Spatial Analysis of Regional Income Inequality

SERGIO J. REY
Department of Geography
San Diego State University
and
Regional Economics Applications Laboratory
University of Illinois

October 15, 2001

Abstract

Questions surrounding regional economic convergence have commanded a great deal of recent attention in economics literature. As in other recent cases in the social sciences, the application of spatially explicit methods of data analysis to the convergence question has yielded important insights on regional economic growth.

By contrast, the literature on regional income inequality, although somewhat older than the convergence literature, has been slower to adopt new spatially explicit methods of data analysis. This chapter helps to speed that adoption by investigating the role of spatial dependence and spatial scale in the analysis of regional income inequality in the US over the 1929-2000 period. The findings reveal a strong positive relationship between measures of inequality in state incomes and the degree of spatial autocorrelation. Additionally, a geographically based decomposition of inequality highlights a strong positive relationship between the interregional inequality share (as opposed to intraregional inequality) and spatial clustering. Finally, a new approach to inference in regional inequality analysis is suggested and demonstrated as providing a formal explanatory framework to complement the broad, but descriptive approaches in the existing literature.

1 Introduction

Just over a decade ago, Barro and Sala-i-Martin (1991) reintroduced mainstream macroeconomics to the concept of a region. That introduction set off an explosion of research on the question of regional economic convergence. [1] Much of this research represented a shift in focus from studying the dynamics of international income disparities to the analysis of intranational dynamics. That is, whether incomes between regions within a given nation state become more, or less, similar over time.

Despite the rich geographical dimensions underlying the data used in regional income convergence analysis, the role of spatial effects has only recently begun to attract

[1] For a recent survey of empirical work on convergence see Durlauf and Quah (1999).

attention (Armstrong and Vickerman, 1995; Chatterji and Dewhurst, 1996; Cuadrado-Roura et al., 1999; Fingleton, 1999). These studies demonstrate how the analysis of spatial dependence and spatial heterogeneity can add to a richer understanding of regional economic growth processes (Goodchild et al., 2000).

The study of regional inequality offers interesting contrasts to, as well as similarities with, the literature on regional convergence. Regional income inequality analysis has its origins in the study of personal income inequality. The latter is "a scalar numerical representation of the interpersonal differences in income within a given population" (Cowell, 1995, pg 12). Kuznets (1955) hypothesized an inverted-U relationship between the level of development and personal income inequality. In early stages of development, the concentration of income generating wealth in the hands of a subset of individuals in the population was seen as a required condition for the accumulation of capital that fueled the expansion of industrial activity. In subsequent stages of development, benefits of growth were passed on to other members of society as higher wages and increased income. Personal income inequality would then begin to slow and eventually decline.

Williamson (1965) applied the inverted-U pattern to the question of unequal regional development. Here the focus is on the distribution of regional incomes, and not the incomes of individuals. Initial concentrations of income in certain geographic regions were attributed to unequal natural resource endowments. Williamson argued that these concentrations attracted selective skilled labor migration from the peripheral regions and generated rapid income growth in the core regions. This led to widening differentials in per-capita incomes between the core and peripheral regions. Over time however, a diffusion of income generating factors leads to the subsequent slowing and eventual decline in regional income inequality.

There has been a great deal of work investigating the inverted-U pattern of regional income inequality in national systems.[2] Most of this work relies on descriptive analysis using measures of dispersion in regional income distributions and relates these to a measure of development. This stands in marked contrast to work on regional economic convergence where there has been a tight linkage between model specification and one or more growth theories. As a result, the empirical analysis of regional inequality relies more heavily on exploratory and descriptive methods in contrast to the more confirmatory and inferential approaches in regional convergence studies.

An important similarity that the inequality literature shares with the convergence literature, is a general neglect of the spatial dimensions of the data underlying the empirical analysis. More specifically, a number of issues associated with spatial effects in regional income inequality analysis remain overlooked by previous studies. These surround the relationship between regional inequality and spatial dependence and the sensitivity of inferences on regional inequality to the choice of spatial scale. Moreover, applications of inequality analysis at the regional scale are currently lacking an inferential basis.

These neglected issues are important for both analytical and substantive reasons. Analytically, the extension of regional convergence analysis to fully consider spatial

[2] Alonso (1980); Amos (1983); Maxwell and Peter (1988); Tsui (1993); Fan and Casetti (1994); Kanbur and Zhang (1999); Nissan and Carter (1999); Azzoni (2001); Zhang and Kanbur (2001).

effects has provided important insights to regional growth processes. It remains to be seen if a similar result will hold for regional inequality analysis.

From a substantive perspective spatial inequalities in income have been identified as a destabilizing force in societies throughout human history.[3] Space can also matter a great deal to policies targeted at reducing regional income disparities. As the work of Baker and Grosh (1994) has shown, not only is the question of the delineation of regional boundaries important, but the level of geographic unit chosen can have an influence on targeting outcomes. At the same time, the spatial distribution of regional incomes as well as the degree of inequality need to be considered jointly. This is because social tensions arising from income inequalities may be heightened by geographical concentration of poorer social groups.

As Zhang and Kanbur (2001) have argued, it is important to move beyond single scalar measures of inequality to consider more disaggregate movements within the income distribution. For example, it is possible that incomes become polarized across groups in a society. If incomes within each of these groups become similar, yet the differences among the mean incomes from each group increase, then an overall index of inequality can in fact decline while polarization increases. Although Zhang and Kanbur (2001) focus on personal income distributions, their concerns can be extended to the case of regional income distributions. A focus on the evolution of the overall measure of regional inequality could mask very important developments within the distribution. Some of these developments could have spatially explicit manifestations reflecting poverty traps, convergence clubs and other forms of geographical clustering that are not captured by an overall regional inequality measure.

This chapter aims to contribute to the regional inequality literature by investigating several spatial dimensions that have been largely ignored. It focuses on three extensions of regional inequality analysis:

- an exploration of the relationship between regional inequality and spatial dependence,
- the analysis of the role of spatial scale and its impact on inequality measurement,
- alternative inferential strategies for regional inequality analysis.

The plan of this chapter is as follows. Section 2 provides an overview of recent work on regional inequality measurement. This is followed by a discussion of several issues related to the spatial characteristics of the data that have been largely overlooked in existing studies. Section 4 presents an empirical investigation of some of these issues using data for the United States over the 1929-2000 period. The chapter concludes with a summary of the key findings.

[3] "Large income and wealth differences between countries and regions generated acts of aggression which inflicted considerable human suffering, loss of resources and knowledge, destruction of civilizations and environmental damage." (Levy and Chowdhury, 1995, pg 17).

2 Regional Inequality Analysis

2.1 Measurement

A wide number of inequality measures are available in the literature.[4] In regional inequality analysis, a popular choice has been Theil's inequality measure (Theil, 1967), given as:

$$T = \sum_{i=1}^{n} s_i \log(ns_i) \tag{1}$$

where n is the number of regions, y_i is per capita income in region i, and:

$$s_i = y_i / \sum_{i=1}^{n} y_i. \tag{2}$$

T is bounded on the interval $[0, log(n)]$, with 0 reflecting perfect equality (i.e., $y_i = y_j \forall i, j$), and a value of $log(n)$ occurring when all the income is concentrated in one region. T measures systematic or, as what we shall refer to now as, global inequality incomes across the regional observations at one point in time.[5]

As T is a member of a generalized entropy class of inequality measures it has the quality of being additively decomposable (Shorrocks, 1984). This is desirable for both analytic and arithmetic reasons. Substantively, the ability to measure the contribution to global inequality (1) that is attributable to inequality between and within different partitions of the observational units can provide a deeper understanding of global inequality. For example, in studies of wage inequality, the partitions are sometimes defined according to industry groupings, such as manufacturing versus agriculture. Mathematically, the decomposition is exhaustive, meaning that the global inequality is completely separated into the two components.

In studies of regional income inequality, the decompositional property has been exploited to investigate the extent to which global inequality is attributable to inequality "between" or "within" regional groupings.[6] By partitioning the n spatial observations into ω mutually exclusive and exhaustive groups, T can be decomposed as follows:

$$T = \sum_{g=1}^{\omega} s_g \log(n/n_g s_g) + \sum_{g=1}^{\omega} s_g \sum_{i \in g} s_{i,g} \log(n_g s_{i,g}) \tag{3}$$

where n_g is the number of observations in group g (and $\sum_g n_g = n$), $s_g = \sum_{i \in g} y_{i,g} / \sum_i^n y_i$ is the share of total income accounted for by group g, and $s_{i,g} = y_{i,g} / \sum_{i=1}^{n_g} y_{i,g}$ is region i's share of group g's income.

The first term on the right hand side of (3) is the "between-group" component of inequality, while the second term is the "within-group" component of inequality. In other words:

$$T = T_B + T_W. \tag{4}$$

[4] See Cowell (1995) for a recent overview.

[5] In what follows, time subscripts are omitted unless explicitly noted.

[6] The decomposition has also been used to study the contribution of different components of income, such as transfer payments versus wages, and how regional inequalities in these components contribute to overall regional inequality. See for example Eff (1999).

In a spatial context, the within-group term measures intraregional inequality, while the between-group component captures interregional inequality. Put another way, the interregional term measures the distance between the mean incomes of the aggregate groups. The intraregional term measures distances between the incomes of regions belonging to the same group.

2.2 Existing Studies

Fan and Casetti (1994) analyzed U.S. state income inequality using four Census Divisions to define the partitions. The within region component was found to account for the largest share of inequality in the US over the 1950-89 period. Using the same partitioning, but county rather than state data, Conceição and Ferreira (2000) also conclude that the within component of inequality was the most important share over the 1969-96 period. Nissan and Carter (1999) analyzed state income inequality over the 1969-95 period. A regional inequality decomposition was employed for the states as a whole, as well as for the subset of metropolitan and the subset of rural states. Inequality between regions was found to decline in the early 1970's, but increased through the 1980's, followed again by convergence in the 1990's. At the same time, they found strong evidence that within region inequality showed a much stronger decline over the study period. This was true for all states, as well as metropolitan and rural states.

Inequality decomposition has been applied in several regional analyses outside of the US as well. Fujita and Hu (2001) analyzed regional income disparities in China over the 1984-94 period, using a coastal-interior partitioning of 30 provinces. They find that overall regional inequality was fairly stable, exhibiting a slight decline in the 1980's. The overall decline was driven by the decline in intraregional inequality; the latter being larger than interregional inequality until the last three years of their sample, accounting for between 77% and 43% of overall inequality.

Azzoni (2001) explored inequality in 20 Brazilian states over the 1939-95 period. Overall regional inequality was substantial up until 1965, at which point a steady decrease began. Partitioning the states into 5 groups, revealed that interregional inequality was the most important contributor to overall regional disparities. Moreover, the interregional component accounted for an increasing share of total inequality, starting from 60% and ending at 87%.

Geographical decomposition of inequality has been applied at the international as well as intranational scale. Theil (1996) applied a decompositional analysis to 100 countries over 1950-90 and found that the majority (roughly 88%) of global inequality was due to differences between, rather than within, regional groupings of countries. In a similar study Levy and Chowdhury (1995) report that the relative importance of the two components has varied over the 1960-90 period, with the between region component dominating from 1960-67, the within region component being larger from 1967-83, and a second reversal from 1983-90.

These studies have illuminated the spatial structure underlying the dynamics of regional inequality in different contexts. However, there is much variation across the studies with respect to the relative importance of the inter versus intraregional inequality components. What is currently unknown, however, is to what extent that variation is due to differences in the structure of the economies in the different studies or to the articulation of methodological issues across the studies. These issues include the choice of regional partitioning and the spatial scale of the observational units.

At the same time, there are several limitations in these studies having to do with a lack of an inferential basis that require additional attention. Moreover, it is possible that much more can be said about the geographical dimensions of regional inequality. In the remainder of the paper, these issues are more fully discussed and an initial attempt at addressing these concerns is presented.

3 Spatial Effects in Regional Inequality Analysis

3.1 Spatial Dependence

Spatial dependence occurs when the values for some phenomenon measured at one location are associated with the values measured at other locations (Anselin, 1988). The issues that spatial dependence raise for econometric analysis of regional income convergence have received recent attention (Fingleton, 2001; Rey and Montouri, 1999), yet the role of spatial dependence in studies of regional inequality has been largely ignored.

The issues associated with spatial dependence may be conveniently split into two groups. From a substantive perspective, spatial dependence can play an important role in shaping the geographical distribution of incomes. From a nuisance perspective, spatial dependence can complicate the application of traditional statistical methods designed to analyze regional inequality.

Lucas (1993) suggests a model that allows for cross-economy interactions in the form of human capital spillovers. The presence of these spillovers (i.e., learning by doing) can radically alter the patterns of cross-economy growth from those suggested by a traditional neoclassical growth model. The basic idea is that if economies interact via human capital spillovers, and if the interacting economies become grouped, it is likely that within group spillovers will be stronger than between group spillovers. This would result in within group convergence but, potentially, divergence between groups.

From a nuisance perspective, the presence of spatial dependence presents a challenge to the use of statistical inference in inequality analysis. This is because the existing approaches to inference are based on an assumption of random sampling which is violated by the presence of spatial dependence. This issue is taken up further below.

Spatial dependence of a nuisance form can also arise from a mismatch between the regional boundaries used to organize the data and the boundaries of the actual socioeconomic process under study. In regional inequality analysis, this could be reflected in a misspecified partitioning, whereby the partitioning imposed by the researcher fails to match the natural groupings of the regional observations.

Interestingly, global inequality measures are insensitive to the underlying spatial distribution of the income values. This reflects a focus on the dispersion of the distribution only. This also brings up an intriguing question regarding the relationship between the level of spatial dependence in regional incomes and spatial income inequality. At first glance, it would appear that strong positive spatial autocorrelation would lead to increasing global inequality, given that we would be able to see clusters of similar incomes on a map. However, the analysis of spatial autocorrelation rests on the assumption of spatial stationarity. Loosely speaking, this requires the mean and variance of the distribution to be constant over space. At the same time, it is well known, that the presence of spatial dependence can induce a form of heteroskedasticity in the error terms of spatial econometric models (Anselin, 1988). The question then becomes one of being able to disentangle any apparent spatial heterogeneity induced by the dependence, from true heterogeneity reflecting a lack of spatial stationarity. This would allow one to distinguish between increasing inequality owing to increasing variance in a distribution versus inequality attributable to the mixing of different distributions.

3.2 Spatial Scale in Regional Convergence Analysis

As is well known in other areas of spatial analysis, the modifiable areal unit problem (MAUP) arises when the inferences drawn about the process under study are sensitive to the spatial scale and partitioning of the data at hand (Opensaw and Alvanides, 1999). Given the wide scope for selecting a spatial partitioning as well as the unit of observation in regional inequality analysis, it would appear that the MAUP would attract much attention. This, however, has not been the case.

That inequality measures will be subject to the MAUP can be seen by an examination of the bounds for the the global T. The theoretical upper bound of T is $log(n)$. Any change in spatial scale, say from the state level to the county level, will change the number of observational units and affect the upper bound of the statistic. The question of how this affects the comparison of inferences drawn about inequality at the two different scales has gone unexamined in the literature.

In addition to affecting the upper bound of the global T, a change in spatial scale may also impact the decomposition of the global measure into its intra and interregional components. Here again, this issue has been neglected in previous regional studies. These issues are taken up in the empirical analysis in section 4.

3.3 Inferential Issues

In regional applications of inequality measures the focus is typically on a descriptive analysis, either reporting the value of the measure at one point in time, tracking the statistic over time, or decomposing global regional inequality into its intra and interregional shares. An important omission in the regional literature is the use of inferential methods that allow for formal hypotheses testing regarding inequality measures. Several interesting hypotheses regarding regional inequality that could be examined from an inferential perspective:

- Is the coefficient different from what would be expected under perfect regional equality
- Is any empirical change in an inequality measure over time significantly different from zero?
- Is the share of intraregional inequality significantly different from some hypothesized value?

- In comparing two (or more) economic systems over time (i.e., the U.S. vs. the EU) is the difference in regional inequality significant?
- Are the within and between regional inequality components significantly different between the two systems?

In the wider inequality literature, two general approaches towards inference have been used. The first rests on theoretical results regarding the asymptotic distributions of different inequality statistics (Maasoumi, 1997). Application of these results to the small samples used in most regional settings is problematic for two main reasons. First, the inequality statistics are typically truncated at 0. Use of asymptotic standard errors to construct confidence intervals around the empirical value for the statistic may produce inadmissible interval bounds. The second problem is that the small sample properties of these statistics are unknown, and as such the usefulness of asymptotic results is also unknown.

The second approach to inference in inequality analysis rests on computational procedures. Mills and Zandvakili (1997) suggest the use of a bootstrap to construct empirical sampling distributions for inequality measures. At first glance, this may appear to offer a way to introduce an inferential component into regional inequality analysis. Unfortunately, regional income data often displays a high degree of spatial autocorrelation (Rey and Montouri, 1999). The presence of such dependence violates the random sampling assumption at the heart of the bootstrap methodology. A similar difficulty applies to the asymptotic results.[7]

Because the presence of spatial dependence rules out the use of asymptotics or bootstrapping, an alternative approach to inference is required. The approach suggested here is based on random spatial permutations of the actual incomes for a given map pattern. This can be used to test hypotheses regarding the decomposition of global inequality into its interregional and intraregional components. This is accomplished with the following steps:

1. Calculate decomposition:
$$T^* = T^*_W + T^*_B \tag{5}$$
2. Randomly reassign incomes to new locations
3. Calculate decomposition for permutated map:
$$T^P = T^P_W + T^P_B \tag{6}$$
4. Repeat steps 2 and 3, K times.

The values for the global inequality measure T^P will be the same for any permutation in a given time period. Because the observations are being randomly reassigned

[7] In testing for changes in personal income inequality indices overtime the two temporal samples may be dependent, since the same individuals may be included in both periods. The focus here is on testing a single regional distribution at one point in time, so this issue is not addressed. For further details see Zheng and Cushing (2001).

to different regional groupings in each permutation, however, the values for the intraregional (T_W^P) and interregional (T_B^P) are likely to vary across the permutations. The actual inequality measure T_W^* can then be compared against the value it would have been expected to take on if regional incomes were randomly distributed in space. The latter would be obtained as the average of the empirically generated measures from step 3:

$$\bar{T}_W = \frac{1}{K} \sum_{P=1}^{K} T_W^P \tag{7}$$

Differences between the actual statistic and its expected value could be compared against the empirical sampling distribution in one of two ways. The first would be based on the assumption that the empirical sampling distribution is approximately normal, in which case the standard deviation for that distribution, given as:

$$s_{T_W} = \frac{1}{K} \sum_{P=1}^{K} \left(T_W^P - \bar{T}_W\right)^2 \tag{8}$$

could be used to define a confidence interval.

The second approach to inference using the random spatial permutations is to use a percentile approach. This simply sorts the empirically generated T_W^P values and then develops a pseudo significance level by calculating the share of the empirical values that are more extreme than the actual value:

$$p(T_W) = \frac{1}{K} \sum_{P=1}^{K} \psi_P \tag{9}$$

where $\psi_P = 1$ if T_W^P is more extreme than T_W, $\psi_P = 0$ otherwise. The advantage of this approach over the first is that the problem of inadmissible interval bounds is avoided.

Because the global inequality measure is invariant to the spatial arrangement of regional incomes, the random permutation approach cannot be used to test inferences regarding the global measure. Future work will focus on developing methods of inference for the global measure in the presence of spatial dependence.

4 Empirical Illustration

To explore some of these issues we focus on US per capita income over the 1929-2000 period for the 48 lower states.[8] Attention is first directed towards the relationship between regional income inequality and spatial dependence. This is followed by an analysis of how changes in spatial scale may affect the measures of regional inequality. Finally, approaches to statistical inference in regional inequality analysis are examined.[9]

[8]The state and county income data used in this study were obtained from the May 3, 2001 release of the BEA state and local personal income series.

[9]The empirical analysis was carried out using the package STARS (Rey, 2001).

4.1 Inequality and Spatial Dependence

Figure 1 portrays the relationship between regional inequality and spatial autocorrelation. Inequality is measured using the global T from (1). Clearly, the long term trend has been one of declining regional income inequality in the U.S., with the majority of the decline coming during the war years in the early 1940's. A slight turn around towards increasing regional inequality is seen through the 1980's, for which numerous explanations have been put forth (Amos, 1988; Fan and Casetti, 1994). However, these explanations focus on US specific causes and ignore presence of a similar turn around in other national systems occurring during the same period. (Paci and Pigliaru, 1997).

Spatial autocorrelation is measured using Moran's I, defined as:

$$I = \frac{n}{\sum_i \sum_j w_{i,j}} \frac{\sum_i \sum_j (y_i - \overline{y})(y_j - \overline{y})}{\sum_i (y_i - \overline{y})^2} \tag{10}$$

where $w_{i,j}$ is an element of a binary spatial contiguity matrix with elements taking on the value of 1 if states i and j are first order neighbors (i.e., share a border), 0 otherwise. y_i is per capita income in state i and $\overline{y}$ is the average per capita income for the 48 states.

Moran's I has an expected value and variance that are function of the structure of the spatial weights matrix only, and are not influenced by the value of the variate in question. As such the moments of I are constant each year: $E[I] = -0.213, V[I] = 0.009$. [10] Basing inference regarding I on a normality assumption results in the statistic being significant in each year in this sample. Thus, personal incomes are highly autocorrelated across the states.

Figure 1 also reveals a strong positive relationship between the inequality measure and the autocorrelation index. The sample correlation between these two statistics over the 72 years is 0.798. It should be noted, however, that a simple re-shuffling of the actual income values about the map for a given year would leave the measure of inequality unchanged, while Moran's I would vary. This highlights the difference in emphasis between the two statistics and suggests that their joint application to the analysis of regional income growth might produce important complementarities offering insights not obtainable when either is used in isolation.

Figure 2 displays the global T and its decomposition into the interregional and intraregional components. The partitioning of the 48 states is based on the US Census Regions which are defined in Table 1. This is the same partitioning as used in Fan and Casetti (1994), although our sample includes a larger number of years. In each of the 72 years, the intraregional component exceeds that of the interregional share. These results are in agreement with those reported by Fan and Casetti (1994).

Figure 3 shows the effect of changing the partitioning scheme from the four Census regions to the nine Census Divisions, as defined in Table 2. The relative importance of the two components of inequality is reversed, with the interregional component now dominating. This reflects an increase in the internal homogeneity of the regions, largely due to the decrease in the number of states found in each region.

[10] For the detailed expressions for the moments of I under the normality assumption see Cliff and Ord (1973).

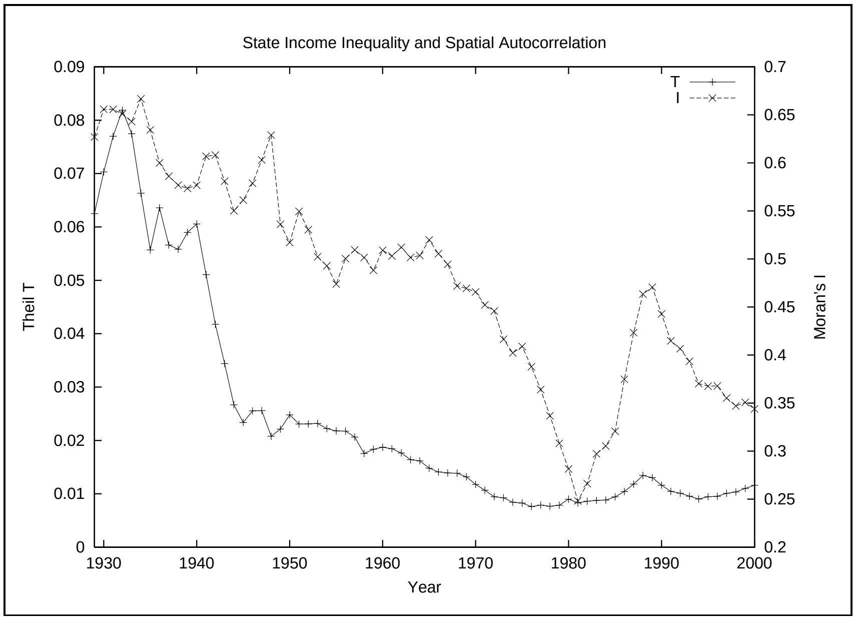

Figure 1: Regional Inequality and Spatial Dependence: US States

Table 1: Census Regions

REGION	STATES
Northeast	Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont
Midwest	Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin
South	Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia
West	Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming

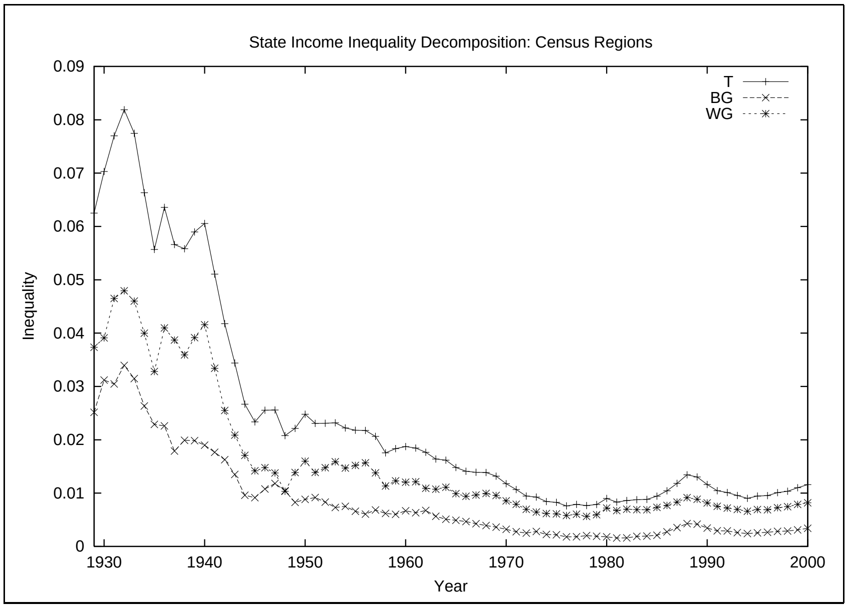

Figure 2: Regional Inequality Decomposition: Census Regions

Figure 4 shows the effect of changing the partitioning scheme to the eight regions defined by the Bureau of Economic Analysis (BEA) and listed in Table 3. The reversal in the relative importance of the two components of inequality is even more pronounced. Although there is a high degree of similarity between the Census Divisions and the BEA Regions, the interregional inequality component is substantially higher in the latter partitioning. Moreover, the share claimed by the interregional component using the BEA Regions in the first half of the study period is higher than that claimed by the intraregional component during the same period when the Census Divisions are used.

These interregional components are isolated in Figure 5 revealing the much higher interregional share each year in the sample when the partitioning is based on the BEA regions. The larger number of groups in the BEA Regions and Census Divisions relative to the Census Regions explains why the former have larger interregional components than the latter. However, the BEA Regions have higher interregional inequality than the Census Divisions, despite having a smaller number of groups of states. Consequently, the interregional share is not a simple function of the number of regional groupings used.

The rankings of the three partitions with respect to the share of total inequality claimed by the interregional component is consistent across the entire 72 year period, with the BEA scheme at the top and the Census Region partition at the bottom. Because an increase in the interregional inequality is due to differences in the mean values becoming more important than intraregional differences, the patterns in figure 5 suggest that the homogeneity of the BEA regions is stronger than that found in the other two

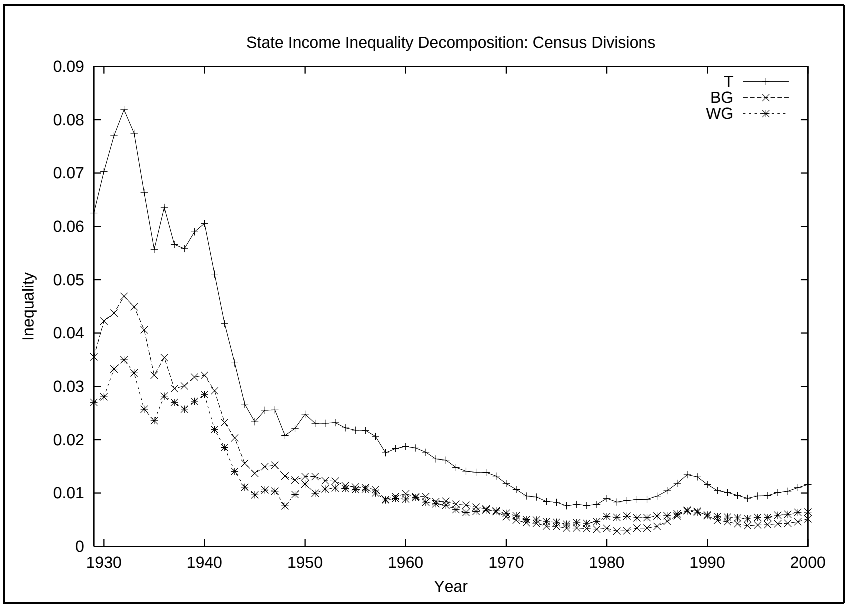

Figure 3: Regional Inequality Decomposition: Census Divisions

Table 2: Census Divisions

DIVISION	STATES
Northeast	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Middle Atlantic	New Jersey, New York, Pennsylvania
East North Central	Illinois, Indiana, Michigan, Ohio, Wisconsin
West North Central	Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota
South Atlantic	Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia
East South Central	Alabama, Kentucky, Mississippi, Tennessee
West South Central	Arkansas, Louisiana, Oklahoma, Texas
Mountain	Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming
Pacific	California, Oregon, Washington

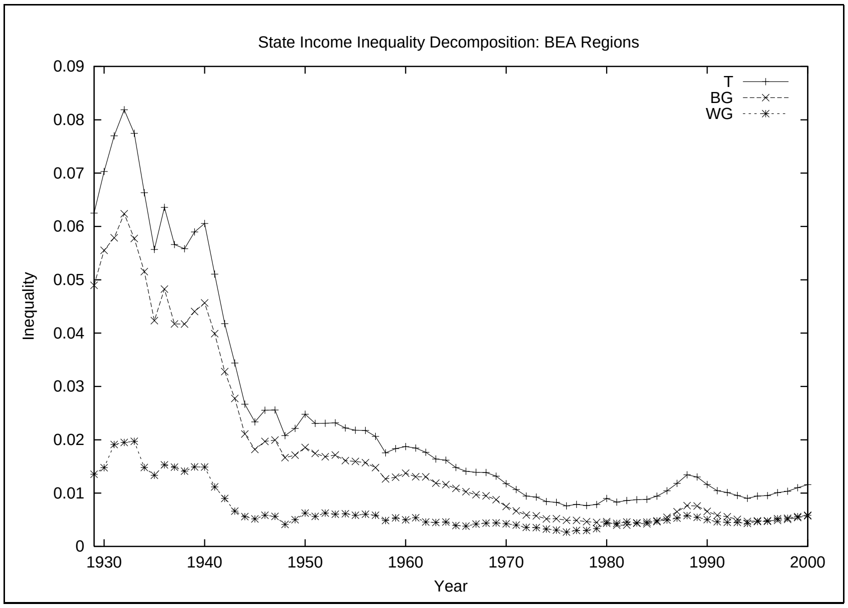

Figure 4: Regional Inequality Decomposition: BEA Regions

sets of regions.

Despite the differences in magnitudes, all three partitions yield interregional inequality shares that are generally declining over time. This general decline was also seen in both the global inequality measure and the level of spatial dependence in Figure 1. The three interregional inequality series also display an increase during the 1980's which is much more pronounced than was the case for the global inequality measure. This coincides with the sharp increase displayed by Moran's I in Figure 1. In fact, each of the interregional inequality shares has a strong positive correlation with the measure of spatial dependence.[11]

4.2 Spatial Scale and Regional Inequality

The results from the previous section indicate that the choice of the partition can fundamentally change the inequality decomposition, both quantitatively and qualitatively. In this section, attention shifts to the effect that a change in the spatial scale of the observational unit may have on the inferences regarding inequality. This is accomplished by using county rather than state data for the 48 lower states.

Moving to the county level of analysis required a truncation of the time series that could be considered. Annual data on per capita income was only available from 1969 through 1999. Despite this shorter time period, the use of counties as the observational

[11] The correlations with Moran's I are: 0.63 (Census Regions), 0.66 (Census Divisions) and 0.64 (BEA Regions).

Table 3: BEA Regions

REGION	STATES
Northeast	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Mideast	Delaware, Maryland, New Jersey, New York, Pennsylvania
Great Lakes	Illinois, Indiana, Michigan, Ohio, Wisconsin
Plains	Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota
Southeast	Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia
Southwest	Arizona, New Mexico, Oklahoma, Texas
Rocky Mountains	Colorado, Idaho, Montana, Utah, Wyoming
Far West	California, Nevada, Oregon, Washington

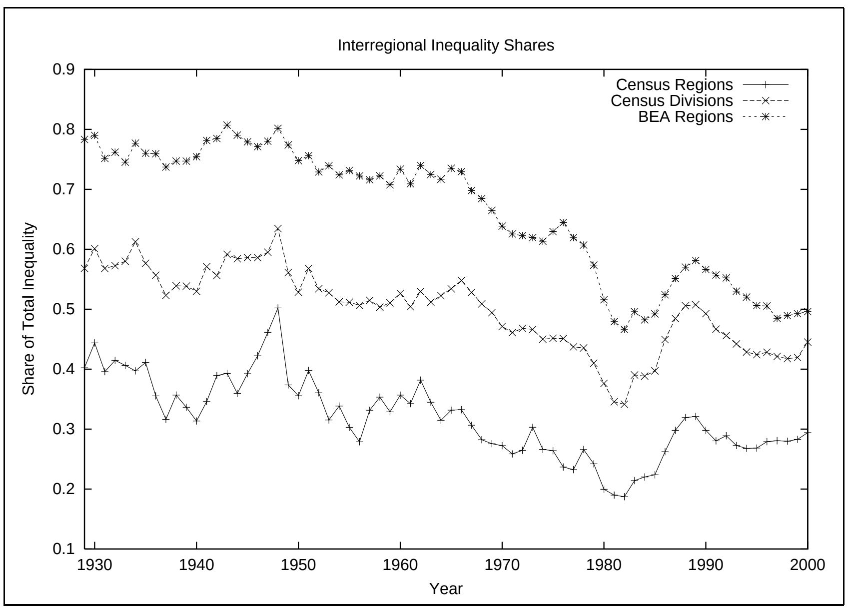

Figure 5: Interregional Inequality Components

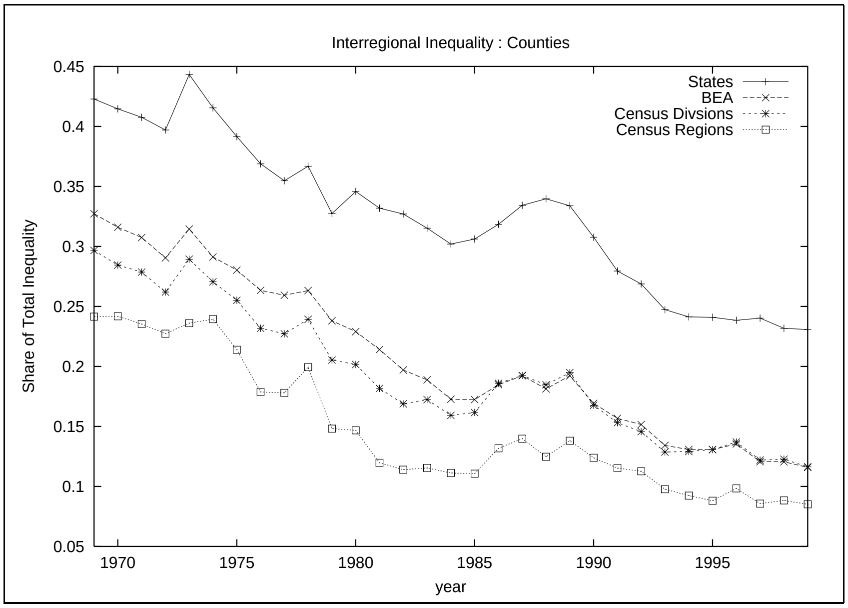

Figure 6: Interregional Inequality: County Unit of Observation

units has two important benefits. First, it vastly increases the number of spatial units, from 48 states to 3079 counties. The second benefit is that a fourth level of partitioning can now be used, as the counties can be assigned to their states, while the states are nested in the three partitions examined in the previous section.

Figure 6 plots the interregional inequality shares for the counties using these four partitions. The most striking pattern is the substantially higher share found for the state partition compared to those for the other three. Intuitively, this reflects the stronger homogeneity of the states. The relative ranking of the three other partition is in general agreement with that found when the states were the unit of analysis, although the pattern is less clear from 1995 onward.

An important difference between the state and county level analysis is that the intraregional inequality component dominates the global decomposition at the county scale. This is true for all of the four partitions. When using the states as the unit of observation, the only partition for which the intraregional component was the largest each year was the Census Regions (see Figure 5). For the BEA regions and Census Divisions, the interregional share was the largest for the majority of the years, again using states. If the focus is on the post-1969 period, the interregional component remains dominant for the states only for the BEA regions partition.

4.3 Inference

The final issue examined is the role of inference in regional inequality analysis. This is an important issue as often interests centers on how much inequality the particular decomposition accounts for. Cowell and Jenkins (1995) suggest a simple measure to get at this question:

$$R(B) = T_B/T \tag{11}$$

where T_B and T are as defined above, and $R(B)$ is the share of inequality accounted for by the between group component. This is similar to the polarization index recently suggested by Zhang and Kanbur (2001):

$$P = T_B/T_W \tag{12}$$

Unfortunately, neither of these studies developed an inferential basis against which the measures could be evaluated, and instead used their measure in a descriptive fashion.

At this point, it could be asked why inference is needed, since the dominance of one component over the other is sometimes readily apparent; for example the interregional component for the states using the Census Region partition is dominated by the intraregional share (See Figure 2). The response is that this question misses an important point. Finding that the interregional component accounts for a smaller share of global inequality should not be taken to mean that the interregional component is irrelevant, or that the partition that it is based upon is somehow erroneous. The question should instead be rephrased as follows: "For a given partition and spatial scale of observation, is the interregional share observed different from what could be expected by random chance?"

Figure 7 provides an answer to this question. It depicts the actual value of the interregional inequality component for the states using the Census Regions as the partition. Also shown are the error bars associated with ± 2 standard deviations around the average values for the shares from 1,000 random spatial permutations of the incomes for each year. In each year of the sample, the interregional share is significantly greater than what would be expected if incomes were randomly distributed in space. This result is of particularly interest, as it was the Census Regions partition that had a smaller interregional share relative to its intraregional share. By extending the traditional decompositional analysis to include an inferential component we find that, although this interregional inequality component is relatively small, the Census Regions do capture some aspect of spatial structure. Without the inferential test, this partition might have been viewed as irrelevant or misspecified given that the interregional share was found to be stronger in the other partitions.

5 Conclusion

In their overview of recent empirical work on economic growth Durlauf and Quah (1999) conclude that the field remains in its infancy (pg. 295). One sign of increasing maturity is the recent attention given to the geographical dimensions of economic growth (Barro and Sala-i-Martin, 1991; Krugman, 1991; Nijkamp and Poot, 1998). Application of recently developed methods of spatial econometrics to the question of regional convergence has yielded a more comprehensive and multidimensional view of regional economic growth (Goodchild et al., 2000).

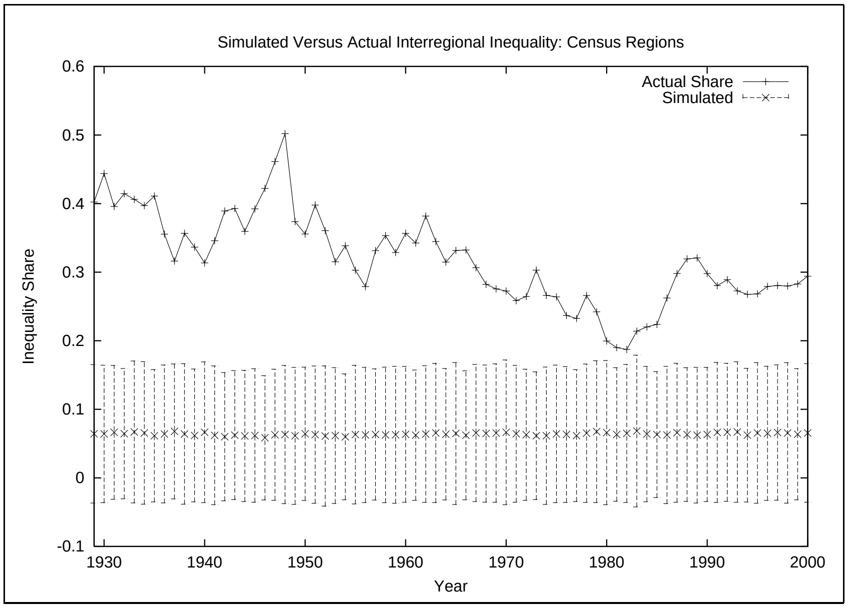

Figure 7: Simulated Versus Actual Interregional Inequality, Census Regions: State Unit of Observation

The literature on regional income inequality, although somewhat older than the convergence literature, has been slower to adopt new spatially explicit methods of data analysis. This chapter has attempted to contribute to such an adoption by investigating the role of spatial dependence and spatial scale in the analysis of regional income inequality in the US over the 1929-2000 period. The key findings include with regard to spatial dependence include:

- A strong positive relationship between measures of inequality in state incomes and the degree of spatial autocorrelation.
- A strong positive relationship between the interregional inequality share (as opposed to intraregional inequality) and spatial clustering.

The analysis of the role of the spatial scale of the observational unit and the choice of regional partitioning of the units revealed the following:

- The qualitative and quantitative results of inequality decomposition are highly sensitive to the scale of the observational unit. Interregional inequality is dominant when state data are used, yet intraregional inequality is most important when county level data are used.
- The relative importance of the interregional inequality component is *not* a simple function of the number of groups used in a partitioning of the regional observations.

Finally, the study also suggested an approach to inference in the decompositional analysis of regional income inequality, offering an important complement to the existing literature that has relied exclusively on descriptive methods.

References

Alonso, W. (1980). Five bell shapes in regional development. *Papers of the Regional Science Association*, 45:5–16.

Amos, Jr, O. (1983). The relationship betwen personal income inequality, regional income inequality, and development. *Regional Science Perspectives*, 13:3–14.

Amos, Jr, O. (1988). Unbalanced regional growth and regional income inequality in the latter stages of development. *Regional Science and Urban Economics*, 18:549–566.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer, Boston.

Armstrong, H. and Vickerman, R. (1995). *Convergence and Divergence Among European Regions*. Pion, London.

Azzoni, C. R. (2001). Economic growth and income inequality in Brazil. *Annals of Regional Science*, 31(1):133–152.

Baker, J. L. and Grosh, M. E. (1994). Measuring the effects of geographic targeting on poverty reduction. *World Bank Living Standards Measurement Study*.

Barro, R. and Sala-i-Martin, X. (1991). Convergence across states and regions. *Brookings papers on Economic Activity*, 1:107–182.

Chatterji, M. and Dewhurst, J. (1996). Convergence clubs and relative economic performance in Great Britan. *Regional Studies*, 30:31–40.

Cliff, A. and Ord, J. (1973). *Spatial Autocorrelation*. Pion, London.

Conceição, P. and Ferreira, P. (2000). The young person's guide to the Theil index: suggesting intuitive interpretations and exploring analytical applications. University of Texas Inequality Project Working Paper 14.

Cowell, F. A. (1995). *Measuring Inequality*. Prentice Hall, London.

Cowell, F. A. and Jenkins, S. P. (1995). How much inequality can we explain? Methodology and an application to the United States. *The Economic Journal*, 105(429):421–430.

Cuadrado-Roura, J., Garcia-Greciano, B., and Raymond, J. (1999). Regional convergence in productivity and productivity structure: The Spanish case. *International Regional Science Review*, 22:35–53.

Durlauf, S. N. and Quah, D. T. (1999). The new empirics of economic growth. In Taylor, J. B. and Woodford, M., editors, *Handbook of Macroeconomics: Volume 1A*, pages 235–308. Elsevier, Amsterdam.

Eff, E. A. (1999). Myrdal contra Ohlin: Accounting for sources of U.S. county per capita income convergence using a flexible decomposition approach. *Review of Regional Studies*, 29(1):13–36.

Fan, C. C. and Casetti, E. (1994). The spatial and temporal dynamics of US regional income inequality, 1950-1989. *Annals of Regional Science*, 28:177–196.

Fingleton, B. (1999). Estimates of time to economic convergence: An analysis of regions of the Eurpean Union. *International Regional Science Review*, 22:5–35.

Fingleton, B. (2001). Regional economic growth and convergence: insights from a spatial econometric perspective. In Anselin, L. and Florax, R., editors, *Advances in Spatial Econometrics*. Springer-Verlag, Berlin.

Fujita, M. and Hu, D. (2001). Regional disparity in China 1985-1994: The effects of globalization and economic liberalization. *Annals of Regional Science*, 35:3–37.

Goodchild, M., Anselin, L., Applebaum, R., and Harthorn, B. H. (2000). Towards a spatially integrated social science. *International Regional Science Review*, 23:139–159.

Kanbur, R. and Zhang, X. (1999). Which regional inequality? The evolution of rural-urban and inland-coastal inequality in China, 1983-1995. *Journal of Comparative Economics*, 27:686–701.

Krugman, P. (1991). *Geography and Trade*. MIT Press, Cambridge.

Kuznets, S. (1955). Economic growth and income equality. *American Economic Review*, 45:1–28.

Levy, A. and Chowdhury, K. (1995). A geographical decomposition of intercountry income inequality. *Comparative Economic Studies*, 37:1–17.

Lucas, R. E. (1993). Making a miracle. *Econometrica*, 61(2):251–271.

Maasoumi, E. (1997). Empirical analysis of inequality and welfare. In Pesaran, M. and Schmnidt, P., editors, *Handbook of Applied Econometrics, Volume II, Microeconomics*, pages 202–245. Blackwell Publishers Ltd., London.

Maxwell, P. and Peter, M. (1988). Income inequality in small regions: A study of Australian statistical divisions. *The Review of Regional Studies*, 18:19–27.

Mills, J. A. and Zandvakili, S. (1997). Statistical inference via bootstrapping for measures of inequality. *Journal of Applied Econometrics*, 12(2):133–150.

Nijkamp, P. and Poot, J. (1998). Spatial perspectives on new theories of economic growth. *Annals of Regional Science*, 32:407–437.

Nissan, E. and Carter, G. (1999). Spatial and temporal metropolitan and nonmetropolitan trends in income inequality. *Growth and Change*, 30(3):407–415.

Opensaw, S. and Alvanides, S. (1999). Applying geocomputation to the anaylsis of spatial distributions. In Longley, P., Goodchild, M., Maguire, D., and Rhind, D., editors, *Geographic Information Systems Volume I: Principles and Technical Issues*, pages 267–282, New York. John Wiley.

Paci, F. and Pigliaru, F. (1997). Structural change and convergence: an Italian perspective. *Structural Change and Economic Dynamics*, 8:297–318.

Rey, S. J. (2001). Space-time analysis of regional systems: STARS. Technical report, Department of Geography, San Diego State University.

Rey, S. J. and Montouri, B. D. (1999). U.S. regional income convergence: A spatial econometric perspective. *Regional Studies*, 33:143–156.

Shorrocks, A. F. (1984). Inequality decomposition by population subgroups. *Econometrica*, 52(6):1369–1385.

Theil, H. (1967). *Economics and Information Theory*. North Holland, Amsterdam.

Theil, H. (1996). *Studies in Global Econometrics*. Kluwer Academic Publishers, Dordrecht.

Tsui, K. Y. (1993). Decomposition of China's regional inequalities. *Journal of Comparative Economics*, 17:600–627.

Williamson, J. (1965). Regional inequality and the process of national development. *Economic Development and Cultural Change*, 4:3–47.

Zhang, X. and Kanbur, R. (2001). What difference do polarisation measures make? An application to China. *Journal of Development Studies*, Forthcoming.

Zheng, B. and Cushing, B. J. (2001). Statistical inference for testing inequality indices with dependent samples. *Journal of Econometrics*, 101:315–335.