Journal of the Korean Geographical Society, Vol.39, No.2, 2004(212~228)

Spatial Data Analysis for the U.S. Regional Income Convergence, 1969-1999: A Critical Appraisal of β-convergence

Sang-Il Lee*

미국 소득분포의 지역적 수렴에 대한 공간자료 분석(1969～1999년)
- 베타-수렴에 대한 비판적 검토 -

이 상 일*

Abstract : This paper is concerned with an important aspect of regional income convergence, β-convergence, which refers to the negative relationship between initial income levels and income growth rates of regions over a period of time. The common research framework on β-convergence which is based on OLS regression models has two drawbacks. First, it ignores spatially autocorrelated residuals. Second, it does not provide any way of exploring spatial heterogeneity across regions in terms of β-convergence. Given that empirical studies on β-convergence need to be edified by spatial data analysis, this paper aims to: (1) provide a critical review of empirical studies on β-convergence from a spatial perspective; (2) investigate spatio-temporal income dynamics across the U.S. labor market areas for the last 30 years (1969-1999) by fitting spatial regression models and applying bivariate ESDA techniques. The major findings are as follows. First, the hypothesis of β-convergence was only partially evidenced, and the trend substantively varied across sub-periods. Second, a SAR model indicated that β-coefficient for the entire period was not significant at the 99% confidence level, which may lead to a conclusion that there is no statistical evidence of regional income convergence in the US over the last three decades. Third, the results from bivariate ESDA techniques and a GWR model report that there was a substantive level of spatial heterogeneity in the catch-up process, and suggested possible spatial regimes. It was also observed that the sub-periods showed a substantial level of spatio-temporal heterogeneity in β-convergence: the catch-up scenario in a spatial sense was least pronounced during the 1980s.

Key Words : β-convergence, ESDA (exploratory spatial data analysis), spatial association measures, spatial autoregressive models, GWR (geographically weighted regression)

요약 : 본 연구는 지역간 소득분포의 수렴/발산의 주요 측면인 베타-수렴을 공간자료분석에 의거하여 비판적으로 검토하고 있다. 베타-수렴에 대한 통상적인 접근법은 두 가지 측면에서 문제점을 갖고 있다. 첫째, 회귀분석 결과 도출되는 잔차의 공간적 자기상관을 고려하지 못한다. 둘째, 베타-수렴의 국지적 변이, 즉 공간적 이질성을 탐색할 어떠한 절차도 제공하지 못한다. 이러한 비판적 검토를 바탕으로, 다양한 공간자료분석 기법들, 즉, 공간적 자기회기 모델(spatial autoregressive models), 이변량 국지통지(bivariate local statistics)를 이용한 탐색적 공간자료분석(ESDA: exploratory spatial data analysis) 기법, 그리고 지리적 가중회귀분석(GWR: geographically weighted regression)을 사용하여 1969-1999년 간의 미국 노동시장지역에 대한 소득 자료를 분석하였다. 주요 결과는 다음과 같다. 첫째, OLS모델을 적용한 결과 베타-수렴은 단지 부분적으로만 드러났고, 베타-수렴 계수도 시기별로 상당한 편차를 보였다. 둘째, 공간적 자기회기 모델의 분석 결과 OLS에 의해 유의한 것으로 나타난 베타-수렴 계수가 99% 신뢰수준에서 유의하지 않은 것으로 드러났다. 셋째, 탐색적 공간자료분석과 지리적 가중회귀분석의 결과는 베타-수렴의 경향에 상당한 정도의 공간적 이질성이 존재한다는 점을 보여주고 있다. 또한 이 공간적 이질성의 양상이 시기별로도 다양하게 드러남이 관찰되었다.

주요어 : 베타-수렴, 탐색적 공간자료분석, 공간적 연관 통계치, 공간적 자기회귀 모델, 지리적 가중회귀분석

* Full-time instructor, Department of Geography Education, Seoul National University, si_lee@snu.ac.kr

1. Introduction

Regional income distribution has increasingly attracted much attention from various academic fields. A profound wave of socio-economic restructuring and the advent of the European Union as an inter-regional integration might precipitate the enormous dedication to the topic. Empirical studies on the subject have sought to investigate whether a national economy has evolved towards a spatial convergence referring to a decreasing gap between richer regions and poorer regions, mainly focusing on two aspects of regional income convergence, σ-convergence and β-convergence.

This paper is focused on the second aspect of regional income convergence, β-convergence which refers to the negative relationship between initial income levels and income growth rates of regions over a period of time. As suggested (Lee, 2004c), β-convergence may be more problematic than σ-convergence, the reduction of dispersions or variances in per capita income across regions, in the sense that the ignorance of spatial effects such as spatial dependence and spatial heterogeneity could yield unsustainable results in empirical studies. There are two obvious drawbacks.

First, statistical tests for regression coefficients in OLS (ordinary least squares) regression models could be flawed in the presence of spatial autocorrelation in regression residuals. In other words, statistically significant β-coefficients could turn out to be insignificant if there exists a substantive level of positive spatial dependence in the dependent variable, that is, regional income growth rates. This precipitates the need of using a particular form of spatial regression.

Second, spatial heterogeneity should be explored. A negative relationship between initial income levels and income growth rates in a global sense does not necessarily mean that the relationship holds for all the regions involved: some regions could show a positive relationship; other regions could report a different level of the negative relationship. Thus, various ESDA (exploratory spatial data analysis) techniques utilizing bivariate spatial association measures such as Lee's L_i (Lee, 2004b) should be actively engaged in a research framework.

Given that empirical studies on β-convergence need to be edified by spatial data analysis, this paper aims to: (1) provide a critical review of empirical studies on β-convergence from a spatial perspective; (2) investigate spatio-temporal income dynamics across the U.S. labor market areas for the last 30 years (1969-1999) by fitting spatial regression models and applying bivariate ESDA techniques.

2. A Critical Review on β-convergence: Numerical Correlation vs. Spatial Co-patterning

β-convergence was introduced by Barro (1991), and Barro and Sala-i-Martin (1991) based on the neoclassical growth theory to capture the catch-up hypothesis that poorer regional economies grow faster than richer regional economies. There will be β-convergence if a negative relation is found between the initial level of income and the growth rate of per capita income (sometimes referred to as 'regression to the mean' or 'mean reversion') (Sala-i-Martin, 1996:1327). This 'weak convergence' (Nijkamp and Poot, 1998:26), as apposed to 'strong convergence' of σ, often takes a regression form as:

$$\frac{1}{k} \cdot \ln\left(\frac{y_{i,t+k}}{y_{i,t}}\right) = \alpha - \beta \cdot \ln(y_{i,t}) + \varepsilon_{i,t} \qquad (1)$$

where k is a year-differential, $y_{i,t}$ is income level in region i at a starting year, $y_{i,t+k}$ is income level in region i at an ending year. However, a more complicated form has been preferred in empirical studies that is given (Sala-i-Martin, 1996; Nijkamp and Poot, 1998):

$$\frac{1}{k} \cdot \ln\left(\frac{y_{i,t+k}}{y_{i,t}}\right) = \alpha - \left(\frac{1 - e^{-\beta \cdot k}}{k}\right) \cdot \ln(y_{i,t}) + \varepsilon_{i,t} \qquad (2)$$

From equation 1, an estimated value of β is a slope coefficient in a regression of regional income growth rates on initial regional income levels (Nijkamp and Poot, 1998; Martin, 2001), which constitutes a reason why works based on β-convergence have been called 'growth regression approach' (Martin, 2001:62). More often, however, the parameter of β is interpreted as the speed at which economies approach their own steady states or poorer regions catch up with richer ones (Sala-i-Martin, 1996; Kangasharju, 1999). When some other shock variables, such as industrial mix and regional dummies, are included in the equation, it calibrates 'conditional convergence' as opposed to 'absolute convergence' attributing to the original specification (Barro and Sala-i-Martin, 1991; Armstrong, 1995b; Sala-i-Martin, 1996; Dickey, 2001). From both equations, a positive value of β indicates economic convergence across regions.

β-convergence is usually preferred over σ-convergence in empirical studies, because the former conveys more information about regional income convergence than the latter. First, β-convergence is a necessary condition for σ-convergence (Nijkamp and Poot, 1998). Without β-convergence, σ-convergence won't happen. In other words, a substantial change in the ranking of regions in economic performance could happen without being captured by σ-convergence. Thus, β-convergence does not imply σ-convergence (Barro and Sala-i-Martin, 1991; Sala-i-Martin, 1996; Nijkamp and Poot, 1998; Kangasharju, 1999; Tsionas, 2000). However, there are disagreements: the convergence rate does not mean that a poorer region catches up with a richer region at that rate (Tsionas, 2000); the convergence may be unrelated to, or uninformative for, the dynamics of economic growth (Quah, 1996b).

Table 1 summarizes empirical studies on regional income dynamics based on β-convergence. Almost all studies listed report consistent convergence at a β of about 0.02, which means that regions, wherever they are, tend to converge at a speed of approximately 2% per year (Barro and Sala-i-Martin, 1991; Armstrong, 1995a; Sala-i-Martin, 1996). For example, β-coefficient for 48 US states was estimated at 0.017 during 1880-1990 (Sala-i-Martin, 1996). This striking coincidence across numerous countries or supranational entities such as the European Union arguably verifies the virtues of neo-classical growth theory. Although the original formulation is based on an assumption of closed Solow economies (Blanchard, 1991, 159), additional considerations such as labor mobility, capital mobility, and technology transfer adjust the theorem equally working across open economies such as countries and supranational regimes (Barro and Sala-i-Martin, 1991; Blanchard, 1991; Armstrong, 1995b).

The alleged myth of regional income convergence, however, has been challenged by theoretical arguments and empirics. Quah (1996b, 1355) contends that "uniformity is due to something relatively uninteresting, namely, the statistical implications of a unit root in the time-series data". The unit root problem refers to non-stationarity in residuals of a time-series OLS regression (Kennedy1998, 268-269). In the same vein, Martin (2001, 62) argues that "the growth regression approach has an inbuilt bias towards identifying convergence, so that the results may even over-estimate what little convergence has occurred." Tsionas (2000) reports that there is a positive relationship between initial levels of regional income and income growth rates. A finding by most ardent advocates for the convergence thesis (Barro and Sala-i-Martin, 1991) says that β-coefficient turns negative during the 1980s implying a possible divergence in recent years.

Again, I would argue that studies based on β-convergence should be edified by findings from spatial data analysis.

First, as far as the regression equation is calibrated based on the OLS algorithm, the approach is obviously subject to problems of spatially autocorrelated errors. Given various types of spatial interactions among regions, adjacent regions tend to show a sim-

Table 1. Empirical studies on β-convergence

Spatial Unites		Studies	Years
Europe	EU regions	Barro and Sala-i-Martin (1991)	1950-1985
		Barro and Sala-i-Martin (1995)	1950-1990
		Armstrong (1995a)	1950-1980
		Armstrong (1995c)	1975-1992
		Dewhurst and Mutis-Gaitan (1995)	1981-1991
		Sala-i-Martin (1996)	1955-1990
		European Commission (1997)	1975-1993
		Button and Pentecost (1999)	1975-1990
	UK regions	Sala-i-Martin (1996)	1950-1990
	France regions	Sala-i-Martin (1996)	1950-1990
	Germany regions	Sala-i-Martin (1996)	1950-1990
	Italy regions	Sala-i-Martin (1996)	1950-1990
	Spain regions	Mas et al. (1995)	1955-1991
		Sala-i-Martin (1996)	1950-1990
		Cuadrado-Roura et al. (1999)	1955-1993
	Finland regions	Kangasharju (1998)	1974-1993
US	US states	Barro and Sala-i-Martin (1991)	1880-1988
		Blanchard and Katz (1992)	1950-1990
		Evans and Karras (1996)	1929-1991
		Sala-i-Martin (1996)	1880-1990
		Sum and Fogg (1999)	1939-1996
		Tsionas (2000)	1978-1996
Others	OECD countries	Andres et al. (1996)	1960-1990
		de la Fuente (1997)	1960-1985
	Canada provinces	Sala-i-Martin (1996)	1961-1991
	Japan prefectures	Sala-i-Martin (1996)	1955-1990

ilar trend in economic performance, which will be reflected in regression residuals. Quah (1996a, 954) correctly points out that "no region can be studied in isolation independently of others". When significant spatial autocorrelation is present in residuals, significance tests for coefficients may be flawed even though coefficients themselves are still unbiased (Anselin and Griffith, 1988; Fotheringham and Rogerson, 1993). Thus, the regression equations in (1) and (2) should be calibrated by a spatial autoregressive model (Armstrong, 1995b; Molho, 1995; Bernat, 1996; European Commission, 1997; Mencken, 1998; Buettner, 1999; Fingleton, 1999; Rey and Montouri, 1999; Pons-Novell and Viladecans-Marsal, 1999; Martin, 2001). When this problem is associated with other statistical symptoms such as non-normality, structural instability, and misspecification (Tsionas, 2000), the whole research becomes unsustainable.

Second, as Martin (2001, 62) correctly points out, the β-convergence approach is based on an unreliable assumption that "the underlying convergence process is identical across all regions, whereas in reality it is may well vary from region to region, or between different types or groups of regions". This resonates with Quah's argument (1996a, 954) that "regression-based approaches, averaging across either cross-section or time series dimensions, are not useful ⋯ such methods construct a representative, and cannot provide a picture of how the entire cross-section distribution evolves". This issue of spatial heterogeneity can be addressed by some other multivariate spatial statistical techniques such as expansion method (Casetti and Jones, 1987), expanded rank-size function (Fan and Casetti, 1994; Lopez-Bazo et al., 1999), Markov chain matrix (Quah, 1993, 1996b; Fingleton, 1997, 1999; Rey, 2001; Bickenbach and Bode, 2003), and geographically weighted regression (GWR) (among others, Brunsdon et al., 1996, 1998; Fotheringham et al., 1998, 2002).

Third, as far as 'absolute convergence' without any other additional variables is concerned, β-convergence is nothing but a correlation between initial income levels and income growth rates. A β coefficient from an OLS regression is directly related to correlation between two variables. When bivariate relations between initial income levels and income growth rates are spatially clustered, a global bivariate spatial association measure should replace aspatial correlation measures such as Pearson's r. Furthermore, there is a good reason to believe that the averaged correlation coefficient does not apply to the whole study area. Rather, local correlations may be highly heterogeneous. In the context of income convergence, some initially poor regions may have accomplished a certain level of catch-up, whereas some others may still fall behind. This spatial heterogeneity can only be tackled by a local bivariate spatial association measure and the related ESDA techniques.

3. Spatio-temporal Income Dynamics across the US Labor Market Areas, 1969-1999: β-convergence

1) Research Design

The main data source for this study is per capita personal income data which are collected and maintained by the Bureau of Economic Analysis and are available via REIS (Regional Economic Information Systems) at the county level from 1969 to 1999. The county level income data are aggregated into 391 labor market areas (LMAs) for the conterminous U.S. The delineation of LMAs is first based on communing flow matrix among counties, and a hierarchical cluster analysis aggregates 3,141 counties into 741 commuting zones (CZs). The CZs are then aggregated into 394 LMAs in terms of a minimum population requirement (100,000) and inter-CZ commuting flows (Tolbert and Sizer, 1996). In this study, three LMAs in Alaska and Hawaii are eliminated. Next, the entire thirty years are divided into three sub-periods, 1969-1979, 1979-1989, and 1989-1999, and four years, 1969, 1979, 1989, and 1999, are utilized as benchmarks to provide particular snapshots, allowing for tracking spatio-temporal evolution.

This empirical study is divided into three parts.

First, β-convergence, a negative relationship between an initial income level and income growth rate between years, is critically evaluated. Global Pearson's r and Lee's L (Lee, 2001a, 2004a) are computed for different sub-periods. Lee's L was developed to capture not only numerical correlation but also spatial co-patterning between two geographical variables or spatial patterns, and the equation is given by:

$$L = \frac{n}{\sum_i \left(\sum_j v_{ij} \right)^2} \cdot \frac{\sum_i \left[\left(\sum_j v_{ij}(x_j - \bar{x}) \right) \cdot \left(\sum_j v_{ij}(y_j - \bar{y}) \right) \right]}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \quad (3)$$

where v_{ij} is an element in a general spatial weights matrix V. Next, spatial autocorrelation in OLS residuals is assessed and a SAR (simultaneous autoregressive) model is introduced to alleviate the problem of spatially autocorrelated errors. Spatial patterns resulting from a decomposition of a SAR model effectively demonstrate the necessity of using spatial autoregressive models when spatial autocorrelation in residuals is significant.

Second, the spatial heterogeneity of β-convergence is investigated. Such bivariate ESDA techniques as Local-r and local-L scatterplot maps and local-L significance maps (for a detailed description about the techniques, see Lee, 2001b, 2004b) are utilized to explore spatial variation in β-convergence. Local-r and local-L scatterplot maps are based on local statistics, local r_i and local L_i which are respectively given by:

$$r_i = n \cdot \frac{(x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \tag{4}$$

$$L_i = \frac{n^2}{\sum_i \left(\sum_j v_{ij} \right)^2} \cdot \frac{\left(\sum_j v_{ij}(x_j - \bar{x}) \right) \cdot \left(\sum_j v_{ij}(y_j - \bar{y}) \right)}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}} \tag{5}$$

where v_{ij} is an element in a general spatial weights matrix V. These measures and the associated ESDA techniques are used to investigate whether the general trend of a negative relationship between initial income levels and income growth rates is spatially even; some LMAs may follow the trend, but others may not.

Third, a geographically weighted regression (GWR) model will be fitted. The technique is expected to capture spatial variations of regression parameters and thus spatial heterogeneity or non-stationarity in bi- or multi-variate situations. The results from this model will be compared to those from bivariate ESDA analyses.

All the measures and ESDA-techniques were implemented in an ESDA-GIS framework; an S-PLUS-ArcView connection plays a main role for the platform. All the scripts were coded in S, the script language in S-PLUS by the author and the results were transferred to ArcView where cartographic visualization tasks were undertaken.

2) β-convergence and Spatially Autocorrelated Errors

A negative relationship between initial income levels and income growth rates during a given period of time constitutes the rationale of β-convergence. Figure 1 shows maps of 1969 per capita incomes and growth rates, 1969-1999. It seems that there is a high negative correlation. Table 2 lists correlations between the two map patterns and other sub-period pairs. Pearson's correlation coefficient column indicates that there is a significant negative relationship between initial income levels and income growth rates during 1969-1999, which may confirm the income convergence hypothesis. However, the last period, 1989-1999, show a very poor correlation between the two variables. This means that many high or low LMAs have respectively experienced high or low income growth, which lowers the tendency towards β-convergence.

The Lee's L column, however, reports contrasting information that needs to be explained. Albeit the lowest Pearson's r (-0.175) in 1989-1999 across LMAs, Lee's L is larger than one for 1979-1989. Since Lee's L not only captures a point-to-point association that Person's r does, but also spatial co-patterning (Lee, 2001b), the higher L value in the period of 1989-1999 than 1979-1989 suggests that the catch-up process happened in a more spatially clustered manner in the former than in the latter.

Another finding is that the highest negative correlation in both columns is found in the period of 1969-1979. Especially Pearson's r for 1969-1979 is higher in magnitude than that for the entire period, 1969-1999. It can be concluded the catch-up forces were most dramatic in the 1970s and then finally faded away in the 1990s.

Figure 2 shows an OLS regression between logarithmic 1969 income levels and annual income growth rates between 1969 and 1999, following equation 1. The slope is -0.010 and R-squared is 0.231 (Table 3). Therefore, β-coefficient for the US LMAs over 30 years is .010 that means that US regional income has converged at a speed of 1% annually. This is too low to conform to the myth of 2% convergence and the exploratory power is rather low. Further, the trend towards income convergence varies sub-period to sub-period. Table 3 lists different β-coefficients for different sub-periods: 0.022 for 1969-1979, 0.012 for 1979-1989, and 0.005 for 1989-

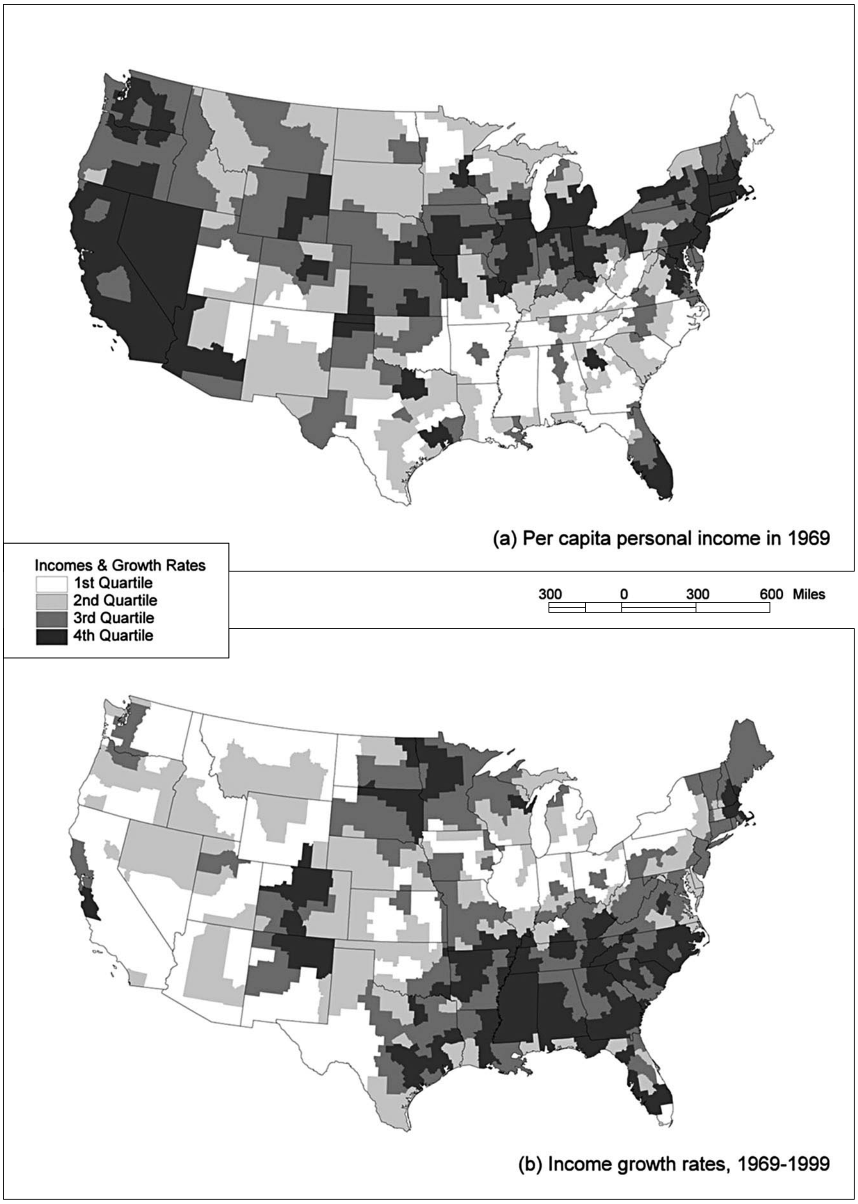


Figure 1. Per capita personal income in 1969 and annual income growth rates, 1969-1999

Table 2. Correlations between initial income levels and income growth rates between years

	Periods	Pearson's *r*	Lee's *L*
Whole period	1969-1999	-0.481*	-0.440*
Sub-periods	1969-1979	-0.548*	-0.385*
	1979-1989	-0.201*	-0.177*
	1989-1999	-0.175*	-0.276*

* significant at $\alpha = 0.01$

1999. Obviously, the thesis of income convergence works best for the 1970s and worst for the 1990s. 1989 income levels only explain 3% of total variance in income growth rates. It thus may be concluded that there has been no β-convergence since the early 1980s.

As discussed, the presence of spatial autocorrela-

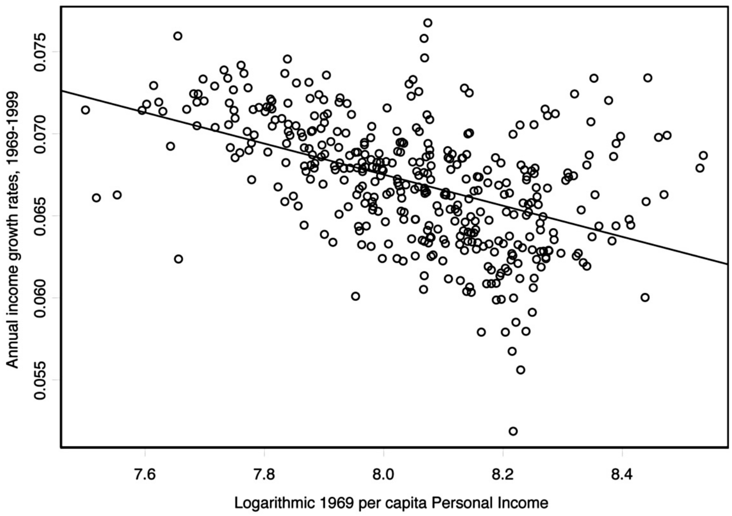


Figure 2. An OLS regression between 1969 logarithmic per capita personal income and annual income growth rates across the US LMAs, 1969-1999

tion in OLS residuals may invalidate the significance of regression coefficients. Table 3 shows that Moran's I tests find a significant spatial autocorrelation in OLS residuals for all sub-periods let alone the entire period. This necessitates the use of spatial autoregressive models. Here, I utilize a SAR (simultaneous autoregressive) model. Following Tiefelsdorf's notation (2000, 43-44), a SAR model is written:

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \text{ and } \varepsilon = \rho\mathbf{V}\varepsilon + \eta, \text{ therefore,} \\ \mathbf{y} = \underbrace{\mathbf{X}\beta}_{trend} + \underbrace{\rho\mathbf{V}\varepsilon}_{signal} + \underbrace{\eta}_{noise} \tag{6}$$

where ε is the correlated error term and η is a random white noise. From equation (6), variation of a dependent variable is decomposed into three parts, respectively what Haining (1990, 258-259; 2003, 333) calls *trend*, *signal*, and *noise*. If there is no spatial autocorrelation, ρ, spatial autocorrelation coefficient, will

Table 3. OLS and SAR models for logarithmic initial per capita personal income levels and annual income growth rates

	Dependent variable: Income growth rate							
	1969-1999		1969-1979		1979-1989		1989-1999	
	OLS	SAR [a]	OLS	SAR [a]	OLS	SAR [a]	OLS	SAR [a]
(intercept)	0.143*	0.085*	0.272*	0.239*	0.170*	0.062*	0.095*	-0.004*
	(20.31)	(10.41)	(19.58)	(15.16)	(6.60)	(2.80)	(6.40)	(-0.25)
Initial income level	-0.010*	-0.002	-0.022*	-0.018*	-0.012*	0.0002	-0.005*	0.005
	(-10.81)	(-2.34)	(-12.93)	(-9.35)	(-4.05)	(0.94)	(-3.45)	(2.92)
R-squared	0.231		0.301		0.040		0.030	
Moran's I [b]	0.283*	-0.050	0.373*	0.003	0.573*	-0.063	0.247*	-0.037
ρ [c]		0.123		0.118		0.161		0.116

* : significant at $\alpha = 0.01$

t-values in parentheses

[a] : Simultaneous autoregressive model

[b] : Moran's I test for regression residuals

[c] : Spatial autocorrelation parameter

be zero, thus, variance of a dependent variable is decomposed into vectors of predicted values and non-correlated errors. Table 3 reports that ρ-coefficient for all sub-periods is not negligible, and is highest in 1979-1989. By applying $\varepsilon = \mathbf{y} - \mathbf{X}\beta$ to the equation, we have:

$$\mathbf{y} = \mathbf{X}\beta + \rho\mathbf{V}\mathbf{y} - \rho\mathbf{V}\mathbf{X}\beta + \eta \tag{7}$$

This equation allows a further decomposition (Tiefelsdorf 2000, 44): (i) the spatially independent influences of the exogenous component $\mathbf{X}\beta$; (ii) the spatially dependent endogenous observations $\rho\mathbf{V}\mathbf{y}$; (iii) the spatial trend values $\rho\mathbf{V}\mathbf{X}\beta$; (iv) independent disturbances η. Further, the variance-covariance matrix $\Omega(\rho)$ among the error terms can be written (Tiefelsdorf, 2000, 44):

Figure 3 shows spatial patterns of decomposition of annual income growth rates between 1969 and 1999 based on equation (6). What a SAR model does is to decompose residuals into spatial autocorrelated errors (signal) and non-autocorrelated ones (noise). The signal map in Figure 3 shows that positive residuals are spatially clustered in the South. By eliminating spatially autocorrelated parts from residuals, the noise map rarely displays spatial autocorrelation. Moran's *I* test for noise resulting from SAR models in Table 3 does not reject the null hypothesis that there is no spatial dependence in residuals. A crucial finding is that SAR models significantly lower t-values of β-coefficients. Even though all the β-coefficients in OLS models are significant at the 99% confidence level, the SAR counterparts are not except for one in the 1969-1979 model. This implies, as Bailey and Gatrell (1995, 285) indicate, that OLS models tend to inflate the significance of regression coefficients. Thus, an ultimate conclusion is that there is

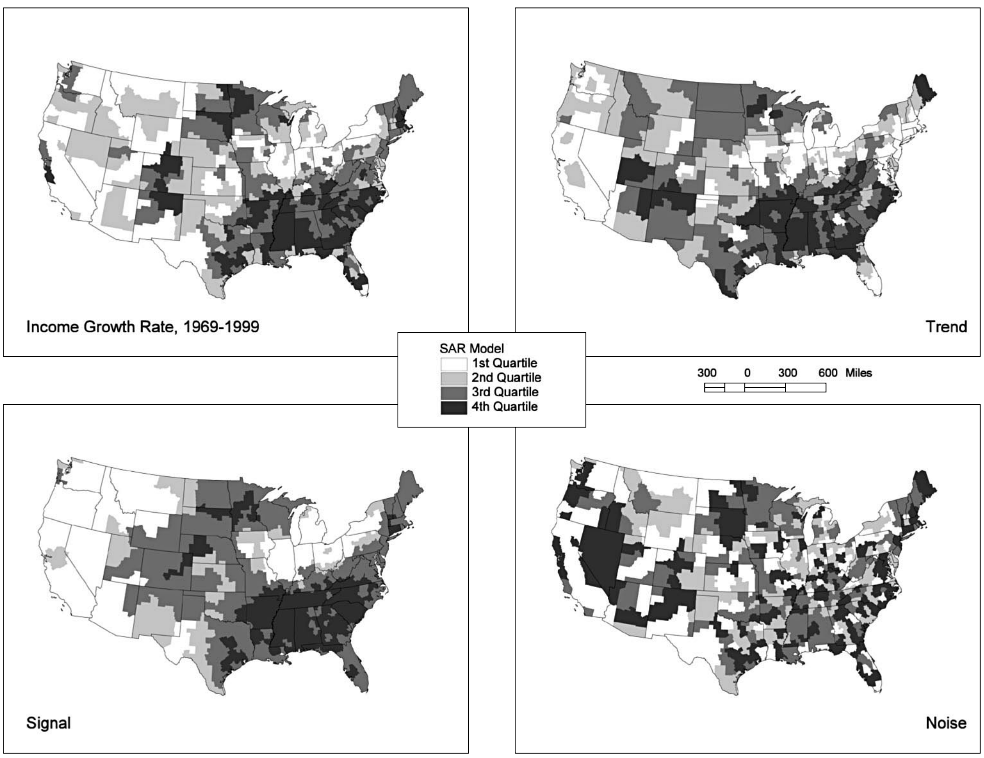


Figure 3. SAR model decomposition: 1969 logarithmic per capita personal income and annual income growth rates, 1969-1999

no statistical evidence of β-convergence in the US over the last three decades.

3) Spatial Heterogeneity in β-convergence

It is noteworthy that the negative relationship between initial income levels and income growth rates should not be assumed to apply to an entire study region. The relationship may be positive for some locales; some characterized by lower-than-average 1969 income level and lower-than-average income growth rates during 1969-1999; others in an opposite direction.

Figure 4 displays local-r and local-L scatterplot maps. The latter is simply a spatially smoothed version of the former, which may benefit pattern detection. From Figure 4(a), one can notice that urban effects are dominant for high-high association. They started at higher-than-average income levels in 1969 and have enjoyed a higher-than-average income growth rates during the last 30 years. Figure 4(b)

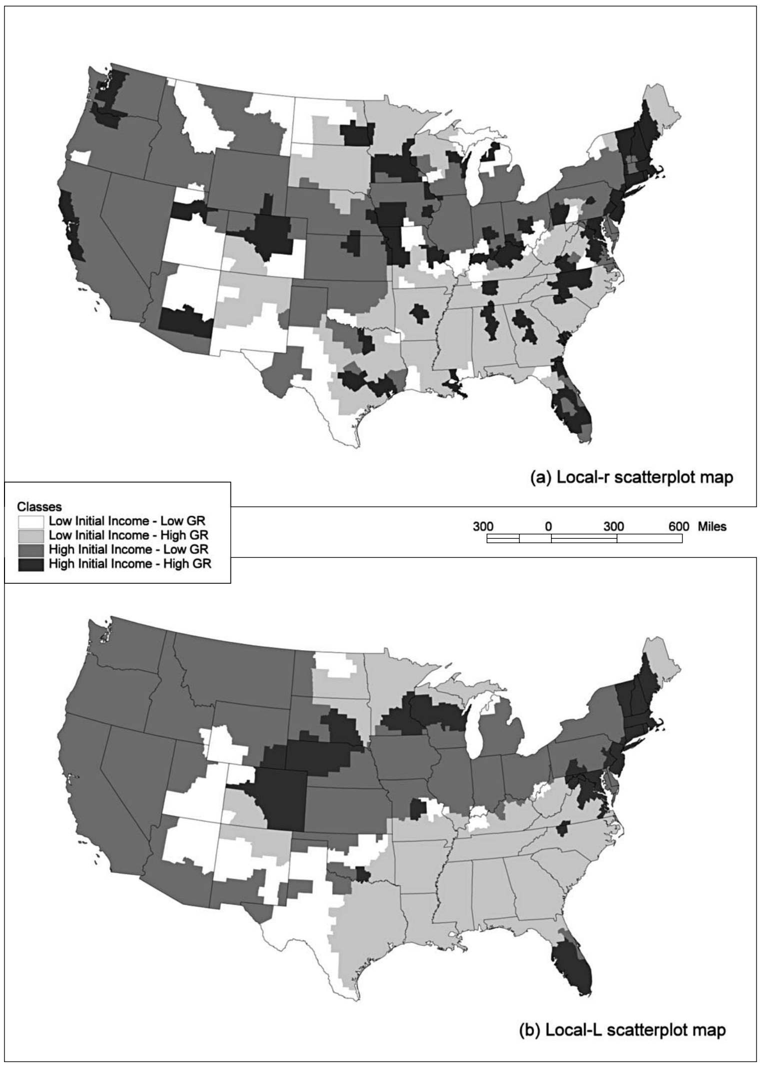


Figure 4. Local-r and local-L scatterplot maps of 1969 logarithmic per capita personal income and annual income growth rates, 1969-1999

may help generalize spatial patterns. Areas with higher-than-average income levels in 1969 are associated with lower-than-average growth rates. In contrast, areas with lower-than-average income levels in 1969 are associated with higher-than-average growth rates. Figure 4(b) also reveals that there is a structural distinction among areas characterized by the continuation of lower income levels; areas from the Mountain region down to western Texas are discernable from ones in the South and the northwestern part of the Midwest; the former has never been involved in the catch-up process; the latter has positively contributed to the catch-up scenario.

Figure 5 shows local-L significance maps for the entire period and three sub-periods. The map for 1969-1999 selects LMAs from Figure 4(b) that are significant. Some interesting patterns are detected. First, the northern part of Megalopolis, southern Florida, and Denver areas have significantly built on their initial higher-than-average income level. Second, economic slowdowns have mostly occurred in the Midwest and the Pacific. Third, most areas in the South except for several urban centers in the Pediment turn out to be significant spatial clusters for the β-convergence; that is, they conform to the 'catch-up' scenario.

Each sub-period, however, displays a particular level and form of spatial heterogeneity in β-convergence. In the 1970s, deindustrialization in the traditional industrial belt (or rust belt) and industrialization in the South centered on the lower Mississippi constituted a dominant pattern. The 1980s experienced re-orientation towards the Megalopolis and southern Florida, marked deindustrialization in the Pacific, and economic slowdowns in the Great Plain, and economic upswings in the South centered on

Figure 5. Local-L significance maps: initial logarithmic per capital personal income and annual income growth rates, 1969-1999

the Pediment. Finally, the 1990s is characterized by economic slowdowns in the Megalopolis, the California region, and southern Florida, and economic revitalization in the South centered on the lower Mississippi.

As a conclusion, the catch-up scenario in a spatial sense was least pronounced during the 1980s, which can be seen from the fact that the 1979-1989 map (Figure 5) displays that LMAs falling into the classes of low-low and high-high associations appears most extensively, and that those areas are most spatially clustered.

4) A Geographically Weighted Regression (GWR): Spatially Drifting β-coefficients

Another way to investigate spatial heterogeneity in statistical parameters is to fit a geographically weighted regression (GWR) model. This approach is simply a combination of weighted least squares (WLS) regression and kernel regression (Schimek 2000). However, it is different from the former in the sense that a weights matrix in WLS is constant across observations, and is different from the latter in the sense that the weights matrix in GWR is based on spatial proximity, rather than numerical similarity. GWR is also different from any spatial autoregressive models because it produces a set of localized estimates. Fotheringham et al. (1998, 1908) contend that "looking at a GWR model estimation gives some insight into how localized effects affect coefficients attached to specific variables".

In the regular OLS regression, regression parameters at ith location are estimated by:

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{8}$$

In the GWR, they are given:

$$\mathbf{b}_i = (\mathbf{X}^T\mathbf{W}_i\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}_i\mathbf{y} \tag{9}$$

where $\mathbf{W}_i$ is an n-by-n local spatial weights matrix, which is a diagonal matrix composed of entries in an ith row in the corresponding global spatial weights matrix. A global spatial weights matrix for GWR is based on inter-distances, and various kernel functions apply to postulate a distance-decay relation. A quartic kernel function is given:

$$w_{ij} = \begin{cases} \{1 - (d_{ij}/h)^2\}^2 & \text{if } d_{ij} < h \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where d_{ij} is a distance between spatial objects, and h is a bandwidth or range beyond which spatial autocorrelation does not exist. In order to determine an h, a cross-validation algorithm can be utilized. However, I fit a variogram for the dependent variable, annual income growth rates between 1969 and 1999 (Figure 6(a)) An exponential function yields a range value around 215 miles.

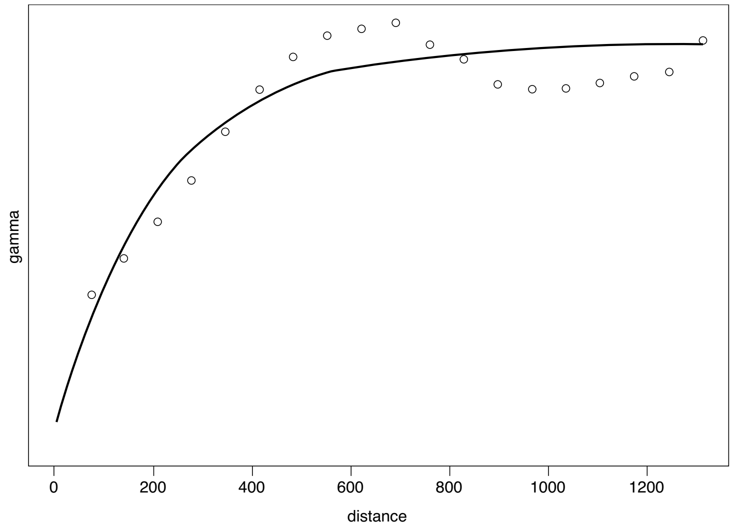
Figure 6(b) shows spatial distribution of β-coeffiencients. Even though ideally the map is expected to be compatible to Figure 4(b), a considerable degree of discrepancy for some areas is observed. This is mainly because they are based on different perspecives in specifying spatial weights matrices; connectivity-based and distance-based. As suggested, a set of spatially adaptive kernel functions, rather than a global kernel function, may perform better to dipict spatial dependence in regression parameters (Brunsden, 1995; Fotheringham et al., 2002).

The first two classes in Figure 6(b) belong to negative β-coefficients which conform to the notion of β-convergence, while the third and fourth classes indicate a positive relationship between 1969 income levels and income growth rates between 1969 and 1999. Note that the global β-coefficient is -0.01 in Table 3. Most areas belong to the first two classes in Figure 4. Focus here is placed on positive values which are against the global trend. High values from California to New Mexico are associated with a combination of high-high association in San Francisco areas and low-low association in the rest of the region (see Figure 4(a)). High values in areas centered on Seattle, the Megalopolis, and southern Florida are related to high-high association in Figure 4.

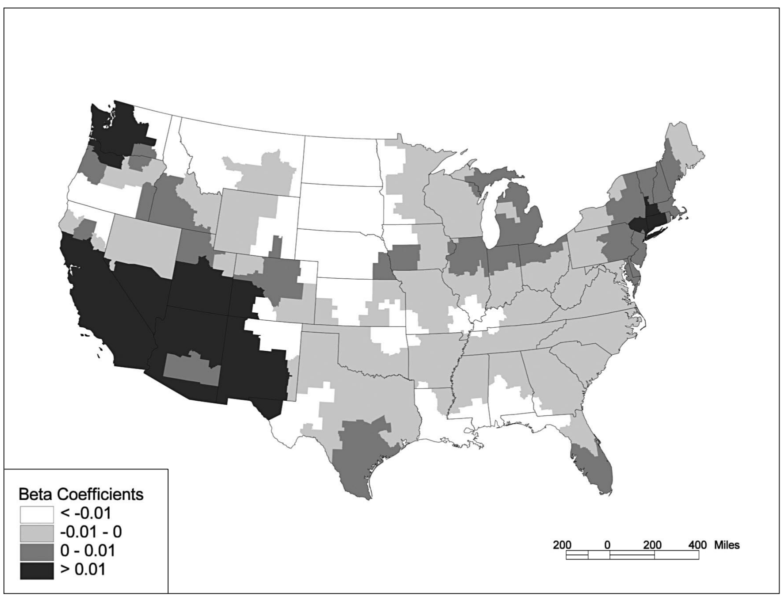
In comparison with Local-r and Local-L scatterplot maps, GWR seems to provide less information. This suggests that GWR may perform better in a multivariate situation, rather than a bivariate situation. Thus, GWR is more suitable for examining conditional convergence that deals with additional shock variables besides an initial income level.

4. Conclusions

In this paper, I analyzed the US annual regional income data from 1969 to 1999 in order to examine β-convergence from a spatial perspective. I fitted a spatial autoregressive model and a GWR model, and utilized bivariate ESDA techniques. The major find-



(a) Variogram



(b) Spatially drifting β-coefficients

Figure 6. A geographically weighted regression (GWR)

ings are as follows.

First, the hypothesis of β-convergence was partially evidenced: the coefficient was 0.01 and significant. However, the trend varies among sub-periods: 2% convergence rate was found in the 1970s, but the coefficient for the 1990s was minimal (0.5%). A SAR model was fitted to deal with spatial autocorrelation in regression residuals. The results indicate that β-coefficient for the entire period is not significant at the 99% confidence level, leading to a conclusion that there is no statistical evidence of regional income convergence in the US over the last three decades.

Second, a local-L scatterplot map and a local-L significance map show that there was a substantive level of spatial heterogeneity in the catch-up process, and suggested possible spatial regimes. Some areas with higher-than-average 1969 income levels, including the northern part of the Megalopolis, southern Florida, and Denver areas, have enjoyed higher-than-average income growth rates, whereas areas with lower-than-average 1969 income levels, LMAs from central Texas to the Four Corners, have experienced lower-than-average income growth rates. A series of local-L significance maps for sub-periods show spatio-temporal heterogeneity in β-convergence: different sub-periods display different facets of spatial restructuring. The catch-up scenario in a spatial sense was least pronounced during the 1980s. A geographically weighted regression (GWR) model also showed significant level of spatial heterogeneity in β-coefficients.

This paper is expected to demonstrate the applicabilities of spatial data analysis techniques, especially ESDA, to a real research topic, here, regional income convergence. The analytical procedures presented in this paper may be benefited if they are coupled with GIS, resulting in an ESDA-GIS framework (Lee, 2001b). Most of all, spatial data analysis needs a generic research platform where data are spatially manipulated, and are effectively explored and visualized. Since major aims of ESDA include spatial pattern detection and the formulation of meaningful hypotheses, ESDA should take more advantage of GIS's capabilities in visualization and spatial data mining (Openshaw, 1990; Goodchild et al., 1992; Fotheringham and Charlton, 1994; Openshaw and Clarke, 1996; Anselin, 1998; Wise et al., 1999). It is obvious that local spatial association measures play a main role in building a feasible ESDA-GIS framework as suggested (Anselin, 1995, 1996, 1999; Anselin and Bao, 1997; Bivand, 1998; Brunsdon, 1998; Dyke, 1998; Unwin and Unwin, 1998; Wilhelm and Steck, 1998; Lee, 2001b, 2004b).

Further, the ESDA-GIS framework orients itself to a broader field of geographical information science (GISc) where various disciplines interact with each other and a new academic division of labor occurs (Goodchild, 1992). I would argue that spatial data analysis in geography need to move into the new terrain with being equipped with various ESDA techniques utilizing local spatial association measures. This implies that analytical geography needs to retreat from technical aspects of GIS and return to the implementation and sophistication of geographical inquiries with substantive research objectives in the GIS environment (Brown, 2000). In this vein, this paper may be regarded as a demonstration.

References

Anselin, L., 1995, Local indicators of spatial association: LISA, *Geographical Analysis*, 27(2), 93-115.

__________, 1996, The Moran scatterplot as an ESDA tools to a assess local instability in spatial association, in Fishcer, M., Scholter, H., and Unwin, D.(eds.), *Spatial Analytical Perspectives on GIS*, Taylor & Francis, London, 111-125.

__________, 1998, Exploratory spatial data analysis in a geocomputational environment, in Longley, P. A., Brooks, S. M., McDonell, R., and MacMillan, B.(eds.), *Geocomputation: A*

Primer, John Wiley & Sons, Chechester, 77-94.

_________, 1999, Interactive techniques and exploratory spatial data analysis, in Longley, P. A., Goodchild, M. F., Maguire, D. J., and Rhind, D. W.(eds.), *Geographical Information Systems, Vol. 1: Principles and Technical Issues*, 2nd Edition, John Wiley & Sons, New York, 253-266.

Anselin, L. and Bao, S., 1997, Exploratory spatial data analysis linking SpaceStat and ArcView, in Fischer, M. and Getis, G.(eds.), *Recent Developments in Spatial Analysis*, Springer-Verlag, Berlin, 35-59.

Anselin, L. and Griffith, D. A., 1988, Do spatial effects really matter in regression analysis?, *Papers of the Regional Science Association*, 65(1), 11-34.

Armstrong, H. W., 1995a, Convergence among regions of the European Union, *Papers in Regional Science*, 74(2), 143-152.

_________________, 1995b, *Trends and Disparities in Regional GDP per capita in the European Union*, United States and Australia, European Commission Report 94/00/74/017, Brussels.

_________________, 1995c, An appraisal of the evidence from cross-sectional analysis of the regional growth process within the European Union, in Armstrong, H. W. and Vickerman, R. W.(eds.), *Convergence and Divergence among European Regions*, Pion, London, 40-65.

Bailey, T. C. and Gatrell, A. C., 1995, *Interactive Spatial Data Analysis*, Longman, New York.

Blanchard, O. J., 1991, Comments and discussion on Barro, R. and Sala-i-Martin, X., Convergence across states and regions, *Brookings Papers on Economic Activity*, 1991(1), 159-174.

Barro, R., 1991, Economic growth in a cross section of countries, *Quarterly Journal of Economics*, 106(2), 407-444.

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

Bernat, G. A., 1996, Does manufacturing matters? A spatial econometric view of Kaldor's laws, *Journal of Regional Science*, 36(3), 463-477.

Bickenbach, F. and Bode, E., 2003, Evaluating the Markov property in studies of economic convergence, *International Regional Science Review*, 26(3), 363-292.

Bivand, R. S., 1998, Software and software design issues in the exploration of local dependence, *Journal of the Royal Statistical Society Series D: The Statistician*, 47(3), 499-508.

Brown, L. A., 2000, The GIS/SA interface for substantive research(ers): a critical need, *Journal of Geographical Systems*, 2(1), 43-47.

Brunsdon, C., 1998, Exploratory spatial data analysis and local indicators of spatial association with XLISP-STAT, *Journal of the Royal Statistical Society Series D: The Statistician*, 47(3), 471-484.

Brunsdon, C., Fotheringham, A. S., and Charlton, M. E., 1996, Geographically weighted regression: A method for exploring spatial nonstationarity, *Geographical Analysis*, 28(4), 281-298.

Brunsdon, C., Fotheringham, A. S., and Charlton, M., 1998, Geographically weighted regression: modelling spatial non-stationarity, *Journal of the Royal Statistical Society Series D: The Statistician*, 47(3), 431-443.

Buettner, T., 1999, The effect of unemployment, aggregate wages, and spatial contiguity on local wages: an investigation with German district level data, *Papers in Regional Science*, 78(1), 47-67.

Casetti, E. and Jones, J. P., III, 1987, Spatial aspects of the productivity slowdown: an analysis of U.S. manufacturing data, *Annals of the Association of American Geographers*, 77(1), 76-88.

Dickey, H., 2001, Regional earnings inequality in Great Britain: a decomposition analysis, *Regional Studies*, 35(7), 605-612.

Dykes, J. A., 1998, Cartographic visualization: exploratory spatial data analysis with local indicators of spatial association using Tcl/Tk and cdv, *Journal of the Royal Statistical Society Series D: The Statistician*, 47(3), 485-497.

European Commission, 1997, *Regional Growth and Convergence*, The Single Market Review, Subseries VI: Aggregate and Regional Impact, Volume 1, European Commission, Luxembourg.

Fan, C. C. and Casetti, E., 1994, The spatial and temporal dynamics of US regional income inequality, 1950-1989, *Annals of Regional Science*, 28(2), 177-196.

Fingleton, B., 1997, Specification and testing of Markov chain models: an application to convergence in the European Union, *Oxford Bulletin of Economics and Statistics*, 59(3), 385-403.

__________, 1999, Estimates of time to economic convergence: an analysis of regions of the European Union, *International Regional Science Review*, 22(1), 5-34.

Fotheringham, A. S, Brunsdon, C., and Charlton, M., 2002, *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*, John Wiley & Sons, Hoboken.

Fotheringham, A. S. and Charlton, M., 1994, GIS and exploratory spatial data analysis, an overview of some research issues, *Geographical Systems*, 1(2), 315-327.

Fotheringham, A. S., Charlton, M. and Brunsdon, C., 1998, Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis, *Environment and Planning A*, 30(11), 1905-1927.

Fotheringham, A. S. and Rogerson, P. A., 1993, GIS and spatial analytical problems, *International Journal of Geographical Information Systems*, 7(1), 3-19.

Goodchild, M. F., 1992, Geographical information science, *International Journal of Geographical Information Systems*, 6(1), 31-45.

Goodchild, M. F., Haining, R., and Wise, S., et al., 1992, Integrating GIS and spatial data analysis: problems and possibilities, *International Journal of Geographical Information Systems*, 6(5), 407-423.

Haining, R., 1990, *Spatial Data Analysis in the Social and Environmental Sciences*, Cambridge University Press, New York.

________, 2003, *Spatial Data Analysis: Theory and Practice*, Cambridge University Press, New York.

Kangasharju, A., 1999, Relative economic performance in Finland: regional convergence, 1934-1993, *Regional Studies*, 33(3), 207-217.

Kennedy, P., 1998, *A Guide to Econometrics*, 4th Edition, MIT Press, Cambridge.

Lee, S.-I., 2001a, Developing a bivariate spatial association measure: an integration of Pearson's r and Moran's I, *Journal of Geographical Systems*, 3(4), 369-385.

______, 2001b, *Spatial Association Measures for an ESDA-GIS Framework: Developments, Significance Tests, and Applications to Spatio-temporal Income Dynamics of U.S. Labor Market Areas*, 1969-1999, Ph.D. Dissertation, The Ohio State University.

______, 2004a, A generalized significance testing method for global measures of spatial association: an extension of the Mantel test, *Environment and Planning A*, 36, in press.

______, 2004b, Exploring the bivariate spatial dependence and heterogeneity: a local bivariate spatial association measure for exploratory spatial data analysis (ESDA), *International Journal of Geographical Information Science*, on revision.

______, 2004c, Exploratory spatial data analysis of σ-convergence in the U.S. regional income distribution, 1969-1999, *Journal of the Korean Urban Geographical Society*, 7(1), in press.

Lopez-Bazo, E., Vaya, E., Mora, A. J., and Surinach, J., 1999, Regional economic dynamics and convergence in the European Union, *Annals of Regional Science*, 33(3), 343-370.

Martin, R., 2001, EMU versus the regions? Regional convergence and divergence in Euroland, *Journal of Economic Geography*, 1(1), 51-80.

Mencken, F. C., 1998, Persistent differences in well-being between Appalachian subregions, *Growth and Change*, 29(4), 469-479.

Molho, I., 1995, Spatial autocorrelation in British unemployment, *Journal of Regional Science*, 35(4), 641-658.

Nijkamp, P. and Poot, J., 1998, Spatial perspectives on new theories of economic growth, *Annals of Regional Science*, 32(1), 7-37.

Openshaw, S., 1990, Spatial analysis and geographical information systems: a review of progress and possibilities, in Scholten, H. and Stillwell, J.(eds.), *Geographical Information Systems for Urban and Regional Planning*, Kluwer, Dordrecht, 153-163.

___________, and Clarke, G., 1996, Developing spatial analysis functions relevant to GIS environments, in Fischer, M., Scholten, H. and Unwin, D.(eds.), *Spatial Analytical Perspectives on GIS*, Taylor & Francis, London, 21-37.

Sala-i-Martin, X., 1996, Regional cohesion: evidence and theories of regional growth and convergence, *European Economic Review*, 40(6), 1325-1352.

Tiefelsdorf, M., 2000, *Modelling Spatial Processes: The Identification and Analysis of Spatial Relationships in Regression Residuals by Means of Moran's I*, Springer, Berlin.

Tolbert, C. M. and Sizer, M., 1996, *U.S. Commuting Zones and Labor Market Areas: A 1990 Update*, Rural Economy Division, Economic Research Service, U.S. Department of Agriculture, Staff Paer 9614, U.S. Government Printing Office, Washington, D.C.

Pons-Novell, J. and Viladecans-Marsal, E., 1999, Kaldor's laws and spatial dependence: evidence for the European regions, *Regional Studies*, 33(5), 443-451.

Quah, D. T., 1993, Galton's fallacy and tests of the convergence hypothesis, *Scandinavian Journal of Economics*, 95(4), 427-423.

__________, 1996a, Regional convergence clusters across Europe, *European Economic Review*, 40(3-5), 951-958.

__________, 1996b, Empirics for economic growth and convergence, *European Economic Review*, 40(6), 1353-1375.

Rey, S. J., 2001, Spatial empirics for economic growth and convergence, *Geographical Analysis*, 33(3), 195-214.

Rey, S. J. and Montouri, B. D., 1999, US regional income convergence: a spatial econometric perspective, *Regional Studies*, 33(2), 143-156.

Schimek, M. G., ed., 2000, *Smoothing and Regression: Approaches, Computation, and Application*, John Wiley & Sons, New York.

Tsionas, E. G., 2000, Regional growth and convergence: evidence from the United States, *Regional Studies*, 34(3), 231-238.

Uniwn, A. and Unwin, D. J., 1998, Exploratory spatial data analysis with local statistics, *Journal of the Royal Statistical Society D: The Statistician*, 47(3), 415-421.

Wilhelm, A. and Steck, R., 1998, Exploring spatial data by using interactive graphics and local statistics, *Journal of the Royal Statistical Society D: The Statistician*, 47(3), 423-430.

Wise, S., Haining, R., and Signoretta, P., 1999, Scientific visualization and the exploratory analysis of area data, *Environment and Planning A*, 31(10), 1825-1838.

Received March 15, 2004
Accepted June 7, 2004