Local versus Global Convergence in Europe: A Bayesian Spatial Econometric Approach

Cem Ertur

Laboratoire d'Economie d'Orléans, LEO - UMR 6221 CNRS, Faculté de Droit d'Economie et de Gestion, Rue de Blois - B.P. 6739, 45067 Orléans Cedex 2, France, e-mail: cem.ertur@univ-orleans.fr

Julie Le Gallo

CRESE, Université de Franche-Comté, 45D, Avenue de l'Observatoire, 25030 Besançon Cedex, France, e-mail: jlegallo@univ-fcomte.fr

James P. LeSage

McCoy Endowed Chair of Urban and Regional Economics, McCoy College of Business Administration, Department of Finance and Economics, Texas State University - San Marcos, San Marcos, TX 78666, e-mail: jlesage@spatial-econometrics.com

Abstract

Numerous studies have pointed to the econometric problems introduced by heterogeneity in cross-sectional data samples used to explore convergence suggested by neo-classical growth models. We introduce a local concept of convergence along with a Bayesian locally linear spatial estimation method to address these problems. The method allows global and local β-convergence to be viewed in a continuous fashion. Inference regarding global convergence can be treated as a mixture distribution arising from local β-convergence estimates from each region in the sample. Taking this approach eliminates the need to specify sub-samples and regimes as well as parameter variation schemes that have been used to model heterogeneity. We illustrate the method using a sample of 138 European regions.

Keywords: Locally linear estimation; Robust; Outliers; Heteroscedastic

JEL classification: C11; R11; R12

1. INTRODUCTION

Since the pioneering contribution of Baumol (1986) and the more formal contributions of Barro and Sala-I-Martin (1991, 1992, 1995) and Mankiw, Romer, and Weil (1992), numerous studies have examined the β-convergence hypothesis based on the neoclassical growth model (Solow 1956) using cross-sectional samples of countries and regions.

The prediction of the neoclassical growth model (Solow 1956) is that the growth rate of an economy is positively related to the distance that separates it from its own steady state. Making the simplistic assumption that economies are structurally similar, characterized by the same steady state and differing only in their initial conditions, we should see *unconditional convergence* to the same steady state. In this case, low-income economies grow faster than those with high incomes and eventually catch up in the long run. Under the more realistic scenario, where economies have different steady states that are conditional on identifiable structural differences, it is possible to draw econometric inferences regarding *conditional convergence*. This requires that we appropriately condition on structural differences that give rise to differences in steady states. In empirical practice, it is difficult to measure and model structural differences; and in theory, heterogenous structures suggest heterogeneity in steady states as well as the structural factors on which we need to condition our econometric models (Durlauf 2000, 2001; Brock and Durlauf 2001).

While the β-convergence hypothesis has been heavily criticized both on theoretical and methodological grounds (Mankiw 1995; Temple 1999; Islam 2003), we focus in this paper on two interrelated issues: parameter heterogeneity and spatial interdependence in conditional convergence models estimated on cross-sectional samples.

First, heterogeneity in the structure of economies suggests that conditioning attempts that rely on smoothly varying variables to describe economic structure might fail to achieve the appropriate conditioning needed to produce valid inferences regarding conditional β-convergence. A theoretical motivation for heterogeneity can be found in endogenous growth theory (Azariadis and Drazen 1990) as well as the neoclassical model with heterogeneous structure (Galor 1996). Econometric methods that attempt to directly accommodate heterogeneity offer an alternative approach to the problem of estimation and inference. Partitioning the cross-sectional sample into regimes based on income levels or other structural characteristics is one approach to modeling heterogeneity, the so-called convergence clubs approach (Desdoigts 1999; Durlauf and Johnson 1995). Allowing for explicit parameter variation over the sample represents another (Durlauf, Kourtellos, and Minkin 2001). In both cases, model specification issues arise beyond those involving which explanatory variables to include in the model. For the case of a multiple regime model, decisions must be made regarding how to partition the cross-sectional sample; for varying parameter models, a specification for this variation must be set forth.

Second, the uneven geographical distribution of economic activities and growth is one of the most striking characteristics of contemporary economies. Indeed, as pointed out by Easterly and Levine (2001), there is a tendency for all factors of production to gather together, leading to a geographic concentration of economic activities. As a consequence, any empirical study on growth and convergence should explicitly acknowledge this phenomenon of spatial interdependence between regions or countries.

In this paper, we develop an empirical methodology that deals with both of these issues through a Bayesian spatial autoregressive locally linear estimation approach. On the one hand, regarding heterogeneity, our locally linear spatial model partitions the cross-sectional sample observations by treating each location along with surrounding locations as a sub-sample. This reduces the need to make arbitrary decisions regarding how to partition the sample observations but allows for variation in the parameter estimates across all observations. On the other hand, our proposed method presumes that similarities in legal and social institutions as well as culture and language might act to create spatially local uniformity in economic structures, leading to similar spatial locality in rates of convergence. We think it useful to define the concept of *local convergence*, which we use to refer to a situation where rates of convergence in economic growth rates are similar for observations located at nearby points in space. In other words, there exists spatial clustering in the magnitudes of the β-convergence parameter estimates. It should be noted that our locally linear spatial estimation method does not impose *a priori* similar rates of convergence for spatially neighboring observations. Rather, we estimate β-convergence parameters for each region/observation in the sample and then examine these estimates in an effort to assess whether there is empirical support for our concept of *local convergence*. This represents an important difference between our approach and a spatially varying parameter estimation scheme that imposes spatial similarity on the estimates. For an example of the latter approach, see LeSage (2004). Furthermore, Bayesian techniques produce robust estimates with regard to potential outliers and heteroscedasticity of unknown form.

The framework that we suggest in this paper therefore allows modeling spatial autocorrelation and heterogeneity of the convergence process. Both heterogeneity in steady states and heterogeneity in the rate of convergence towards this steady state are allowed. Note that the former is usually captured by dynamic panel data models with fixed effects (Islam 1995). However, panel data raises potential problems related to possible small sample bias and short time frequency (Islam 2003). Our approach is then useful when panel data is not appropriate.

The paper is organized as follows. Section 2 describes global versus local spatial autoregressive estimation. Bayesian Markov Chain Monte Carlo (MCMC) estimation of the model is taken up in Section 3 with details provided in an appendix. The model is applied to a sample of 138 European regions in Section 4.

2. ESTIMATION AND INFERENCE REGARDING CONVERGENCE

A spatial autoregressive β-convergence model that can be used to produce global regression estimates in the presence of spatial dependence in a cross-section of observations representing regions or countries is described in Section 2.1. Section 2.2 extends this model to allow for a sequence of locally linear parameter estimates associated with each observation (country or region) in the data sample.

2.1 Global Spatial Autoregressive Estimates

Formally, β-convergence models rely on a cross-section of countries or regions, using the average growth rate of per capita GDP (y) over a given time period as the dependent variable. These models rely on an explanatory variables matrix, $X = [\iota \ y_0]$ consisting of a constant (ι) as well as the initial level of log per capita GDP (y_0) and the associated parameter vector $\gamma = [\alpha \ \beta]'$ as shown in (1):

(1)
$$y = X\gamma + \varepsilon$$
$$\varepsilon \sim N\left(0, \ \sigma_\varepsilon^2 \ I_n\right).$$

Most often, least-squares estimation is used to determine the sign and significance of the parameter β for the case of unconditional β-convergence. For conditional β-convergence, a matrix of explanatory variables that purport to measure and control for structural differences is introduced in (1). Typical variables suggested by Mankiw, Romer, and Weil (1992) in an augmented Solow growth model were: human and physical capital, saving rates, and population growth rates. Additional variables to control for structural differences might include: the ratio of public consumption to GDP, the ratio of domestic investment to GDP, terms of trade, the fertility rate, etc. (see Barro and Sala-I-Martin 1995). In fact, more than 90 of such variables have been included in cross-country regressions using international data sets in the empirical growth literature as surveyed by Durlauf and Quah (1999).

The problems linked to the omission of the spatial dimension of cross-sectional data have recently been highlighted. Indeed, in most cross-country studies, economies are treated as "isolated islands" (Mankiw 1995; Quah 1996), whereas they are often characterized by spatial autocorrelation. Spatial autocorrelation refers to the coincidence of attribute similarity and locational similarity (Anselin 1988). In the context of European regions, for example, positive spatial autocorrelation indicates that wealthier regions tend to be geographically clustered as well as poorer regions. It may arise from the fact that the data are affected by processes touching different locations. Theoretical models from economic geography point to factors such as technology diffusion, factor mobility, and trade, which all have a strong geographic dimension that might interact with growth processes (Kubo 1995; Martin and Ottaviano 1999). Spatial autocorrelation can also arise from model misspecifications (omitted variables, measurement errors) or from a variety of measurement problems, as boundary mismatching between the administrative

boundaries used to organize the data and the actual boundaries of the economic processes believed to generate convergence. In the regional science literature, numerous studies have explicitly taken into account spatial autocorrelation in convergence analysis (see Abreu, de Groot, and Florax 2005 and Rey and Janikas 2005 for reviews).

To accommodate spatial dependence in the growth rates of regions or countries reflected in the cross-sectional dependent variable *y*, estimates might be produced using the spatial autoregressive model (SAR) shown in (2). This model includes what is known as a *spatial lag* of the dependent variable (see Anselin 1988):

(2) $y = \rho W y + X \gamma + \varepsilon.$

This model conventionally assumes that $\varepsilon \sim N\left(0, \sigma^2 I_n\right)$, but we will have more to say about this later. The vector *y*, matrix *X*, and parameter vector γ are as described for the model in (1).

The $n \times n$ matrix *W* is a row-standardized spatial weight matrix. While a number of ways exist to specify *W*, a common specification sets $W_{ij} > 0$ for observations $j = 1 \ldots n$ sufficiently close (as measured by some metric) to observation *i*. For example, we might rely on observations that are spatially contiguous to observation *i*, those that have borders in common, or we might use the five nearest neighbors measured by distance from the centroids of each location. By construction, the main diagonal of *W* is set to zero to preclude an observation from directly predicting itself. Row-standardization of the matrix *W* scales each element in the matrix so that the rows sum to unity, producing an explanatory variable *Wy* that reflects the average of growth rates from neighboring observations. The scalar parameter ρ measures the influence of the variable *Wy* on *y*.

This particular functional form can be motivated theoretically: Ertur and Koch (2005); Lopez-Bazo, Vayá, and Artís (2004); and Vayá et al. (2004) have recently derived neoclassical models with spatial externalities and technological interdependence yielding to convergence models including spatial autocorrelation. The spatial lag structure plays the role of a lagged dependent variable in time-series models, accounting for variation in the dependent variable arising from latent or unobservable variables. In the case of our spatial lag, these latent factors are correlated among cross-sectional observations located nearby in geographic space. Indeed, many empirical studies have found evidence of spatial autocorrelation in the residuals of traditional models (Moreno and Trehan 1997; Fingleton 1999; Conley and Ligon 2002; Le Gallo, Ertur, and Baumont 2003). The spatial lag highlights a spatial spillover effect, where the growth rate in each region is affected by those of neighboring regions after conditioning on initial per capita GDP levels.

This model can be estimated using maximum likelihood methods (see Anselin 1988) assuming that there is a homogeneous relationship between *y* and *X* across the spatial sample of observations. The estimated scalar parameter $\hat{\rho}$ could be used to test for the

presence of significant spatial dependence in the sample of cross-sectional growth rates. If this parameter is not significantly different from zero, the model in (2) collapses to the simple least-squares model in (1). The scalar parameter estimate $\hat{\beta}$ contained in the parameter vector γ is used to produce an inference regarding convergence that we label *global convergence*. Inferences based on this parameter represent a conclusion regarding convergence or non-convergence that averages over sample data evidence from the entire sample of countries or regions.

As an illustration, we provide estimates for the model in (1) based on least-squares alongside maximum likelihood estimates for the SAR model in (2) in Table 1. These estimates were based on a sample of growth rates in log per capita GDP for 138 European regions over the period from 1980 to 1995 (see Section 1 for a detailed description of the sample data). The SAR model used a spatial weight matrix W based on the ten nearest neighbors to each region in the sample. Results based on spatial weight matrices formed using eight to 12 nearest neighbors to each region were similar to those reported in the table. From the homoscedastic model estimates reported in the table, we see strong evidence of spatial dependence as indicated by the estimate of $\rho = 0.75$, that is significant at the 99 percent confidence level. The table also illustrates a difference between the magnitude of the least-squares β and that from the SAR model, pointing to differing rates of convergence. The least-squares estimate suggests more rapid convergence than that from the spatial model.

Another issue that plagues growth regressions is non-constant variance across the sample of countries or regions. Table 1 also presents estimates for the least-squares and SAR model based on a Bayesian heteroscedastic linear model proposed by Geweke (1993) and a spatial autoregressive variant of this model suggested by LeSage (1997). These models allow the disturbances to take the form $\varepsilon \sim N\left(0, \sigma^2 V\right)$, where V is a diagonal matrix containing variance scalars $v_1, v_2, \ldots, v_n$, estimated using Markov Chain Monte Carlo (MCMC) methods.

Specifics regarding the prior assigned to the v_i terms are given in the appendix, but we note here that the role of the v_i terms is to accommodate outliers or observations containing large variances by down-weighting these observations. Note that in cross-country literature, the presence of outliers, affecting the estimation of β-convergence models, have been pointed out by DeLong and Summers (1991) and Temple (1998, 1999). In the context of spatial modeling, outliers or aberrant observations arise due to "enclave effects," where a particular region exhibits divergent behavior from nearby areas. Geweke (1993) shows that this approach to modeling the disturbances is equivalent to a model that assumes a Student–*t* distribution for the errors. We note that this type of distribution has frequently been used to deal with sample data containing outliers, (e.g., Lange, Little, and Taylor 1989). The heteroscedastic estimates reported in Table 1 are based Markov Chain Monte Carlo (MCMC) estimation described in the appendix.

TABLE 1

Least-Squares Versus Spatial Model Estimates

		Homoscedastic Model Estimates		
		OLS *t*-Marginal		SAR *t*-Marginal
Variable	OLS	Probability	SAR	Probability
Constant	0.1294	0.0000	0.0580	0.0009
1980 log-level	-0.0078	0.0000	-0.0048	0.0033
Spatial lag, *Wy*			0.7502	0.0000
		Bayesian Heteroscedastic Model Estimates		
		95% HPD		95% HPD
Variable	OLS	Upper-Lower	SAR	Upper-Lower
Constant	0.1225	0.0933 — 0.1513	0.0474	0.0256 — 0.0726
1980 log-level	-0.0065	-0.1042 — -0.0038	-0.0390	-0.0637 — -0.0146
Spatial lag, *Wy*			0.7676	0.6230 — 0.8750

These robust estimates suggest lower values for the convergence parameter β in both the least-squares and SAR models. It should be noted that a 95 percent highest posterior density (HPD) interval is reported for the Bayesian estimates. These point to a β estimate for the heteroscedastic SAR model that is significantly different from zero based on the 95 percent HPD intervals.

To summarize this discussion, inferences regarding convergence based on what we choose to label global estimates that presume homogeneity in the relationship across the sample of regions or countries are likely to be sensitive to outliers and to influences such as spatial dependence that have the potential to bias least-squares estimates. For this reason, we propose a locally linear spatial autoregressive model described in the next section. This model is capable of producing inferences regarding our concept of *local convergence*.

2.2 Locally Linear Spatial Autoregressive Estimates

McMillen (1996) and McMillen and McDonald (1997) introduced a form of spatial non-parametric locally linear weighted regression (LWR) that Brunsdon, Fotheringham, and Charlton (1996) term geographically weighted regressions (GWR). This approach to modeling spatial dependence relies on separate models estimated using a sub-sample of the data based on observations nearby each observation. The motivation for this approach is that if spatial dependence arises due to inadequately modeled spatial heterogeneity, LWR can potentially eliminate this problem. These models often rely on the estimated parameters to detect systematic variation in the relationship being examined over space. Pace and LeSage (2004) point out that LWR methods exhibit a trade-off between increasing the sample size to produce less volatile estimates that contain increasing spatial dependence. Selecting a smaller sample size reduces the spatial dependence but at the cost of increased parameter variability that impedes detection of systematic patterns of parameter variation over space. Therefore, they establish the spatial autoregressive

local estimation (SALE) method and argue that the SALE method eliminates this problem by extending the LWR approach to include a spatial lag of the dependent variable, which accommodates spatial autocorrelation likely to arise as the sub-sample size is increased. In addition, inclusion of the spatial autoregressive term in the model results in improved prediction and stability of the parameter estimates, decreasing the sensitivity of performance to the bandwidth that is typically observed.

Formally, to accommodate both spatial dependence and heterogeneity, we produce estimates using n–models, where n represents the number of cross-sectional sample observations, using the locally linear spatial autoregressive model in (3):

$$(3) \qquad U(i)y = \rho_i U(i)Wy + U(i)X\gamma_i + U(i)\varepsilon$$

where $U(i)$ represent an $n \times n$ diagonal matrix containing distance-based weights for observation i that assign weights of one to the m nearest neighbors to observation i and weights of zero to all other observations. This results in the product $U(i)y$ representing an $m \times 1$ sub-sample of observed GDP growth rates associated with the m observations nearest in location (using Euclidean distance) to observation i.

Similarly, the product $U(i)X$ extracts a sub-sample of explanatory variable information based on m nearest neighbors. We note that as $m \to n$, $U(i) \to I_n$, so that expanding the sub-sample size m around each locality results in a limiting model where the sub-sample size expands to include all observations in the cross-sectional sample. In other words, these estimates approach the global estimates based on all n observations that would arise from the SAR model in (2). This produces locally linear econometric estimates that vary systematically as the sub-sample size increases towards the global estimates one would achieve using the entire sample. It allows a systematic assessment of the mapping between the locally linear estimates that accommodate heterogeneity in steady states and convergence speeds and estimates based on the global sample reflecting homogeneity (absolute β-convergence model). This allows us to assess empirical evidence in support of local convergence in light of the more traditional global convergence approach.

The SALE model assumes $\varepsilon_i \sim N\left(0, \sigma_i^2 U(i)I_n\right)$, but we will have more to say about this later. On the other hand, the scalar parameter ρ_i measures the influence of the variable $U(i)Wy$ on $U(i)y$. Note that there is a cost associated with introducing the spatial lag since the SALE model requires maximum likelihood methods, whereas the LWR model relies on least-squares. However, Pace and LeSage (2004) present an efficient recursive approach for maximum likelihood estimation of the n spatial autoregressive models for problems involving large numbers of observations and illustrate the method for a sample of 3,107 U.S. counties. Most cross-sectional samples of countries or regions used in the empirical convergence literature involve considerably smaller samples.

We extend the SALE model to accommodate non-constant variances by introducing $\varepsilon_i \sim N\left(0,\ \sigma_i^2\ U(i)V\right)$, $V = \text{diag }(v_1,\ v_2,\ \ldots,\ v_n)$. We label this model BSALE, Bayesian spatial autoregressive local estimation. The specifics of this extension are described in Section 3 with MCMC estimation details provided in the appendix.

The smaller samples used for estimation give rise to another problem with local estimation methods, pointed out by LeSage (2004). Indeed, aberrant observations or outliers arising from spatial enclave effects or shifts in regime can exert a large impact on the locally linear estimates. Since these sub-sample estimates may be based on a small number of observations and the sample data observations are re-used when estimates are produced for each point in space, a single outlier can contaminate estimates covering large areas or sub-regions of the spatial sample. This may create an artifact that resembles a regime shift or spatial clustering pattern in the estimates for β (or in β as well as the parameters on control variables in the conditional β-convergence model). Intuitively, a single outlier will reappear in sub-samples constructed using neighboring locations needed to produce estimates for each point in the spatial sample. This allows a single outlier to produce a contagion effect that can impact estimates for an entire region of the sample.

In the next section, we set forth the BSALE model that can accommodate outliers by down-weighting these observations.

3. BAYESIAN SPATIAL AUTOREGRESSIVE LOCAL ESTIMATION

For each spatial autoregressive model based on a sub-sample of size *m*, we specify our model as shown in (4), where the $n \times n$ diagonal matrix $U(i)$ assigns a weight of unity to the *m* nearest neighbors to observation *i*, and zero weight to all other observations.

$$U(i)y = \rho U(i)Wy + U(i)X\gamma + U(i)\varepsilon$$

(4)

$$\text{U}(i)\varepsilon \sim N\left(0, \sigma^2 U(i)V\right), V = \text{diag }(v_1,\ v_2,\ldots,v_n)$$

The $m \times m$ matrix *W* represents a spatial weight matrix with row-sums normalized to unity. This locally linear Bayesian variant of the basic spatial autoregressive model shown in (4) introduces a set of variance scalars $(v_1,\ v_2,\ \ldots,\ v_n)$, that represent unknown parameters that need to be estimated. This allows us to assume $\varepsilon \sim N\left(0,\ \sigma^2\ U(i)V\right)$, where $V = \text{diag}(v_1,\ v_2,\ \ldots,\ v_n)$, but we note that only *m* of the variance scalars v_i take on non-zero values. As noted, this approach to robust modeling in the face of non-constant variance or outliers was introduced by Geweke (1993) for a least-squares model and LeSage (1997) for the spatial autoregressive model. Details concerning MCMC estimation of this model can be found in the appendix. This Bayesian model relies on a diffuse prior for the parameters α and β, a relatively uninformative Gamma prior for the noise variance, and a uniform prior for ρ over the interval -1 to 1.

The other aspect of our Bayesian SALE model is selection or setting of the sub-sample size *m*. As already noted, variation in this will create a host of parameter outcomes that are: highly volatile over the spatial sample for small values of *m* and nearly constant taking on values near the global estimates as *m* → *n*. This issue typically arises with locally linear non-parametric estimation methods, and cross-validation methods are often used to select an optimal sub-sample size. A plausible range for sub-sample size consideration might be (¼)*n* < *m* < (¾)*n*, so that sub-sample sizes are at least ¼ the number of observations but less than ¾ of the entire sample. Of course, these ranges could be changed depending on the size of the sample data. A related problem is that inference regarding the parameters is conditional on the sub-sample size selected.

One advantage of the SALE method is that a mapping of the parameter estimates is provided that allows an examination of the sensitivity of inferences with regard to choice of sub-sample size. We can examine the sequence of estimates for sub-sample sizes ranging from *m* = (¼)*n* to *m* = (¾)*m* in an effort to see whether inferences would differ as the sub-sample size varies. This is the approach we take here. A cross-validation approach in this setting might involve use of the estimates for observation *i* based on a sub-sample size *m* to predict "fringe observations," those that border the sub-sample of *m* observations. This would represent a spatial analogue to one-step-ahead predictions in time-series. A Bayesian solution to the problem of sub-sample size selection would be to mix over estimates based on alternative sub-sample sizes to produce posterior estimates that reflect uncertainty with regard to the choice of sub-sample size. Unfortunately, this requires determination of weights that would be used in mixing over the estimates from alternative sub-sample sizes. These weights should be based on posterior probabilities associated with models arising from the various sub-sample sizes; but this would require integration over sub-sample sizes, which would be treated as a parameter in the model. This would lead to computationally expensive calculations. We demonstrate that inference regarding convergence versus non-convergence is not sensitive to sub-sample sizes ranging from 40 to 100 observations, which roughly corresponds to (¼)*n* and(¾)*n*.

The parameters γ, *V*, and σ and the sub-sample size *m* in the heteroscedastic SAR model can be estimated by drawing sequentially from the conditional distributions of these parameters, a process known as "alternating conditional sampling," or Markov Chain Monte Carlo (MCMC) sampling. To illustrate how this works, let $\theta = (\theta_1, \theta_2)$ represent a parameter vector and $p(\theta)$ denote the prior, with $L\left(\theta \mid y, X, W\right)$ denoting the likelihood. This results in a posterior distribution $p\left(\theta \mid D\right) = c \cdot p(\theta)L\left(\theta \mid y, X, W\right)$, with *c* a normalizing constant. Consider the case where $p\left(\theta \mid D\right)$ is difficult to work with, but a partition of the parameters into two sets θ_1, θ_2 is easier to handle. Given an initial estimate for θ_1, which we label $\hat{\theta}_1$, suppose we could easily estimate θ_2 conditional on θ_1 using $p\left(\theta_2 \mid D, \hat{\theta}_1\right)$. Denote the estimate $\hat{\theta}_2$, derived by using the posterior mean or mode of $p\left(\theta_2 \mid D, \hat{\theta}_1\right)$. Assume further that we are now able to easily construct a new

estimate of θ_1 based on the conditional distribution $p\left(\theta_1 \mid D, \hat{\theta}_2\right)$. This new estimate for θ_1 can be used to construct another value for θ_2, and so on. On each pass through the sequence of sampling from the two conditional distributions for θ_1, θ_2, we collect the parameter draws that are used to construct a joint posterior distribution for the parameters in our model. Gelfand and Smith (1990) demonstrate that sampling from the sequence of complete conditional distributions for all parameters in the model produces a set of estimates that converge in the limit to the true (joint) posterior distribution of the parameters. That is, despite the use of conditional distributions in our sampling scheme, a large sample of the draws can be used to produce valid posterior inferences regarding the joint posterior mean and moments of the parameters.

To implement this estimation method, we need to determine the conditional distributions for each parameter in our BSALE model. These are developed in the appendix that also describes the MCMC sampling scheme.

4. CONVERGENCE OF EUROPEAN REGIONS

We illustrate the BSALE method using a sample of 138 European regions and data covering the period 1980 to 1995. These local estimation results and inferences regarding convergence are compared to the global estimates and inferences presented in Section 4.1.

4.1 The Sample Data

Data limitations remain a serious problem in the European regional context. Harmonized and reliable data allowing consistent regional comparisons are scarce, in particular for the beginning of the time period under study. There is clearly a lack of appropriate or easily accessible data that could be used to measure and control for structural differences considered by conditional β-convergence models. This represents a departure from the cross-country studies of Barro and Sala-I-Martin (1995) or Mankiw, Romer, and Weil (1992), which rely on an extensive international data set.

We use the log of European regional per capita GDP over the period 1980-1995 expressed in ECUs, the former European Currency Unit, replaced by the Euro in 1999. The data are extracted from the EUROSTAT-REGIO database, which is widely used in empirical studies of European regions. [See for example López-Bazo et al. (1999), Neven and Gouyette (1995), and Quah (1996) among others.] Our sample includes 138 regions in 11 European countries over the 1980-1995 period: Belgium (BE:11), Denmark (DK:1), France (FR:21), Germany (DE:30), Greece (GR:13), Luxembourg (LU:1), Italy (IT:20), the Netherlands (NL:9), Portugal (PT:5), and Spain (ES:16) in NUTS2 level and the United Kingdom (UK:11) in NUTS1 level. (See the data appendix for more details.) NUTS is the French acronym for Nomenclature of Territorial Units for Statistics used by Eurostat. In this nomenclature, NUTS1 refers to European Community Regions and NUTS2 to Basic Administrative Units. NUTS1 is used for the United Kingdom because

there is no official counterpart to NUTS2 units, which are drawn up only for the European Commission use as groups of counties. This explains data non-availability at the NUTS2 level throughout the period for this country. Luxembourg and Denmark may be considered NUTS2 regions according to Eurostat. Our choice to prefer NUTS2 level is mainly driven by policy considerations. Since reform in 1989, NUTS2 is the level at which eligibility for Objective 1 Structural Funds is determined. (See *The European Regions: Sixth Periodic Report on the Socio-Economic Situation in the Regions of the European Union,* European Commission, 1999.)

It is worth mentioning that our sample is far more consistent and encompasses more regions than the one initially used by Barro and Sala-I-Martin (1991, 73 regions; 1995, 90 regions) and Sala-I-Martin (1996a, 73 regions; 1996b, 90 regions) where different sources and different regional breakdowns were mixed. Moreover, the smaller 73 region data set is largely confined to prosperous European regions belonging to Western Germany, France, United Kingdom, Belgium, Denmark, Netherlands, and Italy, excluding Spanish, Portuguese, and Greek regions, which are less prosperous. This may result in a selection bias problem raised by DeLong (1988). Armstrong (1995) attempted to overcome these problems by expanding the original Barro and Sala-I-Martin (1991) 73-region data set to less prosperous southern regions using a more consistent sample of 85 regions. The time period 1980-1995 for our sample results from a need to control for monetary changes that do not allow for consistent measures of income across countries in more recent periods.

4.2 Estimation Results

Throughout the empirical application, the weight matrix used is constructed using the six nearest neighbors to each region in the sample. Estimation results based on a first-order contiguity weighting matrix were also examined. The number of neighbors ranged from a low of just three first-order contiguous neighbors up to 10 contiguous neighbors with an average around six neighbors. Estimates from the model based on a first-order contiguity weighting matrix were nearly identical to those reported here based on the six nearest neighbors. We note that using 10 nearest neighbors in the formulation of *W* places a constraint on the smallest local sample size that can plausibly be used during estimation. It seems advisable to assign non-zero weights using the matrix $U(i)$ in (4) for at least 20 or 30 observations to provide an adequate amount of sample data on which to base estimates of ρ, β, *V*, and σ. This in part motivated our choice of six nearest neighbors and the restriction to 20 observations as the smallest sample size we consider.

Previous investigation using similar datasets of European regions have shown that the spatial distribution of per capita GDP is indeed characterized by spatial autocorrelation and that the convergence process of regions is affected by spatial spillovers (Le Gallo and Ertur 2003; Ertur, Le Gallo, and Baumont 2006; Le Gallo and Dall'erba 2006). The first point we illustrate using our estimation results therefore regards the statistical significance of the spatial dependence parameter ρ. Locally linear non-parametric models attempt to eliminate this dependence by relying on small sample sizes where spatial

dependence would be small or non-existent. We present kernel density estimates of the distribution of 138 estimates for ρ from the SALE model based on sub-sample sizes of *m* = 20, 30, 40 in Figure 1. Even in the case of the small sub-sample size of *m* = 20 shown in Figure 1, we see a multi-modal distribution of the 138 estimates for ρ, suggesting a great deal of variation in spatial dependence across the sample of European regions. The mean of these estimates is -0.07, near zero, lending support to the notion that locally-linear methods based on small sub-samples can overcome spatial dependence. However, there are a number of regions where the spatial dependence estimate appears to take on large (positive or negative) values, indicating the presence of spatial dependence between elements of the *y* vector. This would have an adverse impact on the estimates of β for a number of regions in the 138 region sample. The impact of non-zero ρ values in the spatial autoregressive model is similar to that arising from simultaneity, resulting in biased and inconsistent estimates of β (see Anselin 1988).
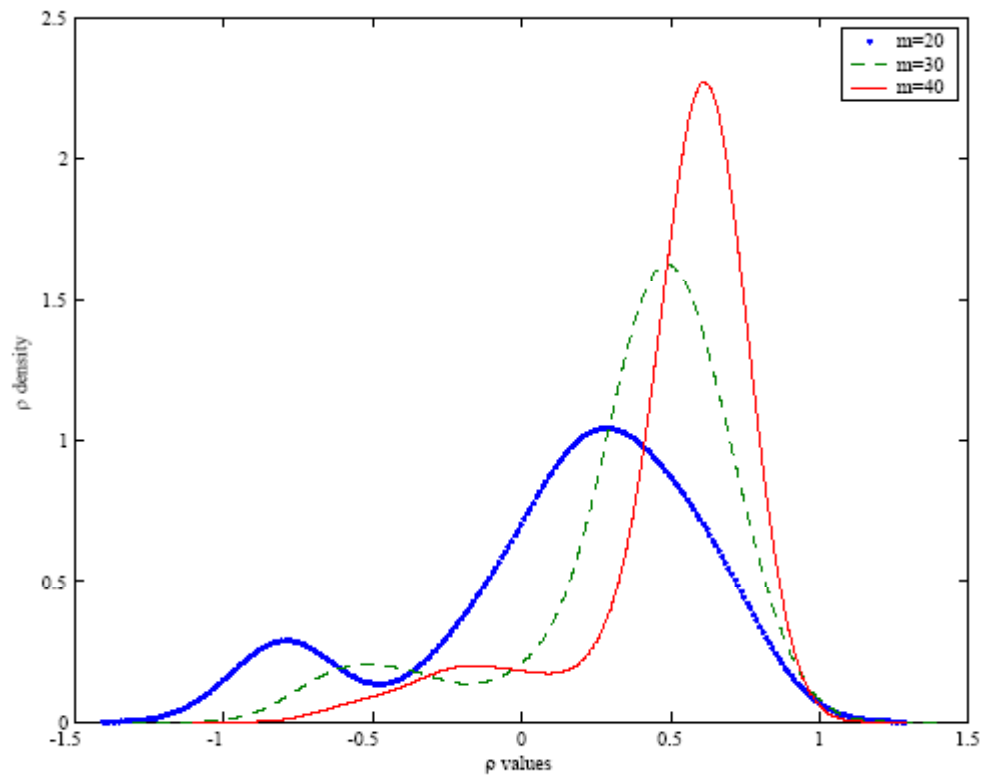

FIGURE 1. Distribution of ρ Estimates for *m* = 20, 30, 40

The distribution of ρ estimates for sub-sample sizes of *m* = 30 and *m* = 40 shown in Figure 1 more clearly point to larger positive modal values, suggesting that spatial dependence increases as the sub-sample size increases, as we would expect. In these cases, the majority of estimates for β would be subject to the biasing impact of spatial dependence among the *y* values. The mean of 138 estimates for ρ based on sub-sample sizes 30 to 80 ranged from 0.35 for the small sub-sample size of 30 up to 0.71 for the large sub-sample size of 80.

There is also variation in the amount of spatial dependence as we move across countries, shown in Figure 2 where individual estimates for ρ are displayed. Observations associated with countries are delimited by vertical lines in the figure, and estimates based on a sub-sample size of 40 and 80 are shown. It should be clear that spatial dependence of a sufficiently large magnitude to create bias in least-squares estimates arises even for the relatively small sub-sample size of 40.
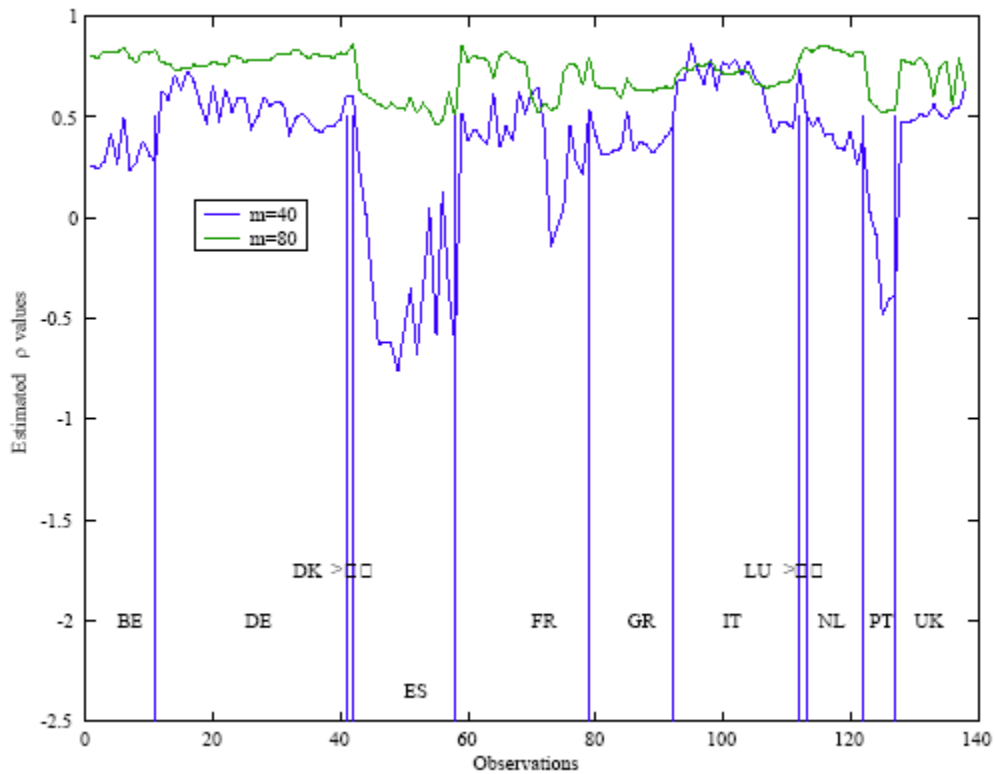

FIGURE 2. Estimates for ρ Based on *m* = 40, 80

These results suggest that inclusion of the spatial lag of the dependent variable serves two useful purposes. First, it acts as a parsimonious proxy for unobserved latent spatial influences that are typically modeled by adding numerous explanatory variables. Second, it allows increasing the sub-sample size used to produce locally linear estimates, which can stabilize the estimates and allow identification of spatial patterns or regimes. This can be done without introducing bias in the estimates for β that typically arises when larger sub-samples are used in local spatial estimation methods.

Estimates for the convergence parameter β are shown in Figure 3, where again observations associated with countries are delimited by vertical lines in the figure. A set of three estimates based on sub-sample sizes of 60, 70, and 80 are presented. Country-level differences are apparent in the figure, where we see estimates change abruptly as we move from one country to another. In addition to distinct variation in the convergence parameter between countries, there is also substantial variation between regions within a country in some cases. Fingleton and McCombie (1998) find similar evidence of heterogeneity across countries when examining European Union growth.
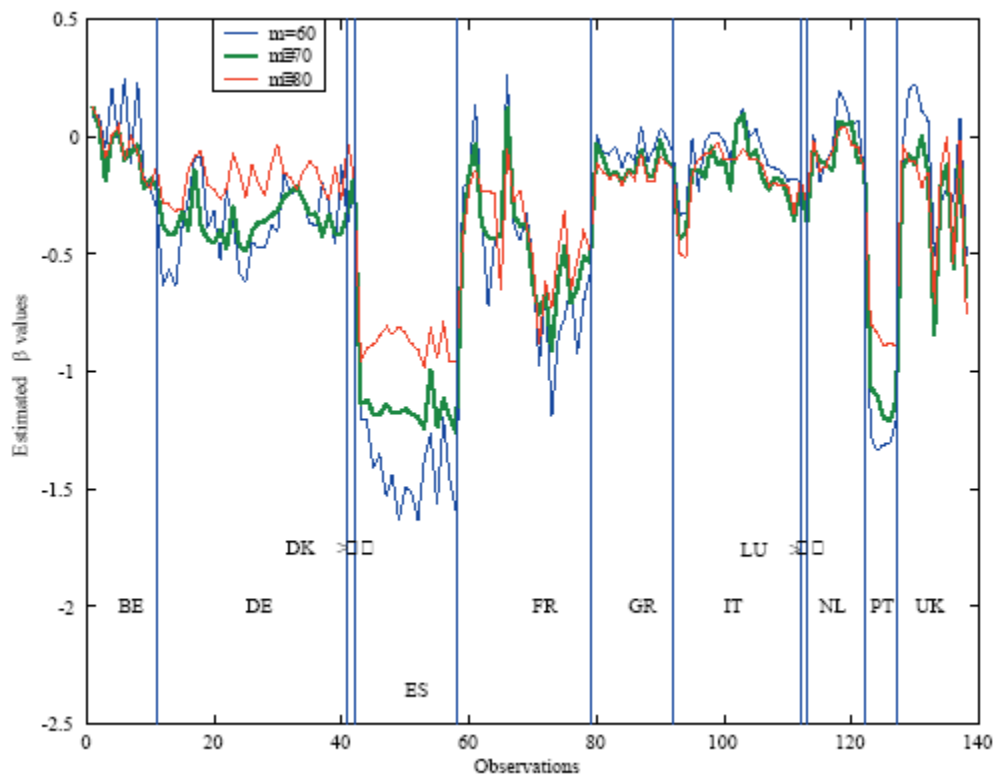

FIGURE 3. Estimates for β Based on *m* = 6, 70, 80

Country-level differences might suggest use of dummy variables, but the approach taken here allows for additional observations in cases where a country consist of only a handful of regions. This is because nearby regions from neighboring countries are included in the sub-sample used to estimate the parameters. In addition, the introduction of dummy variables would only influence the intercept term, and our focus is on the convergence parameter β.

Samples of draws generated during MCMC sampling can be used to produce estimates for the standard deviations of the parameter β and associated confidence intervals. It should be noted that the estimates suffer from sample re-use as in the case of other locally linear non-parametric estimation methods. Sample observations from neighbors are re-used to produce estimates for each location; and in the case of neighboring observations, the amount of sample overlap would be substantial. This inhibits our ability to interpret these measures of dispersion in estimate outcomes in a strict statistical sense. Nonetheless, we provide a graphical depiction of the β estimates based on a sub-sample size of 80 observations along with two standard deviation intervals in Figure 4. We simply note that convergence indicated by negative and significant values of β is likely for the EU regions in Spain and Portugal as well as some regions in France. For observations associated with these regions, the estimates for the convergence parameter β are negative; and the upper confidence interval lies below zero, suggestive of significant negative values for this parameter.
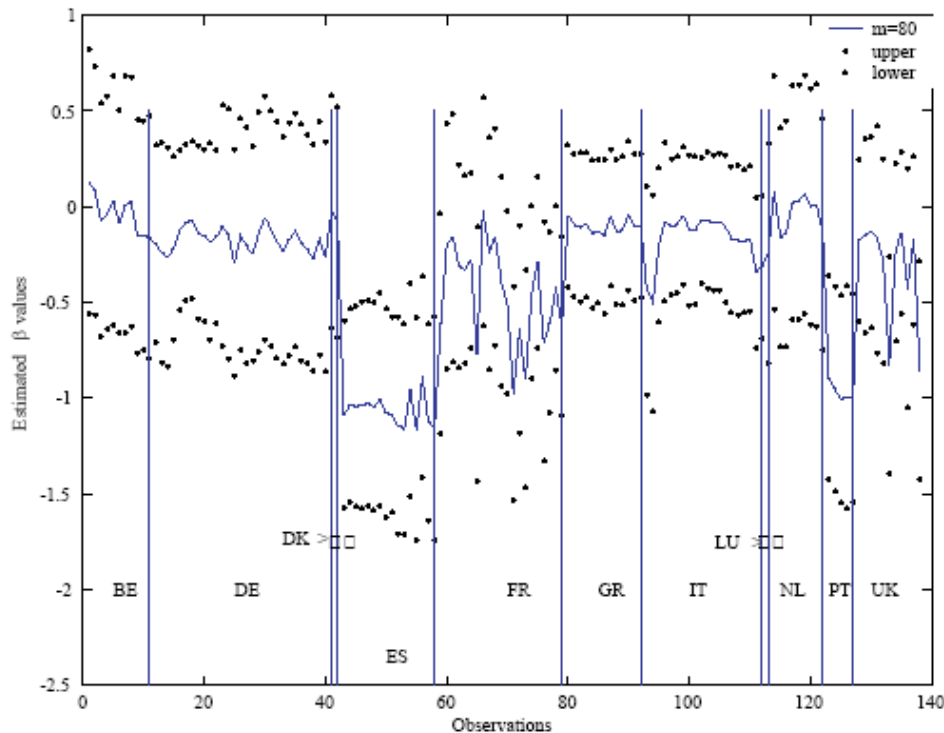

FIGURE 4. Upper and Lower Confidence Intervals for β Based on *m* = 80

Individual β and ρ estimates for the 138 regions based on a sample size of 80 are presented in Table 2 along with standard deviations constructed using the MCMC draws. Regions where the estimate for β is negative and more than two standard deviations away from zero are flagged in the table with the symbol *. These parameter estimates would be consistent with convergence. For the case of the parameter ρ, all values were more than two standard deviations away from zero, so no symbols were added to the table. There are no cases where the positive coefficient values for β are more than two standard deviations away from zero, indicating divergence of the region from surrounding regions.

It is interesting to note that in Table 2, only 31 of the 138 locally linear spatial autoregressive estimates for β are negative and significant (more than two standard deviations from zero), consistent with an inference of convergence. These regions tend to be spatially clustered in Spain, Portugal, and southern France as shown in Figure 5. Use of global least-squares and SAR model estimates such as those presented in Table 1 of Section 1 do not allow for this type of distinction. The β parameter estimates based on the four global models would lead to an inference of global convergence at the 95 percent level or above in all four cases presented in Table 1. The concept of local convergence in conjunction with the BSALE model proposed here provide a great deal of additional information regarding the nature of convergence in growth rates across a spatial sample of observations. The BSALE estimates suggest that convergence is taking place for some regions in our sample but not in others.

TABLE 2

Estimates for β and ρ

OBS	NUTS	β	σ_β	ρ	σ_ρ
1	be1	0.1265	0.3442	0.7673	0.0844
2	be21	0.0703	0.3283	0.7586	0.0818
3	be22	-0.0872	0.3123	0.7738	0.0789
4	be23	0.0039	0.3123	0.7603	0.0790
5	be24	0.0535	0.3251	0.7639	0.0811
6	be25	-0.0982	0.3067	0.7834	0.0762
7	be31	0.0070	0.3463	0.7605	0.0825
8	be32	-0.0606	0.3128	0.7704	0.0820
9	be33	-0.1795	0.3113	0.7614	0.0771
10	be34	-0.2197	0.3159	0.7881	0.0807
11	be35	-0.1197	0.3138	0.7745	0.0808
12	de11	-0.2825	0.2580	0.7235	0.0868
13	de12	-0.2864	0.2880	0.7233	0.0876
14	de13	-0.3147	0.2854	0.7211	0.0888
15	de14	-0.3066	0.2359	0.7171	0.0884
16	de21	-0.1317	0.2162	0.6610	0.1019
17	de22	-0.0809	0.2101	0.6621	0.0983
18	de23	-0.0605	0.2115	0.6452	0.0965
19	de24	-0.2011	0.2269	0.6834	0.0916
20	de25	-0.2221	0.2264	0.7110	0.0909
21	de26	-0.2637	0.2644	0.7006	0.0944
22	de27	-0.2195	0.2256	0.6976	0.0919

OBS	NUTS	β	σ_β	ρ	σ_ρ
23	de5	-0.0696	0.3149	0.7506	0.0849
24	de6	-0.1511	0.3350	0.7412	0.0840
25	de71	-0.2574	0.3051	0.7147	0.0922
26	de72	-0.1204	0.3085	0.7033	0.0884
27	de73	-0.1966	0.3019	0.7137	0.0932
28	de91	-0.2483	0.2851	0.7458	0.0904
29	de92	-0.1110	0.3104	0.7221	0.0896
30	de93	-0.0381	0.3118	0.7522	0.0875
31	de94	-0.1534	0.3047	0.7928	0.0789
32	dea1	-0.1742	0.3297	0.7470	0.0806
33	dea2	-0.2194	0.3139	0.7528	0.0837
34	dea3	-0.1463	0.3112	0.7405	0.0795
35	dea4	-0.1064	0.3103	0.7356	0.0895
36	dea5	-0.1339	0.3171	0.7326	0.0850
37	deb1	-0.2030	0.3111	0.7494	0.0842
38	deb2	-0.2660	0.3123	0.7439	0.0847
39	deb3	-0.1337	0.3013	0.7211	0.0844
40	dec	-0.2652	0.3078	0.7345	0.0850
41	def	-0.0324	0.3050	0.7382	0.0837
42	dk	-0.1784	0.3062	0.7973	0.0738
43	es11*	-0.9529	0.2284	0.6055	0.0934
44	es12*	-0.8981	0.2397	0.5882	0.0956
45	es13*	-0.8828	0.2433	0.5757	0.0972
46	es21*	-0.8472	0.2327	0.5866	0.0955
47	es22*	-0.8067	0.2416	0.5926	0.1046
48	es23*	-0.8367	0.2404	0.5903	0.1020
49	es24*	-0.8080	0.2440	0.5880	0.1064
50	es3*	-0.8296	0.2419	0.5913	0.1001
51	es41*	-0.8826	0.2362	0.5826	0.0971
52	es42*	-0.9053	0.2492	0.5776	0.1048
53	es43*	-0.9778	0.2402	0.6198	0.1005
54	es51*	-0.8109	0.2332	0.5509	0.1085
55	es52*	-0.9414	0.2539	0.5366	0.1139
56	es53*	-0.7892	0.2383	0.5178	0.1102
57	es61*	-0.9604	0.2413	0.5983	0.0963
58	es62*	-0.9545	0.2606	0.5344	0.1144
59	fr1	-0.4370	0.2779	0.7797	0.0784
60	fr21	-0.2209	0.3377	0.7213	0.0928
61	fr22	-0.1429	0.3156	0.7448	0.0878
62	fr23	-0.2334	0.2778	0.7045	0.0907
63	fr24	-0.2298	0.2587	0.6597	0.0919
64	fr25	-0.2424	0.2165	0.6067	0.0929
65	fr26*	-0.6516	0.3209	0.7180	0.0993
66	fr3	-0.0549	0.3161	0.7810	0.0842
67	fr41	-0.2754	0.3073	0.7473	0.0839
68	fr42	-0.2344	0.3008	0.7195	0.0869
69	fr43	-0.3660	0.2929	0.7037	0.0900
70	fr51*	-0.5063	0.2321	0.5635	0.1035

OBS	NUTS	β	σ_β	ρ	σ_ρ
71	fr52*	-0.8784	0.2505	0.5367	0.1021
72	fr53*	-0.6159	0.2475	0.5418	0.1074
73	fr61*	-0.7290	0.2477	0.5638	0.1026
74	fr62*	-0.4777	0.2150	0.5003	0.1194
75	fr63	-0.3155	0.2185	0.6185	0.0978
76	fr71*	-0.6536	0.2949	0.6966	0.0999
77	fr72*	-0.5235	0.2458	0.6629	0.1052
78	fr81	-0.3972	0.2135	0.5754	0.1105
79	fr82	-0.4783	0.2419	0.7016	0.0884
80	gr11	-0.1155	0.1817	0.6062	0.1079
81	gr12	-0.1547	0.1858	0.5879	0.1064
82	gr13	-0.1812	0.1857	0.6188	0.1065
83	gr14	-0.1617	0.1866	0.6220	0.1083
84	gr21	-0.2089	0.1844	0.5850	0.1114
85	gr22	-0.1501	0.1749	0.6405	0.1024
86	gr23	-0.1873	0.1858	0.6210	0.1102
87	gr24	-0.0759	0.1725	0.6008	0.1103
88	gr25	-0.1908	0.1757	0.6208	0.1074
89	gr3	-0.1886	0.1964	0.5836	0.1154
90	gr41	-0.0904	0.1919	0.5845	0.1097
91	gr42	-0.1135	0.1829	0.5821	0.1079
92	gr43	-0.1481	0.1723	0.6428	0.1060
93	it11	-0.4943	0.2629	0.6825	0.0961
94	it12	-0.5118	0.2725	0.7027	0.0926
95	it13	-0.1665	0.2006	0.6727	0.0970
96	it2	-0.0976	0.2019	0.7187	0.0873
97	it31	-0.0756	0.1800	0.6922	0.0909
98	it32	-0.0678	0.1812	0.7005	0.0925
99	it33	-0.0264	0.1749	0.6883	0.0919
100	it4	-0.0933	0.1903	0.6614	0.1012
101	it51	-0.0981	0.1916	0.6748	0.0990
102	it52	-0.0947	0.1589	0.6820	0.0978
103	it53	-0.0565	0.1714	0.6754	0.1023
104	it6	-0.0952	0.1680	0.6704	0.0973
105	it71	-0.0951	0.1769	0.6139	0.1048
106	it72	-0.1186	0.1912	0.5818	0.1137
107	it8	-0.2119	0.1814	0.6346	0.1078
108	it91	-0.2007	0.1810	0.6226	0.1045
109	it92	-0.2078	0.1823	0.6307	0.1079
110	it93	-0.2123	0.1838	0.6351	0.1073
111	ita	-0.3251	0.1786	0.6641	0.0969
112	itb	-0.1837	0.1904	0.6836	0.0939
113	lu	-0.2880	0.2983	0.8040	0.0752
114	nl12	-0.0172	0.3131	0.7954	0.0742
115	nl13	-0.1470	0.2980	0.7908	0.0706
116	nl2	-0.1272	0.3108	0.7866	0.0755
117	nl31	-0.0648	0.3039	0.7947	0.0720
118	nl32	0.0327	0.3051	0.7826	0.0732

OBS	NUTS	β	σ_β	ρ	σ_ρ
119	nl33	0.0414	0.3229	0.7659	0.0795
120	nl34	-0.0369	0.3167	0.7799	0.0827
121	nl41	-0.0431	0.3182	0.7821	0.0772
122	nl42	-0.1558	0.3109	0.7819	0.0816
123	pt11*	-0.7989	0.2427	0.5621	0.0998
124	pt12*	-0.8297	0.2405	0.5388	0.1013
125	pt13*	-0.8889	0.2471	0.5222	0.1051
126	pt14*	-0.8789	0.2725	0.5446	0.1086
127	pt15*	-0.8882	0.2457	0.5369	0.1047
128	uk1	-0.0351	0.2141	0.6803	0.0871
129	uk2	-0.1025	0.2502	0.7078	0.0856
130	uk3	-0.1085	0.2554	0.6858	0.0928
131	uk4	-0.2146	0.3115	0.7332	0.0855
132	uk5	-0.1517	0.2779	0.6553	0.0957
133	uk6*	-0.7126	0.2870	0.5241	0.1071
134	uk7	-0.1408	0.2375	0.6389	0.0942
135	uk8	-0.0041	0.2123	0.6752	0.0877
136	uk9	-0.5423	0.3013	0.5389	0.1107
137	uka	-0.0211	0.2172	0.6869	0.0847
138	ukb*	-0.7504	0.2520	0.6019	0.0990

FIGURE 5. Converging and Non-Converging Regions

5. CONCLUSIONS

We argue that problems created for conventional convergence regressions by shifts in regime as one moves across the spatial regions can be accommodated by a Bayesian spatial autoregressive locally linear estimation approach. Additional problems that arise due to non-constant variance and outliers can also be ameliorated using this approach. We define a local convergence concept and provide an estimation method that we label BSALE to draw inferences regarding this notion of convergence. We demonstrate that inferences regarding convergence differ when using the BSALE methodology and more traditional SAR models based on the entire sample.

One aspect of this methodology is reliance on a spatial autoregressive model to account for latent unobservable factors that influence economic growth but are not typically accounted for in β-convergence models. We argue that as in the case of lagged dependent variables in time-series modeling, spatial lags can filter adverse impacts arising from excluded variables. Another key facet of our BSALE approach is the use of a robust Bayesian variant of the spatial autoregressive local estimation (SALE) model set forth in Pace and LeSage (2004). This type of locally linear sub-sample estimation produces estimates that converge to robust Bayesian spatial autoregressive estimates based on the entire sample as the size of the sub-sample increases towards use of all observations. This allows practitioners to avoid use of a single bandwidth or sub-sample size on which they will ultimately proceed to draw inferences. The continuous nature of the mapping between locally linear and global estimates allows one to consider the role of sub-sample size on the resulting conclusions regarding convergence. For our sample of 138 European regions, we find evidence of substantial spatial dependence for a number of regions. The convergence parameter β varies substantially among countries as well as among regions within a country. More precisely, we find some evidence of convergence for a total of 31 regions in Spain and Portugal as well as some regions in southern France. These conclusions regarding convergence are similar for sub-sample sizes varying from roughly one-fourth to three-fourths of the sample size.

There are several areas where the approach set forth here could be extended or enhanced. These methods could be extended to the case of a spatial Durbin model, where spatial lags of the initial levels are included as an explanatory variable in the model. Spatial error models where the disturbances are modeled as following a spatial autoregressive process would be another extension of the approach. A place for enhancement would be a formal method for identifying the optimal sub-sample size to use in the Bayesian SALE estimation method.

REFERENCES

Abreu, M., H.L.F. de Groot, and R.J.G.M. Florax, 2005. "Space and Growth: A Survey of Empirical Evidence and Methods," *Région et Développement* 21, 13-44.
Anselin, L., 1988. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers: Dordrecht.

Armstrong H., 1995. "Convergence Among the Regions of the European Union," *Papers in Regional Science* 74, 143-152.

Azariadis, C. and A. Drazen, 1990. "Threshold Externalities in Economic Development," *Quarterly Journal of Economics* 105, 501-526.

Barro, R.J. and X. Sala-I-Martin, 1991. "Convergence Across States and Regions," *Brookings Papers on Economic Activity* 1, 107-182.

_____, 1992. "Convergence," *Journal of Political Economy* 100, 223-251.

_____, 1995. *Economic Growth Theory*. McGraw-Hill: New York.

Baumol, W.J., 1986. "Productivity Growth, Convergence and Welfare: What the Long Run Data Show," *American Economic Review* 76, 1072-1085.

Brock, W.A. and S.N. Durlauf, 2001. "Growth Empirics and Reality," *The World Bank Economic Review* 15, 229-272.

Brundson, C., A.S. Fotheringham, and M. Charlton, 1996. "Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity," *Geographical Analysis* 28, 281-298.

Conley, T.G. and E. Ligon, 2002. "Economic Distance, Spillovers, and Cross Country Comparisons," *Journal of Economic Growth* 7, 157-187.

DeLong, J.B., 1988. "Productivity Growth, Convergence and Welfare: Comment," *American Economic Review* 78, 1138-1154.

DeLong, J.B. and L.H. Summers, 1991. "Equipment Investment and Economic Growth," *Quarterly Journal of Economics* 106(2), 445-502.

Desdoigts, A., 1999. "Patterns of Economic Development and the Formation of Clubs," *Journal of Economic Growth* 4, 305-330.

Durlauf, S.N., 2000. "Econometric Analysis and the Study of Economic Growth: A Skeptical Perspective," *Working Paper*, Department of Economics, University of Wisconsin.

_____, 2001. "Manifesto for a Growth Econometrics," *Journal of Econometrics* 100, 65-69.

Durlauf, S.N., A. Kourtellos A., and A. Minkin, 2001. "The Local Solow Growth Model," *European Economic Review* 45, 928-940.

Durlauf, S.N. and P.A. Johnson, 1995. "Multiple Regimes and Cross-Country Growth Behaviour," *Journal of Applied Econometrics* 10, 365-384.

Durlauf, S.N. and D. Quah, 1999. "The New Empirics of Economic Growth," in J. Taylor and M. Woodford (eds.), *Handbook of Macroeconomics.* Elsevier Science: North-Holland.

Easterly, W. and R. Levine, 2001. "It's Not Factor Accumulations: Stylized Facts and Growth Models," *World Bank Economic Review* 15, 177-219.

European Commission, 1999. *The European Regions: Sixth Periodic Report on the Socio-Economic Situation in the Regions of the European Union*, Luxembourg, Official Publication Office.

Ertur, C. and W. Koch, 2005. "Growth, Technological Interdependence and Externalities: Theory and Evidence," *Journal of Applied Econometrics*.

Ertur, C., J. Le Gallo, and C. Baumont, 2006. "The European Regional Convergence Process, 1980-1995: Do Spatial Dependence and Spatial Heterogeneity Matter?" *International Regional Science Review* 29, 1-32.

Fingleton, B., 1999. "Estimates of Time to Convergence: An Analysis of Regions of European Union," *International Regional Science Review* 22, 5-34.

Fingleton, B. and J. McCombie, 1998. "Increasing Returns and Economic Growth : Some Evidence for Manufacturing from the European Union Regions," *Oxford Economic Papers* 50, 89-105.

Galor, O., 1996. "Convergence? Inferences from Theoretical Models," *Economic Journal* 106, 1056-1069.

Gelfand, A. E. and A.F.M Smith, 1990. "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85, 398-409.

Geweke, J., 1993. "Bayesian Treatment of the Independent Student t Linear Model," *Journal of Applied Econometrics* 8, 19-40.

Islam, N., 1995. "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics* 110, 1127-1170.

_____, 2003. "What Have We Learnt from the Convergence Debate?" *Journal of Economic Surveys* 17, 309-362.

Koop, G., 2003. *Bayesian Econometrics*. John Wiley & Sons.

Kubo, Y., 1995. "Scale Economies, Regional Externalities and the Possibility of Uneven Regional Development," *Journal of Regional Science* 35, 318-328.

Lange, K.L., R.J.A. Little, and J.M.G. Taylor, 1989. "Robust Statistical Modeling Using the t Distribution," *Journal of the American Statistical Association* 84, 881-896.

Le Gallo, J. and S. Dall'erba, 2006. "Evaluating the Temporal and Spatial Heterogeneity of the European Convergence Process: 1980-1999," *Journal of Regional Science*.

Le Gallo, J. and C. Ertur, 2003. "Exploratory Spatial Data Analysis of the Distribution of Regional Per Capita GDP in Europe, 1980-1995," *Papers in Regional Science* 82, 175-201.

Le Gallo, J. and C. Baumont, 2003. "A Spatial Econometric Analysis of Convergence across European Regions, 1980-1995," in B. Fingleton (ed.) *European Regional Growth*. Springer-Verlag, Berlin.

LeSage, J.P., 1997. "Bayesian Estimation of Spatial Autoregressive Models," *International Regional Science Review* 20(1&2), 113-129.

_____, 2004. "A Family of Geographically Weighted Regression Models," in L. Anselin, R.J.G.M. Florax, and S.J. Rey (eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Springer-Verlag: Berlin, pp. 241-264.

López-Bazo, E., E. Vayá, and M. Artís, 2004. "Regional Externalities and Growth: Evidence from European Regions," *Journal of Regional Science* 44, 43-73.

López-Bazo, E., E. Vayá, A. Mora, and J. Suriñach, 1999. "Regional Economic Dynamics and Convergence in the European Union," *Annals of Regional Science* 33, 343-370.

Mankiw, N.G., 1995. "The Growth of Nations," *Brooking Papers on Economic Activity* 1, 275-310.

Mankiw, N.G., D. Romer, and D.N. Weil, 1992. "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics* 107, 407-437.

Martin, P. and G.I.P. Ottaviano, 1999. "Growing Locations: Industry Location in a Model of Endogenous Growth," *European Economic Review* 43, 281-302.

McMillen, D.P., 1996. "One Hundred Fifty Years of Land Values in Chicago: A

Nonparameteric Approach," *Journal of Urban Economics* 40, 100-124.

McMillen, D.P. and J.F. McDonald, 1997. "A Nonparametric Analysis of Employment Density in a Polycentric City," *Journal of Regional Science* 37, 591-612.

Moreno, R. and B. Trehan, 1997. "Location and the Growth of Nations," *Journal of Economic Growth* 2, 399-418

Neven, D. and C. Gouyette, 1995. "Regional Convergence in the European Community," *Journal of Common Market Studies* 33, 47-65.

Pace, R.K. and R. Barry, 1997. "Quick Computation of Spatial Autoregressive Estimators," *Geographical Analysis*, 29, 232-246.

Pace, R.K and J.P. LeSage, 2004. "Spatial Autoregressive Local Estimation," in A. Getis, J. Mur, and H. Zoller (eds.), *Recent Advances in Spatial Econometrics.* Palgrave Publishers, pp. 31-51.

Quah, D., 1996. "Regional Convergence Clusters across Europe," *European Economic Review* 40, 951-958.

Rey, S.J. and M.J. Janikas, 2005. "Regional Convergence, Inequality, and Space," *Journal of Economic Geography* 5, 155-176.

Sala-I-Martin, X., 1996a. "Regional Cohesion: Evidence and Theories of Regional Growth and Convergence," *European Economic Review* 40, 1325-1352.

_____, 1996b. "The Classical Approach to Convergence Analysis," *Economic Journal* 106, 1019-1036.

Solow, R.M., 1956. "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 70, 65-94.

Temple, J., 1998. "Robustness Tests of the Augmented Solow Model," *Journal of Applied Econometrics* 13, 361-375.

_____, 1999. "The New Growth Evidence," *Journal of Applied Econometrics* 37, 112-156.

Vayá, E., E. López-Bazo, R. Moreno, and J. Suriňach, 2004. "Growth and Externalities across Economies: An Empirical Analysis Using Spatial Econometrics," in L. Anselin, R.J.G.M. Florax, and S.J. Rey (eds.), *Advances in Spatial Econometrics: Methodology, Tools and Applications.* Springer-Verlag: Berlin.

APPENDIX. MCMC SAMPLER FOR THE HETEROSCEDASTIC SAR MODEL

Extension of the Bayesian heteroscedastic linear regression model of Geweke (1993) to the case of a spatial autoregressive model is described here along with the MCMC estimation scheme. This model takes the form:

(A1) $y = \rho W y + X \beta + \varepsilon$

where y, X, and ε are as described above and the scalar parameter ρ measures the strength of spatial dependence, with the term Wy representing a spatial lag of the dependent variable. This model suggests that growth rates in neighboring areas measured by the spatial lag Wy exert an influence on the growth rate of region i. This dependence is introduced in the model by the N by N spatial weight matrix, which has values of 0.1 in row i column j for observations j representing the nearest 10 neighbors to observation i. Other

elements of the matrix W are set to zero. This produces an explanatory variable Wy reflecting the average growth rates from the 10 nearest neighboring regions.

The likelihood function for this model includes an additional term reflecting the Jacobian of the transformation from ε to y, $|I_n - \rho W|$, taking the form:

$$p\left(y^{\overset{\circ}{a}} \mid \rho, \beta, h, \Omega \right) = |I_n - \rho W|^{N/2} \frac{h^{N/2}}{(2\pi)^{N/2}} \left[\exp\left(-\frac{h}{2} \right. \right.$$

$$\bullet \left(y* - \rho Wy* - X*\beta \right)' \left(y* - \rho Wy* - X*\beta) \right) \Big].$$

The conditional posterior distributions for β and h take similar forms to those for the case of a Bayesian linear regression model with a re-definition of:

$$\tilde{\beta} = \left(X'\Omega^{-1}X \right)^{-1} X'(I_n - \rho W)\Omega^{-1} y$$

$$\bar{s}^2 = \left(y - \rho Wy - X\beta \right)' \Omega^{-1} \left(y - \rho Wy - X\beta \right) + \underline{v s}^2 \ .$$

Sampling for the parameters β, h, λ_i, and d_λ can be achieved using an analogous approach to that for the heteroscedastic linear regression model, e.g., Koop (2003).

For the case of diffuse priors for β, the conditional posterior for the parameter ρ reflecting spatial dependence takes the form:

(A2) $p\left(\rho \mid \beta, h, \lambda \right) \propto |I_n - \rho W| |X'\Omega^{-1} X| \left(s^2 \right)^{-(N-k)/2} p(\rho)$

where $p(\rho)$ denotes the uniform prior on $\rho \in [-1, 1]$. A problem arises here in that this distribution is not one for which established algorithms exist to produce random draws. We can, however, rely on univariate numerical integration of the conditional posterior of ρ. This requires evaluating the conditional posterior over a grid of values from -1 to 1, which can be efficiently done using a vectorization approach described in Pace and Barry (1997) for maximum likelihood estimation of this model. They provide a computationally efficient approach to calculating the log determinant term involving $(I_n - \rho W)$ over a grid of values from -1 to 1, which can be implemented prior to beginning the MCMC sampler.

Applying a log transformation to the conditional posterior we can express the term $\log(s^2)$ as a vector over a grid of $j = 1, \ldots, q$ values for the parameter ρ ranging from -1 to 1, taking the form:

$$
\text{(A3)} \quad \log \begin{pmatrix} s(\rho_1)^2 \\ s(\rho_2)^2 \\ \vdots \\ s(\rho_q)^2 \end{pmatrix} \propto \begin{pmatrix} \log\big(\phi(\rho_1)\big) \\ \log\big(\phi(\rho_2)\big) \\ \vdots \\ \log\big(\phi(\rho_q)\big) \end{pmatrix}
$$

where $\phi(\rho_j) = e_o' e_o - 2\rho_j e_d' e_o + \rho_j^2 e_d' e_d$, $j = 1, \ldots, q$, resulting in:

$$
e_o = y^* - X^* \gamma_o
$$
$$
e_d = Wy^* - X^* \gamma_d
$$
$$
\gamma_o = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}y
$$
$$
\gamma_d = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}Wy .
$$

This vectorization along with a vector of values from calculation of the k by k log determinant term: $\left| X'\Omega^{-1}X \right|$ (which does not depend on ρ) and the vectorized log determinant of $(I_n - \rho W)$ over the grid of q values for ρ results in a simple numerical integration problem that can be solved rapidly using Simpson's rule. Results from this integration are used to construct the cumulative distribution function for the conditional posterior distribution of the parameter ρ that is then used to produce a draw from this distribution using "inversion." Keep in mind that on the next pass through the MCMC sampler, we need to integrate the conditional posterior again. This is because the distribution is conditional on changing values for the other parameters λ_i, β, h in the model, which obviously produce an altered expression for s^2 and $\left| X'\Omega^{-1}X \right|$ in the conditional distribution for ρ.

Following Koop (2003), we implemented this model using a hyperparameter λ to control the degree prior belief in heteroscedasticity. The global sample data was used to produce a posterior distribution for the parameter d_λ. The posterior estimate for d_λ indicated heteroscedasticity, having a mean of 3.98, median equal to 3.74, and mode of 3.60. The posterior distribution for d_λ is shown in Figure A1.

The posterior mean from this global estimation was used as a degenerate prior in the locally linear models. This saves the computational burdens associated with the tuning parameter c for the random-walk Metropolis-Hastings sampling of the parameter d_λ during the locally linear estimation procedures. Experimentation with the global sample indicated that use of the fixed value for d_λ near the posterior mean from the estimate produced estimates and inferences very similar to those for the case where this parameter was estimated rather than fixed.
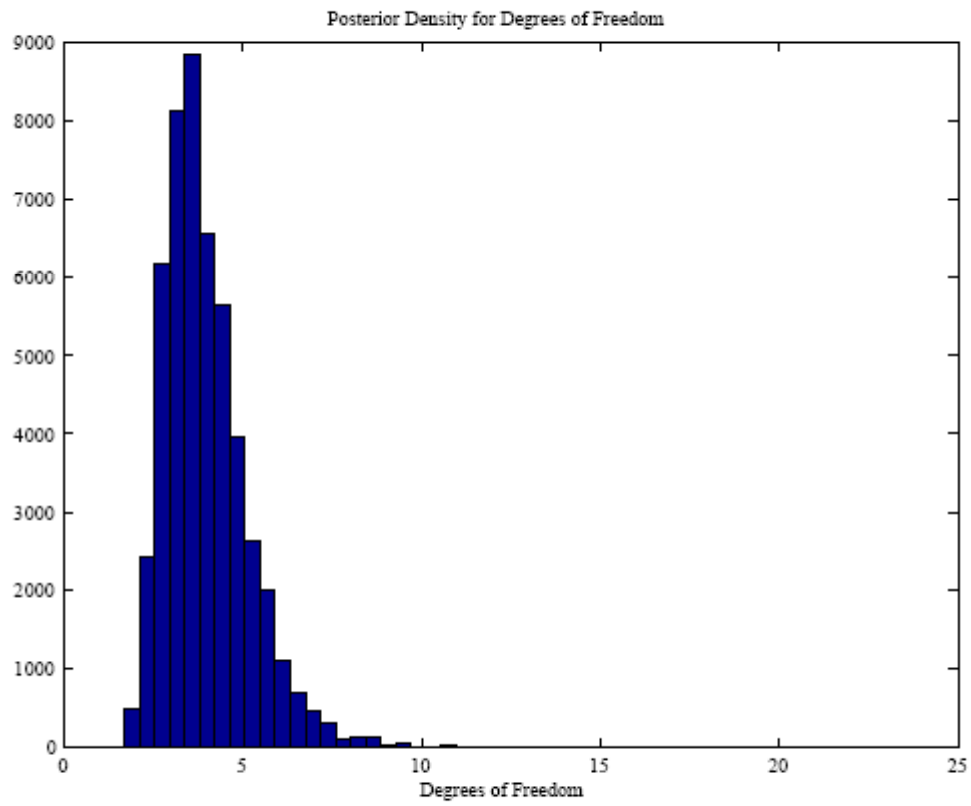
FIGURE A1. Posterior Distribution for SAR Model Parameter λ