geographical analysis

Geographical Analysis ISSN 0016-7363

Growth and Convergence in a Multiregional Model with Space–Time Dynamics

Paul Elhorst,[1] Gianfranco Piras,[2,3] Giuseppe Arbia[4]

[1]Faculty of Economics and Business, University of Groningen, Groningen, The Netherlands, [2]REAL, University of Illinois, Urbana-Champaign, IL, U.S.A., [3]IDEAR, Universidad Catolica del Norte, Antofagasta, Chile, [4]Faculty of Economics, University of Chieti, Pescara, Italy

The goal of this article is to test four distinct hypotheses about whether the relative location of an economy affects economic growth and economic well-being using an extended Solow–Swan neoclassical growth model that incorporates both space and time dynamics. We show that the econometric specification takes the form of an unconstrained spatial Durbin model, and we investigate whether the results depend on some methodological issues, such as the choice of the time span and the inclusion of fixed effects. To estimate the fixed effects spatial Solow–Swan model, we adjust the Arrelano and Bond (1991) generalized method-of-moments (GMM) estimator to deal with endogeneity not only arising from the initial income level, as in the basic model, but also from the initial income levels and economic growth rates observed in neighboring economies.

Introduction

Despite recent developments in the growth literature, the Solow–Swan neoclassical growth model continues to be of great theoretical and empirical interest. One relevant question asks to what extent differences found in the speed of convergence are attributable to the way in which applied econometricians analyze a given body of data. The empirical literature has moved in several different directions. One part of the literature uses data in a cross-section, while another part combines time-series and cross-sectional (TSCS) data. Mankiw, Romer, and Weil (1992) and Barro and Sala-i-Martin (1995) are classic references of the first approach, while Islam (1995) is the most prominent example of the second. Within the literature that combines TSCS data, considerable discussion also occurs about two concerns: the appropriate time length to use when a total sample period is divided into several shorter periods, and the inclusion of fixed effects. An econometric problem related

Correspondence: J. Paul Elhorst, Faculty of Economics and Business, University of Groningen, Groningen, The Netherlands
e-mail: j.p.elhorst@rug.nl

Submitted: April 23, 2008. Revised version accepted: April 27, 2009.

to this second issue is that the ordinary least squares (OLS) estimator of the growth regression equation is inconsistent when the number of observations in the time domain is fixed. This is because the regression equation is indeed equivalent to a dynamic panel model. To remove this inconsistency, generalized method-of-moments (GMM) estimators have been suggested (Arrelano and Bond 1991).

The spatial econometrics literature (Fingleton 2003) points out that a more satisfactory understanding of economic growth requires an appreciation of how economies interact with one another, because economies' income levels are interdependent. Consider, for example, the savings rate. According to standard economic theory, saving and investment are always equal. People cannot save without investing their money somewhere, and they cannot invest without using somebody's savings. This is true for the world in general, but it is not true for individual economies. Capital can flow across borders; hence the amount an individual economy saves does not have to be the same as the amount it invests. In other words, per capita income in one economy depends partly on the savings rates of neighboring economies.

The hypothesis that the relative location of an economy affects economic growth has recently been underpinned by theoretical extensions of the Solow–Swan model (López-Bazo, Vayá, and Artis 2004; Fingleton and López-Bazo 2006; Ertur and Koch 2007). Moreover, a vast empirical literature exists that supports this hypothesis using data in a cross-section (for an overview, see Fingleton 2003; Abreu, de Groot, and Florax 2005; Rey and Janikas 2005). However, empirical research to test this hypothesis using panel data techniques is still scarce because the simultaneous modeling of dynamics in space and time requires quite complex stochastic specifications. One relevant exception is the study by Badinger, Müller, and Tondl (2004), which separates both types of dynamics. First, the variables are spatially filtered to account for spatial dependencies between regional observations, and, second, conditional upon these spatially filtered variables, a dynamic panel data model is estimated using GMM. However, it is more likely that these two types of dynamics are interdependent since short-term dynamics may have major long-run structural impacts, and, conversely, structural changes may alter the short-run dynamics of certain economies.

This paper is organized as follows. The section that follows presents the basic Solow–Swan model and its spatial extension. We use this model to derive four distinct testable hypotheses of all the possible paths along which the relative location of an economy can affect economic growth and economic well-being. Furthermore, we show that its econometric specification takes the form of an unconstrained spatial Durbin model. In presenting the empirical results, particular attention is devoted to the choice of the time span and the inclusion of fixed effects. Finally, to estimate the fixed effects spatial Solow–Swan model, we adjust the Arrelano and Bond (1991) GMM estimation procedure to deal with endogeneity arising not only from the initial income level as in the basic model but also from the initial income levels and growth rates observed in neighboring economies.

The Solow–Swan model and its spatial extension

The basic Solow–Swan model is characterized by the following expression for the steady-state per capita income q_t at time t:

$$\ln(q_t) = \ln(A_{t-T}) + gT + \frac{\alpha}{1-\alpha}\ln[s/(n+g+\delta)] \tag{1}$$

where T denotes the time span of the growth period considered, A_{t-T} is the state of technology at the beginning of the observation period, α is the cost share of capital in production under a Cobb–Douglas technology, s is the savings rate, n is the population growth rate, g is the rate of technological progress, and δ is the depreciation rate.

The convenient growth-initial income-level regression equation is a linear approximation of the growth rate of output per capita, which is derived from a linear approximation of the dynamics around the steady state in equation (1) using a Taylor expansion (see Mankiw, Romer, and Weil 1992, pp. 422–23). This yields

$$\frac{\ln(q_t/q_{t-T})}{T} = \beta_0 + \beta_1 \ln(q_{t-T}) + \beta_2 \ln[s/(n+g+\delta)] + \varepsilon \tag{2}$$

where $\beta_0 = (1-e^{-\lambda T})[\ln(A_{t-T}) + gT]/T$, $\beta_1 = -(1-e^{-\lambda T})/T$, and $\beta_2 = \frac{\alpha}{1-\alpha}(1-e^{-\lambda T})/T$ (see Barro and Sala-i-Martin 1995; Islam 1995; Fingleton 2003). In its simplest version, ε represents a normally distributed and independent error term, as a result of which equation (2) can be estimated by OLS. This model implies that economies tend toward the same equilibrium growth path for capital and hence output per capita, except for differences in s, n, g, and δ. The annual speed of convergence implied by the parameter estimate of β_1 is $\lambda = -\ln(1+\beta_1 T)/T$ (Fingleton 2003, p. 25), and capital's share of income implied by the parameter estimates of β_1 and β_2 is $\alpha = \beta_2/(\beta_2 - \beta_1)$. According to Mankiw, Romer, and Weil (1992, p. 410) and Durlauf and Quah (1999, p. 276), this share can be used as an additional tool to test whether the basic Solow–Swan model is the correct specification (e.g., in addition to the classical goodness-of-fit measures), since its value is expected to be roughly one-third.

Recently, the hypothesis that the relative location of an economy—the effect of being located closer or farther away from other specific economies—is a determinant of economic growth and the steady-state position of an economy has been underpinned by economic–theoretical models (López-Bazo, Vayá, and Artis 2004; Fingleton and López-Bazo 2006; Ertur and Koch 2007). To model interdependence across space, López-Bazo, Vayá, and Artis (2004) assume that spatial externalities derive from (physical and human) capital accumulation, while Ertur and Koch (2007) assume that spatial externalities are generated from technological interdependencies. More formally, Ertur and Koch (2007) model technology as being dependent on three terms in the following way:

$$A_{it} = \Omega_t k_i^{\varphi}(t) \prod_{j=1}^{N} A_{jt}^{\gamma w_{ij}} \tag{3}$$

where $i\,(=1,\ldots,N)$ refers to a particular economy. Just as in the basic Solow–Swan model, part of technological progress is assumed to be exogenous and identical to all economies: $\Omega_t = \Omega_0 e^{gt}$. The level of technology in economy i is also related to the level of physical capital per worker k_{it} in that particular economy, because of knowledge spillovers generated by physical capital. The magnitude of this physical capital externality is measured through the parameter φ (with $0<\varphi<1$). Finally, it is assumed that these externalities affect neighboring economies $j\,(=1,\ldots,N)$ according to some distance decay function γw_{ij}, where γ is an unknown parameter to be estimated and w_{ij} is an element of a $N \times N$ spatial weights matrix W describing the spatial arrangement of the N economies.

Under the assumption that the speed of convergence is identical for all economies, this extension gives the following expression for the steady-state per capita income:

$$\ln(q_t) = \gamma\frac{1-\alpha}{1-\alpha-\varphi}W\ln(q_t) + \frac{1}{1-\alpha-\varphi}\ln(A_{t-T}) + gT + \frac{\alpha+\varphi}{1-\alpha-\varphi}\ln[s/(n+g+\delta)] + \gamma\frac{\alpha}{1-\alpha-\varphi}W\ln[s/(n+g+\delta)] \tag{4}$$

where $\ln(q_t)$, $\ln(A_{t-T})$, gT, and $\ln[s/(n+g+\delta)]$ now denote $N \times 1$ vectors of the corresponding variable in each economy. This model simplifies to the basic Solow–Swan model when both γ and φ are zero.

A linear approximation to the dynamics around the steady state in equation (4), using a Taylor expansion, again produces a growth-initial income-level regression equation:

$$\frac{\ln(q_t/q_{t-T})}{T} = \rho W\frac{\ln(q_t/q_{t-T})}{T} + \beta_0 + \beta_1\ln(q_{t-T}) + \beta_2\ln[s/(n+g+\delta)] + \beta_3 W\ln(q_{t-T}) + \beta_4 W\ln[s/(n+g+\delta)] + \varepsilon \tag{5}$$

where $\beta_1 = -(1-e^{-\lambda T})/T$, $\beta_2 = \frac{\alpha+\varphi}{1-\alpha-\varphi}(1-e^{-\lambda T})/T$, $\beta_3 = \gamma\frac{1-\alpha}{1-\alpha-\varphi}(1-e^{-\lambda T})/T$, $\beta_4 = -\gamma\frac{\alpha}{1-\alpha-\varphi}(1-e^{-\lambda T})/T$, and $\rho = \gamma\frac{1-\alpha}{1-\alpha-\varphi}$. This model is referred to in the spatial econometrics literature as an unconstrained spatial Durbin model, because of the spatially lagged values of both the dependent variable and the independent variables. Since $0<\alpha<1$ and $0<\varphi<1$, γ and ρ are defined on the same interval $(1/\omega_{min}, 1/\omega_{max})$, where ω_{min} denotes the smallest (i.e., most negative) and ω_{max} the largest eigenvalue of W. Note that for row-normalized spatial weights $\omega_{max} = 1$. In addition, W should satisfy the regularity conditions spelled out in Lee (2004) or Yu, de Jong, and Lee (2008).[1] The annual speed of convergence, λ, implied by the parameter estimate of β_1 (or β_3), is the same as in the classical Solow–Swan model. The unknown values of α, φ, and γ implied by the parameter estimates of β and ρ are

$$\alpha = \frac{\beta_4}{\beta_4-\beta_3},\ \varphi = \frac{\beta_2}{\beta_2-\beta_1} - \frac{\beta_4}{\beta_4-\beta_3},\ \text{and}\ \gamma = \frac{\beta_4-\beta_3}{\beta_1-\beta_2} \tag{6}$$

provided that the restriction $\beta_3 + \rho\beta_1 = 0$ holds; otherwise, these unknown parameters are overdetermined.

We presented Ertur and Koch's (2007) extension of the technology term because it leads to a full spatial Durbin specification of equation (5), whereas the extension of the technology term in López-Bazo, Vayá, and Artis (2004) yields spatially lagged values of only the growth and the initial income variables and not of $\ln[s/(n+g+\delta)]$ (see Fingleton and López-Bazo 2006 for details). In contrast, we follow López-Bazo, Vayá, and Artis (2004), as well as Fingleton and López-Bazo (2006), in that we estimate the model only under the condition that the speed of convergence is identical for all economies, whereas Ertur and Koch (2007) estimate their model in the more general context in which this assumption is removed.[2]

In contrast to equation (1), equation (4) does not give an explicit expression for the steady-state per capita income, because the term $\ln(q_t)$ also appears on the right-hand side (see Egger and Pfaffermayr 2006). However, $\ln(q_t)$ may be solved from equation (5). We first multiply equation (5) by T and rewrite it as

$$B\ln(q_t) = A\ln(q_{t-T}) + \beta_X \tag{7}$$

where $\beta_X = T\{\beta_0 + \beta_2\ln[s/(n+g+\delta)] + \beta_4 W\ln[s/(n+g+\delta)]\}$, $A = (1 + T\beta_1)I + T\beta_3 W - \rho W$, and $B = I - \rho W$. For this model to converge, Elhorst (2001) has found that the eigenvalues of the matrix $B^{-1}A$ should lie inside the unit circle (see also assumption 6 in Yu, de Jong, and Lee 2008). If ω_i denotes one of the N eigenvalues of the spatial weights matrix W, then the eigenvalues of $B^{-1}A$ are $(1 + T\beta_1 + T\beta_3 - \rho\omega_i)/(1 - \rho\omega_i)$. If the spatial weights matrix W is row-normalized, so that its largest eigenvalue equals one, then $\beta_1+\beta_3<0$. Elhorst (2001) has also found that the corresponding steady-state value can be obtained by $\ln(q_t) = (B - A)^{-1}\beta_X$. In the present context (and after eliminating T) we obtain following:

$$\ln(q_t) = \left(I + \frac{\beta_3}{\beta_1}W\right)^{-1}\left\{-\frac{\beta_0}{\beta_1} - \frac{\beta_2}{\beta_1}\ln[s/(n+g+\delta)] - \frac{\beta_4}{\beta_1}W\ln[s/(n+g+\delta)]\right\} \tag{8}$$

To test whether s, n, g, and δ in a particular economy affect the rate of economic growth and the steady-state position of that economy, we should verify whether both β_2 and $-\beta_2/\beta_1$ are significantly different from zero. The first test follows from equation (5) and the second test from equation (8). To test whether the relative location of an economy affects economic growth and of the steady-state position of a particular economy, we consider four distinct hypotheses:

H_1. The rate of growth of a particular economy is related to that of its neighbors.

H_2. The rate of growth of a particular economy is affected by s, n, g, and δ in its neighboring economies.

H_3. The steady-state position of a particular economy and s, n, g, and δ in its neighboring economies have a direct relationship.

H_4. The steady-state position of a particular economy is related to s, n, g, and δ in its neighboring economies due to indirect effects.

The first two hypotheses can be tested by verifying whether ρ and β_4 in equation (5) are significantly different from zero, respectively. The last two hypotheses can be tested by verifying whether $-\beta_4/\beta_1$ and β_3/β_1 in equation (8) are significantly different from zero, respectively. Abreu, de Groot, and Florax (2005) provide an explanation of the latter test.

Finally, the unconstrained spatial Durbin model derived in equation (5) generalizes both the spatial lag model and the spatial error model, two models that have been the main focus of the spatial econometrics literature on economic growth for a long time (see Abreu, de Groot, and Florax 2005; Arbia 2006). The spatial lag is obtained when $\beta_3 = \beta_4 = 0$. This implies that the spatial lag model offers the opportunity to test only hypothesis H_1. The spatial error model is obtained when two nonlinear common factor constraints are imposed on the coefficients: $\beta_3 = -\rho\beta_1$ and $\beta_4 = -\rho\beta_1$.[3] In fact, this model is also known as a constrained spatial Durbin model. The first constraint is also dictated by the economic–theoretical model to avoid overdetermining the unknown parameters of this model (see equation [6]). However, the second constraint is unnecessarily restrictive because no theoretical or empirical reason exists to impose it. These constraints will be tested in the subsequent empirical analysis.

Regression results of the basic Solow–Swan model

In this study, we use data drawn from the Cambridge Econometrics European Regional database. The data set consists of 193 regions over the period 1977–2002 across 15 countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Spain, Sweden, and the United Kingdom (see Fig. 1).

The dependent variable is per capita GDP, *s* is measured as the ratio between investment expenditures and gross value-added, and *n* is measured as the growth rate of the working population over time. We assume that $(g+\delta)$ is the same for all

Figure 1. Map of the study area: 193 EU NUTS-2 regions across 15 countries.

economies and equal to 0.05 (Mankiw, Romer, and Weil 1992; Islam 1995; Ertur and Koch 2007); *g* primarily reflects the advancement of knowledge, which is not region- or country-specific. Similarly, no strong reason exists to expect depreciation rates to vary greatly across regions or countries.[4]

Time span of the growth period

The first column of Table 1 reports the estimation results of the basic Solow–Swan model when using the data in one single cross-section; that is, if $T = 25$ and the number of observations equals the number of spatial units in the sample (193). The average annual growth rate over the entire period 1977–2002 is explained by the initial level of per capita GDP in 1977 and the averages of *s*, *n*, *g*, and δ for the entire period. Using the single cross-section approach has three potential drawbacks. First, it only utilizes data at the beginning and the end of the sample period. Second, it assumes that *s* and *n* are constant over the sample period. Third, it overlooks the possibility that different growth paths may lead to similar results in terms of convergence.

Results in the second column of Table 1 are obtained by moving to cross-sections for shorter periods. In such a pooled regression, the average annual growth rate over each period is explained by the initial level of per capita GDP for that particular period, as well as the averages of *s*, *n*, *g*, and δ for that particular period.[5] The question arises: what is the appropriate length of such periods? A time span of just one year is possible since the underlying data set provides annual information. However, often yearly time spans are said to be too short for studying growth convergence because short-term disturbances may loom large in such brief time spans (Islam 1995). Therefore, we consider five-year time spans, $T = 5$, just as in Islam (1995). The number of observations in this regression equals $193 \times 5 = 965$ (1977–1982, 1982–1987, 1987–1992, 1992–1997, 1997–2002).

The first two column entries in Table 1 demonstrate that the coefficient estimates and the *t*-values of the constant and the initial income-level variable of the single cross-section approach and of the TSCS approach do not differ to any great extent. By contrast, the R^2 of the TSCS approach over a five-year time span amounts to 0.05, which is considerably less than the R^2 of the single cross-section approach of 0.26. The explanation is that the increase in the number of observations amplifies the variation in the dependent variable. The capital's share of income according to the single cross-section approach is 0.21, which is <0.30 in the TSCS approach. Thus, the single cross-section approach produces the greatest deviation from capital's supposed one-third share of income. According to the single cross-section approach, the speed of convergence is 1.08% per year, which is higher than the speed of convergence of 0.86% implied by the TSCS approach.

One potential objection to both regressions is that the variable $\ln[s/(n+g+\delta)]$ is not exogenous. For this reason, we used the Wu variable addition test to examine the endogeneity of $\ln[s/(n+g+\delta)]$ (see Greene 2005, pp. 321–25 for details). First, the variable $\ln[s/(n+g+\delta)]$ is regressed on the intercept, the initial income value

Table 1 Estimation Results of the Basic Solow–Swan Model Using Cross-Section Data, Time-Series Cross-Section (TSCS) Data, and Panel Data (Including Fixed Effects in Space and Time)

	Cross-section over 25 years Dep. var $\ln\left(\frac{q(2002)}{q(1977)}\right)/25$			TSCS five-year periods Dep. var $\ln\left(\frac{q(t)}{q(t-5)}\right)/5$			Panel five-year periods Dep. var $\ln\left(\frac{q(t)}{q(t-5)}\right)/5$	
Explanatory variable	Coefficient	*T*-value	Explanatory variable	Coefficient	*T*-value	Explanatory variable	Coefficient	*T*-value
Constant	0.1034	9.40	Constant	0.0942	8.37			
ln (q(1977))	−0.0094	−7.95	ln $(q(t-5))$	−0.0084	−7.00	ln $(q(t-5))$	−0.0646	−1.70
$\ln\left(\frac{(s)}{(n+g+\delta)}\right)/25$	0.0025	1.98	$\ln\left(\frac{(s)}{(n+g+\delta)}\right)/5$	0.0036	2.92	$\ln\left(\frac{(s)}{(n+g+\delta)}\right)/5$	0.0308	1.12
R^2	0.255		R^2	0.052		R^2	0.527	
N	193		N	965		N	772	
Implied λ	1.076		Implied λ	0.855		Implied λ	7.800	
Implied α	0.212		Implied α	0.301		Implied α	0.323	

and the savings rate at the beginning of the observation period. Then the predicted values from this equation are added to the growth-initial income-level regression equation. The t-value on the coefficient of the fitted variable $\ln[s/(n+g+\delta)]$ is 0.65 when using the single cross-section approach, and -1.26 when using the TSCS approach. Therefore, the hypothesis that the variable $\ln[s/(n+g+\delta)]$ is exogenous cannot be rejected.

Fixed effects

A possible objection to pooling TSCS data is that this approach does not control for fixed effects. Islam (1995) argues that fixed effects should be included, since the strategy of replacing the term $(1 - e^{-\lambda T})[\ln(A_{t-T}) + gT]$ in the growth-initial income-level regression (equation [2]) with just a constant and a normally distributed error term, $\beta_0+\varepsilon$, is flawed. Since A_{t-T} not only reflects technology but also such factors as resource endowments, climate, and institutions, A_{t-T} is anything but constant among different economies and could probably be correlated with s, n, g, and δ, and thus with one of the explanatory variables in the regression model. Replacing this variable with a normally distributed error term and then estimating the model by OLS violates the condition that the explanatory variables are independent of the error term. A better solution is to replace A_{t-T} by a conventional error term as well as a dummy variable for each economy in the sample, because the latter does not need to be uncorrelated with the other explanatory variables in the model. As the rate of technological process, gT, may also change over time, this variable is replaced by time-period fixed effects (Islam 1995; Badinger, Müller, and Tondl 2004).

The growth-initial income-level regression in equation (2) extended to include one dummy for every economy in the sample is known as a dynamic panel data model (Islam 1995, p. 1136). As the lagged dependent variable appears as an explanatory variable, the OLS estimator under the fixed effects formulation is no longer consistent in the typical situation where a panel involves a large number of economies, but for a small number of observations over time.[6] To remove this inconsistency, we adopt the GMM estimator developed by Arrelano and Bond (1991; see also Baltagi 2005, pp. 136–42). This estimator takes first-differences to eliminate the intercept and the spatial fixed effects, and then applies GMM using a set of appropriate instruments.[7] Due to the reformulation in first-differences, 1987–1992 is the first period for which we can observe the fixed-effects model. Consequently, the number of observations equals $193 \times 4 = 772$ (1987–1992, 1992–1997, 1997–2002).

Column 3 in Table 1 reports the estimation results for the fixed-effects model. The R^2 increases to 0.527, which can be explained by the fixed-effects model focusing only on the time-series variation between observations. Furthermore, the deviation from capital's supposed one-third share of income becomes smaller when including fixed effects, 0.323 instead of 0.301. The speed of convergence becomes almost 10 times as large than that of the TSCS approach, 7.80% per year. This tremendous increase is nonetheless in line with Islam's (1995) findings: annual convergence rates between 3.8% and 9.1%. It is questionable, however, whether these convergence parameters are really comparable. The basic point here is that the faster convergence that we observe is conditional upon the fixed effects included in the model. The growth model with fixed effects measures the time required for the system to return to equilibrium due to shocks caused by changes in *s*, *n*, *g*, and δ (conditional on the fixed effects), while the growth model without fixed effects measures the time required before these relatively persistent differences disappear as well (see Durlauf and Quah 1999 for further details).

Regression results of the spatial Solow–Swan model

Corrections for endogenous interaction effects

Table 2 reports the estimation results of the spatial Solow–Swan model. The spatial weights matrix used in these estimations is a row-normalized binary contiguity matrix.[8] Just as in Table 1, the first column refers to cross-sectional data over the entire sample period, the second column to TSCS data pooled over periods of five years, and the last column to panel data including fixed effects. Due to the spatially lagged dependent variable, the spatial Solow–Swan model may no longer be estimated by OLS or by the Arrelano and Bond GMM estimator because it would render these estimators biased and inconsistent. To estimate the spatial Solow–Swan model based on cross-sectional data or pooled TSCS data, we used maximum likelihood (Anselin 1988). LeSage (1999) furnishes a MATLAB routine to estimate the cross-sectional version of this model, while Elhorst (2003) extends this routine for spatial panels with and without fixed effects.[9]

The spatial growth-initial income-level regression equation (5) extended to include fixed effects can be rewritten as[10]

$$\begin{aligned}\ln(q_t) =& \rho W \ln(q_t) + T\beta_0 + (1 + T\beta_1)\ln(q_{t-T}) + T\beta_2 \ln[s/(n+g+\delta)]_t \\ &+ (T\beta_3 - \rho)W\ln(q_{t-T}) + T\beta_4 W \ln[s/(n+g+\delta)]_t + \mu + \varepsilon_t\end{aligned} \tag{9}$$

where $\mu = (\mu_1, ..., \mu_N)'$ because this equation is in vector form. This equation cannot be consistently estimated by Elhorst's (2003) ML estimator because it does not cover the lagged dependent variable $\ln(q_{t-T})$. Nor can it be consistently estimated by the Arrelano and Bond's (1991) GMM estimator because it does not cover the spatially lagged dependent variable $W\ln(q_t)$. Therefore, we have developed a mixture of both estimators.

Let $Y_t = \ln(q_t)$, $Y_{t-T} = \ln(q_{t-T})$, $\tau = 1 + T\beta_1$, $\eta = T\beta_3 - \rho$, $X_t = [\ln[s/(n+g+\delta)]_t\ W\ln[s/(n+g+\delta)]_t]$, and $\beta = T[\beta_2, \beta_4]'$. Then equation (9) becomes $Y_t = \rho WY_t + T\beta_0 + \tau Y_{t-T} + \eta WY_{t-T} + X_t\beta + \mu + \varepsilon_t$. When taking first differences to eliminate the intercept and the fixed effects, we get

$$\begin{aligned}\Delta Y_t = \rho W\Delta Y_t + \tau\Delta Y_{t-T} + \eta W\Delta Y_{t-T} + \Delta X_t\beta + \Delta\varepsilon_t \\ t = 1987,\ 1992,\ 1997,\ 2002, \text{ and } T = 5\end{aligned} \tag{10}$$

Table 2 Estimation Results of the Spatial Solow–Swan Model Using Cross-Section Data, Time-Series Cross-Section (TSCS) Data, and Panel Data (Including Fixed Effects in Space and Time)

	Cross-section 25 years Dep. var $\ln\left(\frac{q(2002)}{q(1977)}\right)/25$			TSCS five-year periods Dep. var $\ln\left(\frac{q(t)}{q(t-5)}\right)/5$			Panel five-year periods Dep.var $\ln\left(\frac{q(t)}{q(t-5)}\right)/5$	
Explanatory variable	Coefficient	*T*-value	Explanatory variable	Coefficient	*T*-value	Explanatory variable	Coefficient	*T*-value
$W\times\ln\left(\frac{q(2002)}{q(1977)}\right)/25$	0.5190	7.51	$W\times\ln\left(\frac{q(t)}{q(t-5)}\right)/5$	0.6490	25.53	$W\times\ln\left(\frac{q(t)}{q(t-5)}\right)/5$	0.6230	23.36
Constant	0.0508	3.87	Constant	0.0392	3.99			
$\ln(q(1977))$	-0.0104	-5.52	$\ln(q(t-5))$	-0.0080	-4.72	$\ln(q(t-5))$	-0.0634	-1.45
$\ln\left(\frac{(s)}{(n+g+\delta)}\right)/25$	0.0010	0.92	$\ln\left(\frac{(s)}{(n+g+\delta)}\right)/5$	0.0014	1.44	$\ln\left(\frac{(s)}{(n+g+\delta)}\right)/5$	0.0316	1.27
$W\times\ln(q(1977))$	0.0055	2.38	$W\times\ln(q(t-5))$	0.0042	2.16	$W\times\ln(q(t-5))$	0.0372	2.00
$W\times\ln\left(\frac{(s)}{(n+g+\delta)}\right)/25$	0.0031	1.62	$W\times\ln\left(\frac{(s)}{(n+g+\delta)}\right)/5$	0.0031	1.93	$W\times\ln\left(\frac{(s)}{(n+g+\delta)}\right)/5$	-0.0341	-2.21
R^2	0.284		R^2	0.481		R^2	0.746	
N	193		N	965		N	772	
Implied λ	1.198		Implied λ	0.819		Implied λ	7.628	
Implied α	<0		Implied α	<0		Implied α	0.478	
Implied φ	>1		Implied φ	>1		Implied φ	<0	
Implied γ	0.213		Implied γ	0.117		Implied γ	0.750	
$\beta_1+\beta_3$	-0.0049		$\beta_1+\beta_3$	-0.0038		$\beta_1+\beta_3$	-0.0262	
$\beta_3=-\rho\beta_1$	$\chi^2(1)=0.01$	Not rej.	$\beta_3=-\rho\beta_1$	$\chi^2(1)=0.79$	Not rej.	$\beta_3=-\rho\beta_1$	$\chi^2(1)=0.07$	Not rej.
$\beta_4=-\rho\beta_1$	$\chi^2(1)=3.95$	Rej.	$\beta_4=-\rho\beta_1$	$\chi^2(1)=6.76$	Rej.	$\beta_4=-\rho\beta_1$	$\chi^2(1)=0.67$	Not rej.

where Δ is an operator that takes first-differences over a period of T years. To estimate τ, Arrelano and Bond (1991) suggest using Y_{1977} as an instrument for $\Delta Y_{1982} = Y_{1982} - Y_{1977}$ because it is highly correlated with ΔY_{1982} and not correlated with $\Delta\varepsilon_{1987} = \varepsilon_{1987} - \varepsilon_{1982}$ as long as ε is not serially correlated. X_{1977} and X_{1982} are valid instruments, too, since they are not correlated with $\Delta\varepsilon_{1987}$. Similarly, valid instruments for ΔY_{1987} are Y_{1977}, Y_{1982}, X_{1977}, X_{1982}, and X_{1987}. One can continue in this fashion, adding two extra valid instruments, one for Y and one for X, for each forward period.

In equation (10), $W\Delta Y_t$ and $W\Delta Y_{t-T}$ also need to be instrumented. To estimate the parameter ρ of the endogenous variable WY_t, Kelejian, Prucha, and Yuzefovich (2004) suggest using the instruments $[X_t WX_t \dots W^d X_t]]$, where d is a preselected constant.[11] The set of instruments satisfying the conditions spelled out both in Arrelano and Bond (1991) and in Kelejian, Prucha, and Yuzefovich (2004) consists of the variables Y_{1977} to Y_{t-2T} and X_{1977} to X_{t-T} (as in the dynamic panel data literature) and WY_{1977} to WY_{t-2T} and WX_{1977} to WX_{t-T} (as in the spatial econometrics literature when $d = 1$). Note that this setup was first considered by Revelli (2001). The Arrelano and Bond GMM estimator extended to include $W\Delta Y_t$ and $W\Delta Y_{t-T}$ therefore takes the form

$$(\rho\,\tau\,\eta\,\beta')' = [\bar{X}'Z(Z'(G \otimes I_N)Z)^{-1}Z'\bar{X}]^{-1}\bar{X}'Z(Z'(G \otimes I_N)Z)^{-1}Z'\Delta Y \qquad (11)$$

where $\otimes$ denotes the Kronecker product, and G is a matrix of order T_{GMM}, whose diagonal elements are 2 and subdiagonal elements are -1 (see Baltagi 2005, p. 137). T_{GMM} is the number of observations for each economy in the sample after taking first-differences of time-series observations over five-year time spans. In this particular case, $T_{GMM} = 4$, $\bar{X} = [(I_T \otimes W)\Delta Y\ \Delta Y_{-T}\ (I_T \otimes W)\Delta Y_{-T}\ \Delta X]$, $Z = \text{diag}(Z_{1987}, \dots, Z_{2002})$, and the block-diagonal submatrices of Z_t ($t = 1987$, 1992, 1997, 2002 and $T = 5$) are

$$Z_t = [Y_{1977}\ WY_{1977} \dots Y_{t-2T}\ WY_{t-2T}\ X_{1977}\ WX_{1977} \dots X_{t-T}\ WX_{t-T}] \qquad (12)$$

Because we assume that the data are sorted first by time and then by cross-sectional units, we finally have $G \otimes I_N$ instead of $I_N \otimes G$ as in Baltagi (2005, p. 137).

Recently, Elhorst (2010) performed Monte Carlo simulation experiments to study the performance of this extended GMM estimator. Unfortunately, he found that the bias in ρ is unacceptable, whereas the bias in ρ, when using Elhorst's ML estimator, is rather small, even though the latter estimator does not cover dynamic panel data models. On the basis of these findings, he suggests estimating ρ by ML and the remaining parameters, given ρ, by

$$(\tau\,\eta\,\beta')' = [\bar{X}'Z(Z'(G \otimes I_N)Z)^{-1}Z'\bar{X}]^{-1}\bar{X}'Z(Z'(G \otimes I_N)Z)^{-1}Z'(\Delta Y - \rho W\Delta Y) \qquad (13)$$

To demonstrate the validity of this mixed estimator, we carried out a simple Monte Carlo simulation experiment based on the reduced form of equation (9).[12] Table 3 reports the bias and root mean squared error (RMSE) of the different

Table 3 Bias* and RMSE of Elhorst's ML Estimator, the Extended Arrelano and Bond GMM Estimator (Equation [11]), and the Mixed Estimator (Equation [13])

Parameter	Experimental value†	Bias (RMSE) ML	Extended A–B GMM	Mixed estimator
ρ	0.6230	– 0.0149 (0.0329)	0.3007 (0.0879)	– 0.0149 (0.0329)
τ	– 0.0634	– 0.0317 (0.0048)	– 0.0091 (0.0133)	– 0.0081 (0.0125)
η	0.0372	0.0181 (0.0080)	0.0318 (0.0201)	– 0.0029 (0.0190)
B_2	0.0316	– 0.0032 (0.0045)	– 0.0038 (0.0125)	– 0.0038 (0.0122)
B_4	– 0.0341	– 0.0006 (0.0072)	0.0017 (0.0196)	0.0022 (0.0224)

*Based on 1,000 replications.†Based on the parameter estimates in the last column of Table 2.

estimators based on 1,000 replications. Results reported in this table reveal that the bias in the coefficient τ of the lagged dependent variable, ΔY_{t-T}, when using Elhorst's (2003) ML estimator, is large and negative (– 0.0317), corresponding to the Nickell (1981) bias in a dynamic panel data model without endogenous interaction effects. As expected, this bias in τ is reduced when using the extended Arrelano and Bond GMM estimator (– 0.0091). By contrast, whereas the bias in the coefficient ρ of the spatially lagged dependent variable ΔWY_t when using Elhorst's ML estimator is small (– 0.0149),[13] it is large and unacceptable (0.3007) when using the extended Arrelano and Bond GMM estimator. Because the mixed estimator produces the smallest bias in both the parameters τ and ρ (– 0.0081 and – 0.0149, respectively), and because its properties (bias and RMSE) with respect to the other parameters are comparable with the extended Arrelano and Bond GMM estimator, we have used this mixed estimator to estimate the fixed effects spatial Solow–Swan model.

Discussion of the results

Results reported in Table 2 show that the explanatory power of the spatially extended growth-initial income-level regression increases by 0.24, on average. A necessary and sufficient condition for convergence is $\beta_1+\beta_3<0$. For all regression results summarized in Table 2, this condition is satisfied. Furthermore, the coefficient of the spatially lagged dependent variable is significantly different from zero. This result strongly supports hypothesis H_1, which states that the rate of economic growth of a particular economy is related to that of its neighbors. The econometric literature points out that if a relevant explanatory variable is omitted from a regression equation, the OLS estimator of the coefficients for the remaining variables is biased and inconsistent (Greene 2005, pp. 133–34). This also holds for the spatially lagged dependent variable and for the spatially lagged independent variables. The difference found between the speed of convergence in Tables 1 and 2 may shed more light on the size of this bias. For the single cross-section regression results summarized in Table 2, the speed of convergence is estimated at 1.20% per year. If spatial interaction effects are omitted, the speed of convergence reduces to 1.08% (see Table 1), which, with respect to 1.20%, is an underestimation of 10.2%. Similarly, we can calculate that the speed of convergence is overestimated by 4.4% when using TSCS data over five-year time spans. If fixed effects are added to the regressions, this percentage drops to 2.3%.

The speed of convergence of 1.20% obtained when using the data in a cross-section is somewhat lower than the speed of convergence of 1.5–1.7% found in Ertur and Koch (2007) and of 2.0% found in López-Bazo, Vayá, and Artis (2004). However, each of these two studies investigated convergence among a smaller group of economies (91 countries, and 108 regions across 12 EU countries, respectively). When pooling TSCS data over five-year time spans, the speed of convergence changes to 0.82% per year, and when using the panel data with fixed effects, it changes to 7.63%. The latter percentage is in line with Badinger, Müller, and Tondl (2004), who report a figure of 6.9% using a relatively large sample (194 regions across 14 EU countries).

The initial income-level variable in neighboring economies has a positive and significant effect for all three of the regressions whose results are reported in Table 2. When the coefficient of this variable is divided by the coefficient of the initial income-level variable in the economy itself, we can test our fourth hypothesis, which states that the steady-state position of a particular economy is related to *s*, *n*, *g*, and δ in neighboring economies due to indirect effects (see equation [8]). This calculation gives $0.0055/(-0.0104) = -0.53$ (*t*-value -3.40) for the first, -0.52 (-3.26) for the second, and -0.59 (-0.02) for the third regression reported in Table 2. These results imply that our fourth hypothesis is supported by the results from the first two regressions.

The variable $\ln[s/(n+g+\delta)]$ observed in neighboring economies has a positive but insignificant effect for all three of the regressions whose results are reported in Table 2. When the coefficient of this variable is divided by the coefficient of the initial income-level variable in the economy itself, we also obtain the impact of $\ln[s/(n+g+\delta)]$ observed in neighboring economies on the steady-state per capita income level. This calculation gives $0.0031/(-0.0104) = -0.30$ (*t*-value -1.53) for the first, -0.38 (-1.76) for the second, and 0.54 (0.01) for the third regression shown in Table 2. Overall, these results imply that our second and third hypotheses, which states that the rate of growth and the steady-state position of an economy are affected by *s*, *n*, *g*, and δ in its neighboring economies, are not supported by the data.

We also tested whether the unconstrained spatial Durbin model can be simplified to a spatial lag or a spatial error model. Because for all of the regression results reported in Table 2, the coefficient of either the initial income-level variable or the variable $\ln[s/(n+g+\delta)]$ observed in neighboring economies is significant, we can reject the spatial lag model. To determine whether the spatial error model is acceptable, we tested the two constraints $\beta_3 = -\rho\beta_1$ and $\beta_4 = -\rho\beta_1$. The first constraint is acceptable for the data for all regressions whose results are summarized in Table 2. The second constraint must be rejected for the first two regression specifications but not for the last specification. This implies that the spatial error model is consistent with the data for regressions with fixed effects but is unnecessarily restrictive if fixed effects are left aside.

Although the unconstrained spatial Durbin model appears to be a useful extension of the commonly used spatial lag and spatial error models, it is not so for the economic-theoretical model that laid the groundwork for this empirical model. In all of the regression results summarized in Table 2, one or more of the parameters α, φ, and γ take values outside the interval on which they are defined ($0<\alpha<1$, $0<\varphi<1$, and $1/\omega_{min}<\gamma<1$). This problem improves for the regression including fixed effects. In the latter case, only the parameter φ takes a value outside the interval on which it is defined, whereas the parameters α and γ take reasonable values. A possible explanation is that we only allowed spatial heterogeneity in the intercept. The performance of the model might further increase if the assumption that the speed of convergence is identical in all economies is relaxed by allowing spatial heterogeneity in the slope parameters too (Ertur and Koch 2007).

Conclusions

We found empirical evidence in favor of the hypotheses that the rate of growth of a particular economy is related to that of its neighbors (H_1) and that the steady-state position of a particular economy is related to *s*, *n*, *g*, and δ in its neighboring economies due to indirect effects (H_4). By contrast, we found no empirical evidence in favor of the hypotheses that the rate of growth of a particular economy is affected by *s*, *n*, *g*, and δ in its neighboring economies (H_2) and that the steady-state position of a particular economy and *s*, *n*, *g*, and δ in its neighboring economies have a direct relationship (H_3).

We also found that the speed of convergence when ignoring spatial interaction effects is biased but that this bias decreases by including fixed effects and by reducing the time interval over which the growth rate is measured. When pooling TSCS data over five-year time spans, the speed of convergence amounts to 0.82% per year and to 7.63% when using panel data with fixed effects. But the speed of convergence obtained from a model without fixed effects is not comparable to that of a model with fixed effects. Therefore, the strengths and weaknesses of fixed effects, put forward, respectively, by Islam (1995) and Durlauf and Quah (1999), can be maintained when spatial interaction effects are added to economic growth models.

Acknowledgements

Gianfranco Piras ackowledges financial support from the Nuclei of the Millennium Science Initiative Program ''Regional Science and Public Policy'' (MIDEPLAN, Chile).

Notes

1 In a cross-sectional setting, W should be a nonnegative matrix of known constants. The diagonal elements are set to zero by assumption. The matrix $I_N - \rho W$ should be nonsingular and its inverse be uniformly bounded. For a row-normalized W, this condition is satisfied as long as ρ is in the interior of $(1/\omega_{\min}, 1/\omega_{\max})$. Furthermore, the row and column sums of the matrices W and $(I_N - \rho W)^{-1}$ should be uniformly bounded in absolute value as N goes to infinity. Finally, the row and column sums of W before row normalization should not diverge to infinity at a rate equal to or faster than the rate of the sample size N. In a panel data setting, the latter condition is not relevant (but note Kelejian, Prucha, and Yuzefovich 2006).

2 Although Ertur and Koch (2007) find strong evidence supporting parameter heterogeneity, the upsurge in the use of heterogeneous panel data estimators has been criticized, both inside and outside of the growth literature (Quah 1996; Baltagi, Griffin, and Xiong 2000).

3 See Anselin (1988, pp. 226–99) for mathematical details.

4 Several studies have also split the variable $\ln[s/(n+g+\delta)]$ into $\ln(s)$ and $\ln(n+g+\delta)$ to test whether their coefficients are different. We do not report the results of these analyses, because they did not change our overall conclusions.

5 Usually, s, n, g, δ, and ε are not indexed in the convergence literature (e.g., Mankiw, Romer, and Weil 1992; Islam 1995). However, since s, n, g, and δ are measured as averages over the period $t-T$ to t, these variables are different for different time periods and therefore should be indexed by t. The same applies to ε.

6 This also holds for the least-squares dummy variables estimator, which first eliminates the intercept and the spatial fixed effects (see Baltagi 2005, pp. 135–36).

7 Note that first differencing a regression equation to eliminate spatial fixed effects (as well as the intercept) does not eliminate time-period fixed effects. Also note that the structure of these first-differenced time-period fixed effects is such that common time-period fixed effects can replace them.

8 Note that islands (such as those associated with southern European countries) are assumed to be connected to the mainland, so that each region has at least one neighbor. We have also tested the robustness of our results using spatial weights matrices based on the six nearest neighbors and on the 10 nearest neighbors. Although the parameter estimates of λ, α, φ, and γ appear to be sensitive to the spatial weights matrix used, none of the conclusions with respect to the four hypotheses discussed below changed (acceptance or rejection).

9 These routines are freely downloadable from http://www.spatial-econometrics.com and http://www.regroningen.nl/elhorst

10 To ease the notation, time-period fixed effects are left aside. Since their number is small, they can be added as additional explanatory variables, that is, as part of ΔX in equation (10).

11 Typically, one would take $d = 1$ or 2, dependent on the number of regressors. Lee (2003) introduced the optimal instrument 2SLS estimator, but Kelejian, Prucha, and Yuzefovich (2004) found that the 2SLS estimator based on this set of instruments has quite similar small sample properties.

12 To simulate the error term, we took $\varepsilon_{it} \sim N(0, \sigma^2)$, using the parameter estimate for σ^2, which was 0.00083. Similarly, to simulate the spatial and time-period fixed effects, we used the estimated values based on formulas (3.6) and (3.7) in Baltagi (2005).

13 A similar result was found in the Monte Carlo simulation experiment by Yu, de Jong, and Lee (2008). They experimented with τ, ρ, and $\eta = 0.2$ and 0.3 for $N = 49$ and 196, and $T = 10$ and 50, and found that the bias in ρ for this set of experimental parameter values does not exceed 0.0105 (absolute value).

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*. Dordrecht, The Netherlands: Kluwer.

Arbia, G. (2006). *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence*. Berlin: Springer.

Arrelano, M., and S. Bond. (1991). ''Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations.'' *Review of Economic Studies* 58, 277–97.

Badinger, H., W. G. Müller, and G. Tondl. (2004). ''Regional Convergence in the European Union (1985–1999): A Spatial Dynamic Panel Analysis.'' *Regional Studies* 38, 241–53.

Baltagi, B. H. (2005). *Econometric Analysis of Panel Data*, 3rd ed. Chichester, UK: Wiley.

Baltagi, B. H., J. M. Griffin, and W. Xiong. (2000). ''To Pool or Not to Pool: Homogeneous Versus Heterogeneous Estimators Applied to Cigarette Demand.'' *The Review of Economics and Statistics* 82, 117–26.

Barro, R. J., and X. Sala-i-Martin. (1995). *Economic Growth*. New York: McGraw-Hill.

Durlauf, S. N., and D. T. Quah. (1999). ''The New Empirics of Economic Growth.'' In *Handbook of Macroeconomics*, Vol. 1: 235–308, edited by J. B. Taylor and M. Woodford. Amsterdam: Elsevier.

Egger, P., and M. Pfaffermayr. (2006). ''Spatial Convergence.'' *Papers in Regional Science* 85, 199–215.

Elhorst, J. P. (2001). ''Dynamic Models in Space and Time.'' *Geographical Analysis* 33, 119–40.

Elhorst, J. P. (2003). ''Specification and Estimation of Spatial Panel Data Models.'' *International Regional Science Review* 26, 244–68.

Elhorst, J. P. (2010). ''Dynamic Panels with Endogenous Interaction Effects when T is Small.'' *Regional Science and Urban Economics*, in press. Available at http://dx.doi.org/10.1016/j.regsciurbeco.2010.03.003.

Ertur, C., and W. Koch. (2007). ''Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence.'' *Journal of Applied Econometrics* 22, 1033–62.

Fingleton, B. (2003). *European Regional Growth*. Berlin: Springer.

Fingleton, B., and E. López-Bazo. (2006). ''Empirical Growth Models with Spatial Effects.'' *Papers in Regional Science* 85, 177–98.

Greene, W. H. (2005). *Econometric Analysis*, 6th ed. Englewood Cliffs, NJ: Pearson Prentice Hall.

Islam, N. (1995). ''Growth Empirics: A Panel Data Approach.'' *Quarterly Journal of Economics* 110, 1127–70.

Kelejian, H. H., I. R. Prucha, and Y. Yuzefovich. (2004). ''Instrumental Variable Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances: Large and Small

Sample Results.'' In *Spatial and Spatiotemporal Econometrics*, 163–98, edited by J. P. LeSage and K. Pace. Amsterdam: Elsevier.

Kelejian, H. H., I. R. Prucha, and Y. Yuzefovich. (2006). ''Estimation Problems in Models with Spatial Weighting Matrices which have Blocks of Equal Elements.'' *Journal of Regional Science* 46, 507–15.

Lee, L.-F. (2003). ''Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances.'' *Econometric Reviews* 22, 307–35.

Lee, L.-F. (2004). ''Asymptotic Distribution of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models.'' *Econometrica* 72, 1899–925.

LeSage, J. P. (1999). Spatial Econometrics. Available at http://www.spatial-econometrics.com/html/sbook.pdf (accessed June 8, 2008).

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

Mankiw, N. G., D. Romer, and D. Weil. (1992). ''A Contribution to the Empirics of Economic Growth.'' *Quarterly Journal of Economics* 107, 407–37.

Nickell, S. (1981). ''Biases in Dynamic Models with Fixed Effects.'' *Econometrica* 49, 1417–26.

Quah, D. (1996). ''Aggregate and Regional Disaggregate Fluctuations.'' *Empirical Economics* 21, 137–62.

Revelli, F. (2001). ''Spatial Patterns in Local Taxation: Tax Mimicking or Error Mimicking?'' *Applied Economics* 33, 1101–07.

Rey, S. J., and M. J. Janikas. (2005). ''Regional Convergence, Inequality, and Space.'' *Journal of Economic Geography* 5, 155–76.

Yu, J., R. de Jong, and L. Lee. (2008). ''Quasi-Maximum Likelihood Estimators for Spatial Dynamic Panel Data with Fixed Effects When Both n and T are Large.'' *Journal of Econometrics* 146, 118–34.