. A further reason – and perhaps the most important one – is that something might have changed in the relationship between spatial inequality and development since the 1950s or 1960s, in that most countries have experienced very dynamic growth since then. For this purpose, a unique panel data set was collected covering 56 countries at all stages of economic development. The period covered by the unbalanced panel is 1980–2009. Cross-country and panel regressions have been carried out using a parametric as well as a semiparametric approach. I find clear evidence of an inverted-U in models focusing on between- and within-country variations. There is also some evidence that spatial inequalities increase again at very high levels of economic development, but this finding is sensitive to measurement issues and the econometric methodology.

What do we learn from this study? When countries shift from agricultural to industrial economies, spatial inequalities increase. If a certain development level is reached, the relationship is reversed until high levels of economic development are reached. The good news is that spatial inequalities are thus a temporal phenomenon, since they tend to diminish in time. But within this process, spatial inequalities might be harmful for the development process itself. Recent studies on internal conflicts – such as Deiwiks et al. (2012), Buhaug et al. (2012), and Lessmann (2013b) – show that spatial inequalities increase the risk of civil wars thereby harming development. Unfortunately, what we do not know is, which level of spatial inequalities is acceptable for a country and which is not. The "tunnel effect" as discussed by Hirschman and Rothschild (1973), suggests that people accept inequalities as long as they believe that they can participate in the economic progress in the future. But if these beliefs are disappointed, people demand for redistribution or they try to take their piece of the cake forcibly. In light of this, policies are justifiable, which aim at a more equal factor distribution among the different regions of a country. The results of Kessler et al. (2011) suggest that unconditional equalization transfers are not a suitable instrument to reduce regional inequalities. But a recent study on regional development by Gennaioli et al. (2013a) suggests that the differences in the human capital endowment are one of the major determinants of the regional differences in development, therefore governments may try to help lagging regions by investing in human capital.

[16] The regressions are based on only 55 countries, since sectoral data is missing in case of Malta.

[17] Note that this result also holds if the ln(GDP p.c.) is added to the list of control variables.

Appendix A

Table A.1
Sources of regional data by country.

Country	Territorial level	Period	Source
Argentina	23 provinces; 1 capital region	1991–2002	Dirección Provincial de Estadística
Australia	8 TL2 regions	1990–2008	OECD Regional Statistics
Austria	9 NUTS2 regions	1980–2004	Cambridge Econometrics
Belgium	11 NUTS2 regions	1980–2004	Cambridge Econometrics
Bolivia	9 departments	1988–2009	Instituto Nacional de Estadistica
Brazil	26 states; 1 federal district	2002–2007	Instituto Brasileiro de Geografica e Estatistica
Bulgaria	6 TL2 regions	1995–2007	OECD Regional Statistics
Canada	12 provinces and territories (Northwest Territories including Nunavut)	1981–2004	Statistics Canada
Chile	13 regions	1996–2006	National Statistics Institute
China	30 provinces, autonomous regions, and cities	1994–2008	National Bureau of Statistics of China
Colombia	33 departments	1990–2007	Departamento Administrativo Nacional de Estadística
Croatia	Nuts3 regions	2001–2007	Eurostat
Czech Republic	8 TL2 regions	1990–2007	Cambridge Econometrics and OECD Regional Statistics
Denmark	3 NUTS2 regions	1980–2004	Cambridge Econometrics
Estonia	5 NUTS3 regions	1996–2009	Eurostat
Finland	6 NUTS2 regions	1980–2004	Cambridge Econometrics
France	22 NUTS2 regions	1980–2004	Cambridge Econometrics
Georgia	9 provinces	2003–2009	National Statistics Office of Georgia
Germany	30 NUTS2 regions	1991–2004	Cambridge Econometrics
Greece	13 NUTS2 regions	1980–2004	Cambridge Econometrics
Hungary	7 NUTS2 regions	1990–2009	Cambridge Econometrics and OECD Regional Statistics
India	28 states and union territories	1980–2005	Directorate of Economics & Statistics of respective State Governments, and Central Statistical Organisation
Indonesia	33 provinces	2004–2008	Badan Pusat Statistik
Iran, Islamic Rep.	28 provinces	2000–2003	Statistical Center of Iran
Ireland	2 NUTS2 regions	1980–2004	Cambridge Econometrics
Italy	20 NUTS2 regions	1980–2004	Cambridge Econometrics
Japan	10 TL2 regions	1990–2005	OECD Regional Statistics
Kazakhstan	16 regions and cities	1998–2009	Agency of Statistics of the Republic of Kazakhstan
Korea, South	7 TL2 regions	1990–2007	OECD Regional Statistics
Latvia	6 NUTS3 regions	1996–2007	EUROSTAT
Lithuania	10 NUTS3 regions	1995–2007	EUROSTAT
Malta	2 NUTS3 regions	2000–2007	EUROSTAT
Mexico	32 states; 1 capital region	1980–2006	Instituto Nacional de Estadística y Geografía, and OECD Regional Statistics
Mongolia	21 provinces; 1 capital region	2000–2006	National Statistical Office
Netherlands	12 NUTS2 regions	1986–2004	Cambridge Econometrics
New Zealand	2 TL2 regions	2000–2003	OECD Regional Statistics

Table A.1 (*continued*)

Country	Territorial level	Period	Source
Norway	7 NUTS2 regions	1980–2004	Cambridge Econometrics
Panama	9 provinces	2002–2007	Instituto Nacional De Estadistica
Peru	24 departments	2001–2009	Instituto Nacional de Estadística e informática — Dirección Nacional de Cuentas Nacionales
Philippines	17 districts	2002–2008	National Statistics Office
Poland	16 NUTS2 regions	1990–2007	Cambridge Econometrics and OECD Regional Statistics
Portugal	7 NUTS2 regions	1980–2004	Cambridge Econometrics
Romania	8 NUTS2 regions	1995–2007	EUROSTAT
Russian Federation	7 federal regions	1998–2008	Federal State Statistics Office
Slovak Rep.	4 TL2 regions	1990–2007	Cambridge Econometrics and OECD Regional Statistics
Slovenia	2 NUTS2 regions	1995–2007	EUROSTAT
South Africa	9 provinces	2001–2008	Statistics South Africa
Spain	18 NUTS2 regions	1980–2004	Cambridge Econometrics
Sweden	8 NUTS2 regions	1980–2004	Cambridge Econometrics
Switzerland	7 NUTS2 regions	1980–2004	Cambridge Econometrics
Tanzania	21 administrative regions	2002–2007	National Bureau of Statistics
Thailand	7 geographic regions	2001–2009	National Statistics Office Thailand
Turkey	26 TL2 regions	1990–2006	OECD Regional Statistics
U.S. of America	51 states	1980–2008	U.S. Department of Commerce, OECD Regional Statistics
Ukraine	27 districts	2004–2008	State Statistics Committee of Ukraine
United Kingdom	37 NUTS2 regions	1980–2004	Cambridge Econometrics
Uzbekistan	12 provinces; 1 republic; 1 capital region	2008	Uzbekistan in Figures — UinF
Venezuela	23 states; 1 federal district	2007	Banco Central de Venezuela

Table A.2
Descriptive statistics: territorial classification.

Country	Area (km^2)	Spatial units (no.)	Area per unit (km^2)	Population per unit (people)
Australia	7,680,000	8	960,000	2,547,500
Austria	82,500	9	9167	910,333
Belgium	30,300	11	2755	952,727
Bolivia	1,080,000	9	120,000	1,010,222
Brazil	8,460,000	27	313,333	6,837,037
Bulgaria	109,444	6	18,241	1,295,833
Canada	9,090,000	12	757,500	2,682,500
Chile	744,000	13	57,231	1,246,923
China	9,330,000	30	311,000	43,300,000
Colombia	1,110,000	33	33,636	1,294,546
Croatia	54,622	3	18,207	1,479,000
Czech Republic	77,300	8	9663	1,285,000
Denmark	42,400	3	14,133	1,806,667
Estonia	42,400	5	8480	270,000
Finland	304,667	6	50,778	874,500
France	548,000	22	24,909	2,761,364
Georgia	69,500	9	7722	499,444
Germany	349,000	30	11,633	2,744,000
Greece	129,000	13	9923	852,308
Hungary	89,600	7	12,800	1,444,286
India	2,970,000	28	106,071	38,800,000
Indonesia	1,810,000	33	54,848	6,596,970
Iran, Islamic Rep.	1,630,000	28	58,214	2,446,786
Ireland	68,900	2	34,450	2,067,000
Italy	294,000	20	14,700	2,919,000
Japan	365,000	10	36,500	12,800,000
Kazakhstan	2,700,000	16	168,750	950,625
Korea, Rep. (South)	97,044	7	13,863	6,858,572
Latvia	62,267	6	10,378	385,167
Lithuania	62,700	10	6270	342,200
Malta	320	2	160	201,250
Mexico	1,940,000	32	60,625	3,203,125
Mongolia	1,550,000	22	70,455	115,046
Netherlands	33,800	12	2817	1,353,333
New Zealand	268,000	2	134,000	2,047,000
Norway	304,000	7	43,429	661,143
Panama	74,300	9	8256	355,889
Peru	1,280,000	24	53,333	1,151,250
Philippines	298,000	17	17,529	4,985,294
Poland	304,667	16	19,042	2,386,875
Portugal	91,500	7	13,071	1,495,714
Romania	230,000	8	28,750	2,717,500
Russian Federation	16,400,000	7	2,342,857	20,500,000
Slovak Republic	48,100	4	12,025	1,348,000
Slovenia	20,100	2	10,050	1,003,000
South Africa	1,210,000	9	134,444	5,205,556
Spain	499,000	18	27,722	2,394,445
Sweden	410,000	8	51,250	1,130,000
Switzerland	40,000	7	5714	1,061,000
Tanzania	886,000	21	42,190	1,840,476
Thailand	511,000	7	73,000	9,334,286
Turkey	770,000	26	29,615	2,718,846
Ukraine	579,000	27	21,444	1,755,926
United Kingdom	242,000	37	6541	1,626,487
United States of America	9,160,000	51	179,608	5,776,471
Uzbekistan	425,000	14	30,357	1,865,714

Table A.3
Summary statistics.

Variable	Obs.	Mean	Stand. dev.	Min.	Max.
WCV	56	0.35	0.20	0.06	0.89
CV	56	0.38	0.23	0.09	1.23
Gini	56	0.20	0.10	0.06	0.46
Spatial units (no.)	56	14.64	10.88	2.00	51.00
Total area (1000 km^2)	56	1,554,597	3,179,162	320	16,400,000
Average unit size	56	119,169	345,182	160	2,342,857
Average population per unit	56	4,079,632	7,944,227	115,046	43,300,000
Population density	56	129.10	191.91	1.63	1257.81
ln(units)	56	2.39	0.81	0.69	3.93
ln(area)	56	12.69	1.95	5.77	16.61
ln(area)/ln(units)	56	6.06	2.83	3.43	18.03
Ethnic fractionalization	56	0.34	0.23	0.00	0.75
Trage/GDP	56	84.58	39.52	25.18	171.38
Urbanization	56	67.01	15.72	24.07	97.26
Federal dummy	56	0.21	0.41	0.00	1.00
Infant mortality rate	56	18.35	21.96	3.43	119.33
Inflation rate	56	6.08	5.91	−1.18	28.79
Years of democracy since 1800	56	65.86	52.60	0.00	210.00

Table A.4
Data sources & definitions.

Variable	Definition	Source
WCV	Population-weighted coefficient of variation of regional GDP per capita	Various sources
CV	Coefficient of variation of regional GDP per capita	Various sources
Gini	Adjusted Gini coefficient of regional GDP per capita	Various sources
GDP p.c.	GDP per capita in 2005 $ prices.	World Bank (2011)
spatial units	ln of the number of regions considered for the calculation of measures of regional inequality.	Various sources
Total area	Ln of area in square kilometers.	Weltbank (2011)
Trade/GDP	Sum of imports and exports (total trade) as a share of the GDP.	Weltbank (2011)
Ethnic fractionalization	Ethnolinguistic fractionalization is computed as one minus Herfindahl index of ethnolinguistic group shares, and reflects the probability that two randomly selected individuals from a population belonged to different groups.	Alesina et al. (2003)
Urbanization	Share of urban living population in total population.	Weltbank (2011)
Federal dummy	Dummy for countries with a federal constitution.	Treisman (2008)
Infant Mortality	Mortality rate under-5 (per 1000)	Weltbank (2011)
Inflation	Inflation, GDP deflator (annual %)	Weltbank (2011)
Democracy	Number of years of democracy since 1800. A country is assumed to be democratic if the Polity2 index provided by the PolityIV project is positive	See Marshall and Jaggers (2009)

Table A.5
Cross-section regressions excluding capital cities or regions.

Dependent variable: Weighted coefficient of variation of regional GDP p.c. (*WCV*), mean 2000–2009					
	(1)	(2)	(3)	(4)	(5)
ln(GDP p.c.)	−0.100***	0.156	0.343	0.375**	2.219
	(−6.19)	(0.66)	(1.63)	(2.09)	(1.21)
(ln(GDP p.c.))^2		−0.015	−0.025**	−0.022**	−0.247
		(−1.09)	(−2.02)	(−2.14)	(−1.09)
(ln(GDP p.c.))^3					0.010
					(0.98)
ln(units)			0.084	0.188**	0.192**
			(0.96)	(2.20)	(2.25)
ln(area)			0.017	0.051**	0.047**
			(0.94)	(2.67)	(2.45)
ln(area)/ln(units)			0.011	0.036	0.036
			(0.32)	(1.11)	(1.13)
ethnic				0.289***	0.276**
				(2.73)	(2.70)
trade/GDP				0.003***	0.003***
				(4.35)	(4.32)
urbanization				−0.004*	−0.004*
				(−1.96)	(−1.99)
federal dummy				−0.096**	−0.084*
				(−2.34)	(−1.95)
constant	1.162***	0.103	−1.229	−2.674***	−7.561
	(7.53)	(0.10)	(−1.45)	(−3.02)	(−1.58)
Obs. (N)	52	52	51	51	51
adj. R-sq	0.363	0.364	0.433	0.637	0.635

Note: *t*-values are reported in parentheses; standard errors are calculated using White correction; ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

Table A.6
Cross-section data: robustness test on different spatial issues.

	Robustness test: spatial issues																	
	Dependent variable: *WCV*												Dependent variable:			Dependent variable:		
	Area and units without ln			NUTS1 regions			Units > 4			Units < 30			*CV*			*GINI*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
ln(GDP p.c.)	0.281	0.302	4.738***	0.327*	0.376**	2.847**	0.309	0.376*	3.830**	0.281	0.319	4.397***	0.382	0.426	6.121***	0.253**	0.275***	2.359***
	(1.42)	(1.28)	(2.96)	(1.74)	(2.13)	(1.33)	(1.53)	(1.73)	(2.51)	(1.53)	(1.68)	(3.10)	(1.58)	(1.66)	(3.72)	(2.58)	(2.74)	(3.50)
$(\ln(\text{GDP p.c.}))^2$	−0.022*	−0.020	−0.560***	−0.023**	−0.023**	−0.324*	−0.023*	−0.023*	−0.444**	−0.021*	−0.020*	−0.516***	−0.028*	−0.026*	−0.719***	−0.017***	−0.017***	−0.270***
	(−1.90)	(−1.45)	(−2.84)	(−2.21)	(−2.26)	(−1.72)	(−2.00)	(−1.84)	(−2.33)	(−1.97)	(−1.81)	(−2.95)	(−2.00)	(−1.79)	(−3.50)	(−3.07)	(−2.93)	(−3.19)
$(\ln(\text{GDP p.c.}))^3$			0.022***			0.012*			0.017**			0.020***			0.028***			0.010***
			(2.70)			(1.58)			(2.16)			(2.80)			(3.30)			(2.90)
Units	0.004*	0.006**	0.006**	0.076	0.114*	0.113*	0.305	0.362*	0.336*	0.017	0.074	0.106**	0.093	0.137**	0.160***	0.022	0.045**	0.053**
	(1.97)	(2.34)	(2.22)	(1.20)	(2.00)	(1.88)	(1.58)	(1.83)	(1.76)	(0.25)	(1.49)	(2.13)	(1.37)	(2.47)	(2.87)	(0.77)	(2.09)	(2.64)
Area	−0.001	0.001	0.001	0.001	0.035**	0.029*	−0.052	−0.012	−0.013	0.012	0.048***	0.036**	0.014	0.045**	0.028	0.006	0.021***	0.015*
	(−1.48)	(1.05)	(0.32)	(0.06)	(2.18)	(1.84)	(−1.23)	(−0.26)	(−0.28)	(0.76)	(3.01)	(2.46)	(1.04)	(2.42)	(1.48)	(1.01)	(2.92)	(2.01)
Area/units	−0.007	−0.011**	−0.009*	0.001	−0.003	−0.001	0.119	0.123	0.107	−0.009	−0.005	0.001	0.007	0.007	0.014	0.004	0.005	0.007*
	(−1.09)	(−2.28)	(−1.77)	(0.01)	(−0.29)	(−0.11)	(1.22)	(1.33)	(1.20)	(−0.71)	(−0.65)	(0.10)	(0.48)	(0.74)	(1.37)	(0.51)	(1.00)	(1.79)
Ethnic		0.222**	0.183**		0.161	0.152		0.191*	0.170*		0.129	0.109		0.262**	0.233**		0.122***	0.112***
		(2.14)	(2.09)		(1.63)	(1.61)		(1.94)	(1.83)		(1.30)	(1.24)		(2.16)	(2.24)		(2.84)	(2.95)
Trade/GDP		0.002***	0.002***		0.003***	0.003***		0.003***	0.003***		0.003***	0.003***		0.003***	0.002***		0.001***	0.001***
		(3.27)	(3.35)		(5.21)	(4.81)		(4.12)	(4.09)		(4.94)	(4.72)		(3.83)	(3.66)		(4.80)	(4.78)
Urbanization		−0.003	−0.003*		−0.003**	−0.003**		−0.003*	−0.004**		−0.003*	−0.003**		−0.003*	−0.003**		−0.001*	−0.002**
		(−1.51)	(−1.78)		(−2.12)	(−2.26)		(−1.88)	(−2.07)		(−1.96)	(−2.44)		(−1.70)	(−2.07)		(−1.97)	(−2.34)
Federal dummy		−0.073	−0.038		−0.077**	−0.063*		−0.089**	−0.070*		−0.080**	−0.059*		−0.094*	−0.061		−0.034*	−0.021
		(−1.49)	(−0.82)		(−2.19)	(−1.91)		(−2.04)	(−1.77)		(−2.23)	(−2.01)		(−1.91)	(−1.46)		(−1.83)	(−1.29)
Constant	−0.436	−0.793	−12.710***	−0.834	−1.916**	−8.452**	−1.254	−2.572**	−11.660***	−0.612	−1.706*	−12.610***	−1.255	−2.344*	−17.480***	−0.813*	−1.341***	−6.880***
	(−0.51)	(−0.79)	(−3.00)	(−1.00)	(−2.36)	(−2.16)	(−1.26)	(−2.36)	(−2.90)	(−0.74)	(−1.97)	(−3.41)	(−1.18)	(−1.97)	(−4.13)	(−1.90)	(−2.89)	(−4.05)
Obs. (N)	56	56	56	56	56	56	49	49	49	49	49	49	56	56	56	56	56	56
adj. R-sq	0.485	0.582	0.624	0.559	0.746	0.758	0.445	0.623	0.648	0.448	0.675	0.717	0.463	0.598	0.646	0.449	0.634	0.669

(continued on next page)

Table A.6 (*continued*)

	Sample adjustments and control for EU membership																	
	Dependent variable: *WCV*																	
	GDP p.c. > 1000 US$			GDP p.c. < 30,000 US$			Transition countries excluded			Sovjet ctrys excluded			EU membership			EU membership		
	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)	(34)	(35)	(35)
ln(GDP p.c.)	−0.098	0.035	8.022*	0.472**	0.644***	2.057	0.228	0.384	5.228***	0.242	0.337	4.439***	0.293	0.393**	3.566**	0.293	0.323	3.885**
	(−0.25)	(0.09)	(1.85)	(2.39)	(3.47)	(1.22)	(0.86)	(1.43)	(3.26)	(1.06)	(1.50)	(2.95)	(1.51)	(2.07)	(2.37)	(1.51)	(1.62)	(2.53)
$(\ln(\text{GDP p.c.}))^2$	−0.001	−0.005	−0.925*	−0.033***	−0.040***	−0.215	−0.019	−0.024	−0.619***	−0.019	−0.021	−0.523***	−0.022*	−0.024**	−0.412**	−0.022*	−0.020*	−0.453**
	(−0.02)	(−0.23)	(−1.88)	(−2.87)	(−3.66)	(−1.01)	(−1.21)	(−1.58)	(−3.10)	(−1.48)	(−1.66)	(−2.80)	(−1.98)	(−2.20)	(−2.19)	(−1.98)	(−1.71)	(−2.37)
$(\ln(\text{GDP p.c.}))^3$			0.035*			0.007			0.024***			0.020**			0.016**			0.017**
			(1.89)			(0.81)			(2.93)			(2.64)			(2.03)			(2.23)
Units	0.033	0.080	0.097**	0.102*	0.187***	0.184***	0.059	0.110*	0.118*	0.062	0.121**	0.134**	0.056	0.113**	0.122**	0.056	0.108**	0.117**
	(0.66)	(1.64)	(2.04)	(1.70)	(3.76)	(3.59)	(0.77)	(1.87)	(1.96)	(0.96)	(2.31)	(2.52)	(1.00)	(2.26)	(2.38)	(1.00)	(2.11)	(2.20)
Area	0.002	0.028*	0.020	0.001	0.032**	0.028*	0.001	0.030	0.015	0.001	0.028*	0.016	0.006	0.033**	0.026*	0.006	0.038**	0.029*
	(0.17)	(1.90)	(1.62)	(0.03)	(2.21)	(1.79)	(0.08)	(1.65)	(1.00)	(0.04)	(1.74)	(1.18)	(0.43)	(2.30)	(1.76)	(0.43)	(2.54)	(1.94)
Area/units	−0.007	−0.004	0.001	0.006	0.014	0.014	−0.002	−0.001	0.003	−0.002	−0.001	0.004	−0.003	−0.001	0.002	−0.003	−0.001	0.003
	(−0.71)	(−0.53)	(0.07)	(0.50)	(1.59)	(1.50)	(−0.15)	(−0.14)	(0.25)	(−0.20)	(−0.01)	(0.47)	(−0.23)	(−0.17)	(0.24)	(−0.23)	(−0.13)	(0.28)
Ethnic		0.165*	0.119		0.176*	0.177*		0.146	0.108		0.152	0.122		0.148	0.144		0.174*	0.163*
		(1.81)	(1.37)		(1.77)	(1.78)		(1.41)	(1.14)		(1.50)	(1.34)		(1.42)	(1.49)		(1.76)	(1.80)
Trade/GDP		0.002***	0.002***		0.003***	0.003***		0.003***	0.002***		0.003***	0.002***		0.003***	0.003***		0.003***	0.003***
		(3.80)	(4.10)		(6.18)	(5.52)		(3.65)	(3.46)		(4.34)	(4.41)		(4.81)	(4.42)		(4.91)	(4.55)
Urbanization		−0.003*	−0.002*		−0.003**	−0.003**		−0.003	−0.003		−0.003*	−0.003*		−0.003**	−0.003**		−0.003*	−0.003**
		(−1.79)	(−1.78)		(−2.23)	(−2.30)		(−1.41)	(−1.48)		(−1.70)	(−1.91)		(−2.05)	(−2.16)		(−1.81)	(−2.06)
Federal dummy		−0.059	−0.039		−0.050	−0.045		−0.088*	−0.051		−0.088*	−0.060		−0.089**	−0.071**		−0.092**	−0.070**
		(−1.43)	(−1.04)		(−1.36)	(−1.22)		(−1.93)	(−1.22)		(−2.01)	(−1.50)		(−2.45)	(−2.09)		(−2.40)	(−2.03)
EU27													−0.010	−0.057	−0.036			
													(−0.21)	(−1.17)	(−0.78)			
EU15																−0.002	−0.017	0.003
																(−0.05)	(−0.36)	(0.06)
Constant	1.194	−0.132	−22.970*	−1.486*	−3.210***	−6.880	−0.330	−1.801	−14.490***	−0.403	−1.645	−12.460***	−0.685	−1.947**	−10.370**	−0.685	−1.724*	−11.190***
	(0.65)	(−0.07)	(−1.81)	(−1.70)	(−3.90)	(−1.59)	(−0.29)	(−1.50)	(−3.47)	(−0.40)	(−1.59)	(−3.15)	(−0.79)	(−2.29)	(−2.68)	(−0.79)	(−1.92)	(−2.83)
Obs. (N)	49	49	49	50	50	50	40	40	40	47	47	47	56	56	56	56	56	56
adj. R-sq	0.492	0.648	0.674	0.480	0.722	0.718	0.524	0.661	0.715	0.518	0.669	0.708	0.486	0.667	0.684	0.486	0.657	0.680

Note: *t*-values are reported in parentheses; standard errors are calculated using White correction; ***, **, and * indicate significance at 1%, 5%, and 10%, respectively.

References

Ades, A.F., Glaeser, E.L., 1995. Trade and circuses: explaining urban giants. Q. J. Econ. 110 (1), 195–227.
Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., Wacziarg, R., 2003. Fractionalization. J. Econ. Growth 8 (2), 155–194.
Amos, O.J., 1988. Unbalanced regional growth and regional income inequality in the latter stages of development. Reg. Sci. Urban Econ. 18 (4), 549–566.
Baltagi, B.H., 2005. Econometric Analysis of Panel Data. Wiley, Chichester.
Baltagi, B.H., Li, D., 2002. Series estimation of partially linear panel data models with fixed effects. Ann. Econ. Finance 3, 103–116.
Barrios, S., Strobl, E., 2009. The dynamics of regional inequalities. Reg. Sci. Urban Econ. 39 (5), 575–591.
Barro, R.J., 2000. Inequality and growth in a panel of countries. J. Econ. Growth 5 (1), 5–32.
Barro, R.J., Sala-i-Martin, X., 1992. Convergence. J. Polit. Econ. 100 (2), 223–251.
Bendel, R.B., Higgins, S.S., Pyke, D.A., 1989. Comparison of skewness coefficient, coefficient of variation, and Gini coefficient as inequality measures within populations. Oecologia 78 (3), 394–400.
Boadway, R., Flatters, F., 1982. Efficiency and equalization payments in a federal system of government: a synthesis and extension of recent results. Can. J. Econ. 15 (4), 613–633.
Brambor, T., Clark, W.R., Golder, M., 2006. Understanding interaction models: improving empirical analyses. Polit. Anal. 14 (1), 63–82.
Buhaug, H., Gleditschm, K., Holtermann, H., Østby, G., Tollefsen, A.F., 2012. It's the local economy, stupid! Geographic wealth dispersion and conflict outbreak location. J. Confl. Resolut. 55 (5), 814–840.
Chen, J., Fleisher, B.M., 1996. Regional income inequality and economic growth in China. J. Comp. Econ. 22 (2), 141–164.
Dalton, H., 1920. The measurement of the inequality of incomes. Econ. J. 30 (119), 348–361.
Deiwiks, C., Cederman, L., Gleditsch, K., 2012. Inequality and conflict in federations. J. Peace Res. 49 (2), 289–304.
Desbordes, R., Verardi, V., 2012. Refitting the Kuznets curve. Econ. Lett. 116 (2), 258–261.
DiNardo, J., Tobias, J.L., 2001. Nonparametric density and regression estimation. J. Econ. Perspect. 15 (4), 11–28.
Durlauf, S.N., 2001. Manifesto for a growth econometrics. J. Econ. 100 (1), 65–69.
Easterly, W., 2001. The middle class consensus and economic development. J. Econ. Growth 6, 317–335.
Elbers, C., Lanjouw, P., Mistiaen, J., Özler, B., Simler, K.R., 2005. Are neighbours equal? Estimating local inequality in three developing countries. In: Kanbur, R., Venables, A. (Eds.), Spatial Inequality and Development. Oxford University Press, New York, pp. 37–76.
Engel, C., Rogers, J.H., 1996. How wide is the border? Am. Econ. Rev. 86 (5), 1112–1125.
Ezcurra, R., Rapun, M., 2006. Regional disparities and national development revisited: the case of Western Europe. Eur. Urban Reg. Stud. 13 (4), 355–369.
Fan, C.C., Casetti, E., 1994. The spatial and temporal dynamics of U.S. regional income inequality, 1950–1989. Ann. Reg. Sci. 28 (2), 177–196.
Gennaioli, N., Porta, R.L., de Silanes, F.L., Shleifer, A., 2013a. Human capital and regional development. Q. J. Econ. 128 (1), 105–164.
Gennaioli, N., Porta, R.L., de Silanes, F.L., Shleifer, A., 2013b. Growth in regions. NBER Working Papers 18937. National Bureau of Economic Research, Inc.
Hale, T., 2003. The Theoretical Basics of Popular Inequality Measures; Online Computation of Examples. University of Texas Inequality Project.
Hirschman, A.O., Rothschild, M., 1973. The changing tolerance for income inequality in the course of economic development. Q. J. Econ. 87 (4), 544–566.
IMF, 2000. Transition economies: an IMF perspective on progress and prospects. Discussion Paper 00/08. International Monetary Fund.
Jian, T., Sachs, J.D., Warner, A.M., 1996. Trends in regional inequality in China. China Econ. Rev. 7 (1), 1–21.
Kanbur, R., Venables, A., 2005a. Rising spatial disparities and development. Number 3 in United Nations University Policy Brief. Helsinki.
Kanbur, R., Venables, A., 2005b. Spatial inequality and development. In: Kanbur, R., Venables, A. (Eds.), Spatial Inequality and Development. Oxford University Press, New York, pp. 3–11.
Kanbur, R., Zhang, X., 2005. Fifty years of regional inequality in China: a journey through central planning, reform, and openness. Rev. Dev. Econ. 9 (1), 87–106.
Kanbur, R., Venables, A.J., Wan, G., 2005. Introduction to the special issue: spatial inequality and development in Asia. Rev. Dev. Econ. 9 (1), 1–4.
Kessler, A.S., Hansen, N.A., Lessmann, C., 2011. Interregional redistribution and mobility in federations: a positive approach. Rev. Econ. Stud. 78 (4), 1345–1378.
Krugman, P., Elizondo, R.L., 1996. Trade policy and the Third World metropolis. J. Dev. Econ. 49 (1), 137–150.
Kuznets, S., 1955. Economic growth and income inequality. Am. Econ. Rev. 45 (1), 1–28.
Lessmann, C., 2009. Fiscal decentralization and regional disparity: evidence from cross-section and panel data. Environ. Plan. A 41 (10), 2455–2473.
Lessmann, C., 2012. Regional inequality and decentralization: an empirical analysis. Environ. Plan. A 44 (6), 1363–1388.
Lessmann, C., 2013a. Foreign direct investment and regional inequality: a panel data analysis. China Econ. Rev. 24 (C), 129–149.
Lessmann, C., 2013b. Regional inequality and internal conflict. CESifo Working Paper Series 4112. CESifo Group, Munich.
Libois, F., Verardi, V., 2012. Semiparametric fixed-effects estimator. Working Paper. University of Namur.
List, J.A., Gallet, C.A., 1999. The Kuznets curve: what happens after the inverted-U? Rev. Dev. Econ. 3 (2), 200–206.
Lucas, R.E., 2000. Some macroeconomics for the 21st century. J. Econ. Perspect. 14 (1), 159–168.
Marshall, M.G., Jaggers, K., 2009. Polity IV Project. Political regime characteristics and transitions (1800 to 2007). unpublished manual, Center for Systemic Peace.
Mehran, F., 1976. Linear measures of income inequality. Econometrica 44 (4), 805–809.
Milanovic, B., 2013. All the Ginis. Data set, World Bank, Research Department. Available from http://go.worldbank.org/9VCQW66LA0.
Mills, E.S., Ferranti, D.M.d., 1971. Market choices and optimum city size. Am. Econ. Rev. 61 (2), 340–345.
Nordhaus, W., Azam, Q., Corderi, D., Hood, K., Makarova, N., Mohammed, M., Miltner, A., Weiss, J., 2006. The G-Econ database on gridded output: methods and data. Discussion paper. Yale University.
Pigou, A.C., 1912. Wealth and Welfare. Macmillan, London.
Ram, R., 1991. Kuznets's inverted-U hypothesis: evidence from a highly developed country. South. Econ. J. 57 (4), 1112–1123.
Ravallion, M., 2001. Growth, inequality and poverty: looking beyond averages. World Dev. 29 (11), 1803–1815.
Robinson, P.M., 1988. Root-N-consistent semiparametric regression. Econometrica 56 (4), 931–954.
Rodríguez-Pose, A., 2012. Trade and regional inequality. Econ. Geogr. 88 (2), 109–136.
Rodríguez-Pose, A., Ezcurra, R., 2010. Does decentralization matter for regional disparities? A cross-country analysis. J. Econ. Geogr. 10 (5), 619–644.
Rodriguez-Pose, A., Gill, N., 2006. How does trade affect regional disparities? World Dev. 34 (7), 1201–1222.
Sala-i-Martin, X., 1996. Regional cohesion: evidence and theories of regional growth and convergence. Eur. Econ. Rev. 40 (6), 1325–1352.
Sen, A., 1973. On Economic Inequality. Norton, New York.
Stewart, F., 2000. Crisis prevention: tackling horizontal inequalities. Oxf. Dev. Stud. 28 (3), 245–262.
Stewart, F., 2002. Horizontal inequalities: a neglected dimension of development. QEH Working Paper Series 81. University of Oxford.
Terrasi, M., 1999. Convergence and divergence across Italian regions. Ann. Reg. Sci. 33 (4), 491–510.
Treisman, D., 2008. Decentralization dataset. http://www.sscnet.ucla.edu/polisci/faculty/treisman/.
Venables, A.J., 2005. Spatial disparities in developing countries: cities, regions, and international trade. J. Econ. Geogr. 5 (1), 3–21.
Verardi, V., Debarsy, N., 2012. Robinson's square root of N consistent semiparametric regression estimator in Stata. Stata J. 12 (4), 726–735.
Wei, K., Yao, S., Liu, A., 2009. Foreign direct investment and regional inequality in China. Rev. Dev. Econ. 13 (4), 778–791.
World Bank, 2011. World Development Indicators 2010. World Bank, Washington D. C.
Williamson, J.G., 1965. Regional inequality and the process of national development: a description of patterns. Econ. Dev. Cult. Chang. 13 (4), 3–45.
Wooldridge, J.M., 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, Cambridge, MA.
Yemtsov, R., 2005. Quo Vadis? Inequality and poverty dynamics across Russian regions. In: Kanbur, R., Venables, A. (Eds.), Spatial Inequality and Development. Oxford University Press, New York, pp. 348–408.
Zeng, D.-Z., Zhao, L., 2010. Globalization, interregional and international inequalities. J. Urban Econ. 67 (3), 352–361.