Spatial Convergence of Regions Revisited: A Spatial Maximum Likelihood Systems Approach

Michael Pfaffermayr*

11th February 2009

Abstract

This paper suggests that one should account for the persistence of technology shocks and the endogeneity of important explanatory variables when analyzing spatial convergence among regions. In a cross-section a systems approach relating the average growth rate to the initial income level in a two equations framework is called for. In an empirical study of growth and convergence of 212 European regions, the estimation results for the period 1980-2002 reveal a substantial correlation between the disturbances of the equation explaining initial income per capita and that of its subsequent average growth rate. Moreover, the estimated speed of convergence is found substantially higher in a systems framework. This holds true for both spatial conditional and unconditional convergence.

Keywords: Spatial β-convergence; Spatial Solow model; Spatial systems maximum likelihood estimation; European regions.

JEL: R11; C31; O47

*Department of Economics, University of Innsbruck, Universitaetsstrasse 15, 6020 Innsbruck, Austria, Austrian Institute of Economic Research, Vienna, Austria and IFO, Germany. Financial support from the Austrian Science Foundation grant 17028 is gratefully acknowledged. Email: Michael.Pfaffermayr@uibk.ac.at.

1 Introduction

The estimated convergence parameters and their implied speed of convergence in the studies on regional growth based on the spatial Solow model are surprisingly low, even among European regions that are apparently more integrated than a group of countries. On the theoretical side this may reflect the ignorance of capital mobility and migration. Essentially, this model treats regions as closed units that are only interrelated by non-pecuniary spillovers arising from learning in the course of capital accumulation. Empirical research based on this model so far seems to ignore the endogeneity of important explanatory variables like the investment ratio as proxy of the savings rate and the growth rate of population in case of capital and/or labor mobility. While this may have been an acceptable assumption for empirical work at the country level, for regions this assumption is not plausible as mobility of factors is partly among regions belonging to the same country.

On the empirical side, the finding of low convergence speeds might also originate from misspecified convergence equations which fail to account for the persistence of technology shocks. Available evidence of growth equations derived from the Solow model at the country level by Caselli et al. (1997) and McQuinn and Whelan (2007) supports this view.[1]

For empirical research on regional convergence these two issues imply that the commonly used specification originating from Mankiw, Romer and Weil (1992) and reformulated in a spatial setting with regional knowledge spillovers by Ertur and Koch (2007) and Pfaffermayr (2009) may be misspecified. This specification introduces the investment to GDP ratio as a proxy of

[1] In a somewhat different approach Goetz and Hu (1996) show that inference on the convergence coefficient can be misleading if the population growth rate and savings rates are themselves functions of income.

the savings rate and the rate of population growth as explanatory variables that account for steady state determinants of income per capita. However, both variables are clearly endogenous under factor mobility. Furthermore, the possible persistence in technology shocks suggests that initial income should be treated as an endogenous variable, too. In cross-section studies this latter issue calls for a spatial systems approach relating the average growth rate of income per capita to its initial level in a spatial simultaneous two equations framework.

Accounting for these endogeneity issues, the estimation results for European regions for the period 1980-2002 indicate substantial correlation between disturbances of the initial income per capita equation and the convergence equation that explains average income growth. Moreover, the estimated speed of convergence is found substantially higher in the spatial systems framework as compared to the available estimates of spatial convergence. This holds true for both conditional and unconditional convergence.

2 Regional growth under knowledge spillovers and factor mobility

The spatial Solow model is based on a constant returns to scale production function, exhibiting a diminishing marginal product of capital and spatial knowledge spillovers across regions. The knowledge spillovers are modelled as pure externalities, which can either be local or global in nature (see Anselin, 2003). These knowledge spillovers originate form learning effects in the course of capital accumulation assuming that knowledge is embodied in capital interpreted in a broad sense. Here, we concentrate on local knowledge spillovers, since the immediate diffusion of knowledge among all regions as implied by global knowledge spillovers seems less plausible. Rather, it is assumed that initially knowledge spillovers are local by nature, but develop into global ones over time as knowledge diffuses. This implies that the spatially weighted capital labor ratio enters the production function so that total factor productivity is higher for those regions that are surrounded by rich neighbor regions.[2] Formally, the macroeconomic production function of region i under local knowledge spillovers is given by

$$Y_i = C_i \left(\frac{K_i}{L_i A_i}\right)^{\phi} \left(\Pi_{j \neq i} \left(\frac{K_j}{L_j A_j}\right)^{\rho w_{ij}}\right) K_i^{\varphi} \left[L_i A_i\right]^{1-\varphi} \tag{1}$$

where Y_i denotes output, K_i stands for the broad measure of (physical and human) capital and L_i for labor. A_i denotes the state of labor augmenting technological progress. Knowledge spillovers are assumed to exhibit a spatial decay represented by spatial weights w_{ij} with $w_{ii} = 0$. These spatial weights are either based on contiguity or inversely related to some measure of distance. The term $\left(\frac{K_{il}}{L_i A_i}\right)^{\phi}$ captures intra-region spillovers (see Pfaffermayr, 2009 for more details). Lastly, C_i is a normalizing constant. Convergence will be observed under diminishing returns to scale, which will occur if $1 - \varphi - \phi - \rho > 0$ (see the Appendix). Below, the spatial weights are collected in the $(N \times N)$ spatial weighting matrix $\mathbf{W}$, which may be row sum normalized or maximum row sum normalized. Furthermore, the natural log of $\frac{K_i}{L_i A_i}$ is denoted by k_i. Similarly, y_i is the natural log of $\frac{Y_i}{L_i A_i}$.

In almost all empirical studies on regional growth and convergence population growth has been taken as an exogenous determinant of the steady state level of income per worker. At the regional level this assumption is

[2]For example, Chua (1993), Egger and Pfaffermayr (2006), López-Bazo et al. (2004), Moreno and Trehan (1997), Pfaffermayr (2009) and Vayá et al. (2004) assume local knowledge spillovers.

not plausible, however, as barriers to migration are low within countries or between regions in an integrated area such as Europe. To derive an empirical specification of the spatial income convergence equation that accounts for labor mobility among regions, Barro and Sala-i-Martin (2004, p. 384) and Braun (1993) augment the Solow model by a net-migration function, but assume that economies are closed in any other respect.[3]

To illustrate the impact of migration in an exogenous spatial growth model let us assume that migration is partially restricted in the sense that migration costs are high enough to render migration unattractive in the steady state. Hence, migration is a transitory phenomenon and extreme agglomerations are ruled out. In the presence of migration population growth in a region is composed of the natural population growth rate as determined by fertility net of mortality (n), which is assumed constant across regions, and the net-migration rate. The latter is captured by the net-migration function $\xi(k_1, ..., k_N)$. In the spirit of Barro and Sali-i-Martin (2004) and Faini (1996, eq. 8) one can approximate the net migration function by

$$\xi(k_1, ..., k_N) \approx \varepsilon \left(k_i - k_i^* - \sum_{j=1}^{N} w_{ij} \left(k_j - k_j^* \right) \right) \tag{2}$$

where ε denotes a constant scalar capturing the sensitivity of the willingness to migrate with respect to wage differentials, which in turn depend on regional differences in the capital to efficiency units of labor.[4] Then the law

[3]Faini (1996) among others analyzes the impact of migration on income growth in a more general model. With constant returns there is unconditional convergence and migration increases the speed of convergence. However, the model implies divergence if economies of scale are strong and labor is sufficiently mobile.

[4]Following Faini (1996) and others one can motivate this specification by the assumptions that individuals differ in their preferences and that they choose their destination region in a random utility framework (see also Fields, 1979), accounting for wage differentials and distance depending migration costs. Then, each region receives a net-migration

of motion of capital per efficient worker is given by

$$\dot{k}_i = s_i e^{z_i} - g - \varepsilon \left(k_i - k_i^* - \sum_{j=1}^{N} w_{ij} \left(k_j - k_j^* \right) \right), \tag{3}$$

where $g = x + \gamma + n$, $z_i = \sum_{j=1}^{N} \rho w_{ij} k_j + c_i + (\alpha - 1) k_i$ and $\alpha = \varphi + \phi$. In the steady state one obtains $\sum_{j=1}^{N} \rho w_{ij} k_j^* + c_i + (\alpha - 1) k_i^* = \ln(\frac{g}{s_i})$ and the augmented Solow equation with local knowledge spillovers reads

$$\dot{k}_i = s_i e^{z_i} - g - \varepsilon \left(k_i - k_i^* - \sum_{j=1}^{N} w_{ij} \left(k_j - k_j^* \right) \right). \tag{4}$$

Linearizing $s_i e^{z_i}$ around the steady state with $s_i e^{z_i^*} = g$ yields

$$\begin{aligned} \dot{k}_i &= g + s_i e^{z_i^*} \sum_{j=1}^{N} \left. \frac{\partial z_i}{\partial k_j} \right|_{\mathbf{k}=\mathbf{k}^*} \cdot (k_j - k_j^*) - g - \varepsilon \left(k_i - k_i^* - \sum_{j=1}^{N} w_{ij} \left(k_j - k_j^* \right) \right) \\ &= g \sum_{j=1}^{N} \left. \frac{\partial z_i}{\partial k_j} \right|_{\mathbf{k}=\mathbf{k}^*} (k_j - k_j^*) - \varepsilon (k_i - k_i^*) + \varepsilon \sum_{j=1}^{N} w_{ij} (k_j - k_j^*). \end{aligned}$$

Using

$$\left. \frac{\partial \mathbf{z}}{\partial \mathbf{k}'} \right|_{\mathbf{k}=\mathbf{k}^*} = \rho \mathbf{W} + (\alpha - 1) \mathbf{I}$$

the law of motion can be compactly described as

$$\dot{\mathbf{k}} - \dot{\mathbf{k}}^* = \beta_M \mathbf{B}_M (\mathbf{k} - \mathbf{k}^*), \tag{5}$$

with $\mathbf{B}_M = (\mathbf{I} - \theta_M \mathbf{W})$, $\beta_M = -((1-\alpha)g + \varepsilon)$ and $\theta_M = \frac{g\rho + \varepsilon}{(1-\alpha)g + \varepsilon}$. In line with the literature, net-migration speeds up convergence in the absence of knowledge spillovers (β_M). However, migration also affects the spillover

flow from every other region.

parameter θ_M and enhances the spillovers across regions. In every other respect the specification of the spatial convergence equation is similar to the available ones in the literature.

Obstfeld and Rogoff (1996) and Barro and Sali-i-Martin (2004) demonstrate in a model without knowledge spillovers that it is possible to also cover the case of mobile capital in the Solow model as long as one accumulated factor (e.g., human capital) is immobile and accumulated. These authors introduce a broad concept capital including human and physical capital and derive essentially the same specification results when regions initially start their growth process from below the steady state. The main assumption is that countries or regions are debt constraint as only physical capital can be used as a collateral. If the initial stock of capital is smaller than its steady state counterpart and convergence is from below, capital inflows into or outflows from a region are proportional to the existing stock of capital. In a similar model with labor and broad capital as production factors, Cohen and Sachs (1986) and Escot and Galindo (2000) obtain similar results.

In a similar vein, one may assume that small regions can borrow or lend in a fixed proportion to their stock of capital in international capital markets at a fixed interest rate r^*. Their net debt is given by $B_i = m_i K_i$ with $m_i = \widetilde{m} > 0$ if $k_i^* > k_i(0) - b_i(0)$ and $m_i = -\widetilde{m} < 0$ if $k_i^* > k_i(0) - b_i(0)$, where $b_i(0)$ denotes B_i/L_iA_i at time 0. At $m_i > 0$ a region borrows from the world capital market, otherwise it lends to world capital market (see Escot and Galindo, 2000).[5] In both cases it is assumed that this amounts to a constant small enough fraction of a region's existing capital stock and

[5]This specification differs from Barro et al. (1995) who impose the borrowing condition in terms of the stock of capital but not in terms of the capital to efficiency units of labor ratio. If $K_i(0) + B_i(0) \geq K_i^*$ the borrowing constraint is not binding and region i would immediately jump to its steady state value of k_i, see Barro et al. (1995, p.110).

that the capital constraint is always binding. The basic law of motion is then generalized to

$$\begin{aligned} \dot{K}_i &= s_i(Y_i - r^* B_i) + \delta K_i + \dot{B}_i \qquad (6) \\ &= s_i(Y_i - r^* m_i K_i) + \delta K_i + m_i \dot{K}_i \end{aligned}$$

Measured in efficiency units of labor this law of motion now reads

$$\dot{k}_i = \tfrac{s_i}{(1-m)}(e^{z_i} - r^* m_i) - (x + n_i + \tfrac{\delta}{1-m_i}). \qquad (7)$$

Denoting $g_i = x + n_i + \frac{s_i r^* m_i + \delta}{(1-m_i)}$ and linearizing around $\bar{k}_i$ defined implicitly by $e^{z_i} = g = x + \bar{n} + \frac{\bar{s} r^* \bar{m} + \delta}{(1-\bar{m})}$ (see Pfaffermayr, 2009 for details) yields the linearization

$$\tfrac{dk_i}{dt} = g \sum_{j=1}^{N} \tfrac{\partial z_i}{\partial k_j}\Big|_{\mathbf{k}=\mathbf{k}^*} (k_j - \bar{k}_j) - (g_i - g), \qquad (8)$$

where

$$\tfrac{\partial \mathbf{z}}{\partial \mathbf{k}'}\Big|_{\mathbf{k}=\mathbf{k}^*} = \rho \mathbf{W} + (\alpha - 1)\,\mathbf{I}.$$

In vector form, the law of motion is then given by

$$\dot{\mathbf{k}} - \dot{\mathbf{k}}^* = \beta_K \mathbf{B}_K (\mathbf{k} - \mathbf{k}^*) + \Delta \mathbf{g}_K \qquad (9)$$

with $\mathbf{B}_K = (\mathbf{I} - \theta_K \mathbf{W})$, $\Delta \mathbf{g}_K = \mathbf{g}_i - \mathbf{g}$, $\beta_K = -(1-\alpha)g$ and $\theta_K = \frac{\rho}{1-\alpha}$. This specification likewise implies that the basic structure of the spatial Solow model remains the same under this form of capital mobility. The main difference is that capital mobility increases the speed of convergence, since $\frac{\partial g}{\partial \bar{m}} > 0$.

So far, capital mobility and net-migration seem to be neglected in the empirical literature on regional growth and convergence. The spatial Solow model serving as the workhorse model to analyze regional convergence in many studies exhibits enough economic structure to analyze regional convergence processes under regional knowledge spillovers and (to a limited extent) factor mobility. The presumption is that regions are not too far away from their steady state in the initial period considered in the empirical application and that factor mobility is restricted in some way.

One can easily show that for output per efficient worker (denoted by $\mathbf{q}$, see the Appendix) the same law of motion holds. To obtain the empirical convergence equation, one can solve the system of first order differential equations. However, due to the knowledge spillovers regions cannot be treated as independent units. Rather the full system of differential equations that describes the law of motion of income per efficient worker (see Egger and Pfaffermayr, 2006) has to be solved.

As shown in the Appendix under local knowledge spillovers and labor mobility the convergence equation is given by

$$\mathbf{q}_t - \mathbf{q}_{t-\tau} = \mathbf{a}_t - \mathbf{a}_{t-\tau} - \left(\mathbf{I} - e^{\beta_M \mathbf{B}_M \tau}\right)(\mathbf{q}_{t-\tau} - \mathbf{q}^*), \tag{10}$$

where $\mathbf{a}_t$ denotes labor saving technical progress and $e^{\beta_M \mathbf{B}_M t} = \mathbf{I} + \beta_M \mathbf{B}_M t + \beta_M^2 \mathbf{B}_M^2 \frac{t^2}{2!} + ...$, with $\mathbf{B}_M$ defined above. In case we allow for capital mobility, the convergence equation reads

$$\mathbf{q}_t - \mathbf{q}_{t-\tau} \approx \mathbf{a}_t - \mathbf{a}_{t-\tau} - \left(\mathbf{I} - e^{\beta_K \mathbf{B}_K \tau}\right)(\mathbf{q}_{t-\tau} - \mathbf{q}^*) + \tau\left(\alpha \mathbf{I} + \rho \mathbf{W}\right) \mathbf{\Delta g}_K. \tag{11}$$

Without knowledge spillovers, heterogeneity in n and the possibility of labor migration or capital mobility, i.e. $\rho = 0$ and $\varepsilon = 0$ or $m_i = 0$, the spatial convergence equation collapses to the well known β-convergence equation as derived by e.g. McQuinn and Whelan (2007), since in this case $\mathbf{B}_l = \mathbf{I}$ and $\beta_l = -(1-\alpha)(x+n+\delta)$, $l = M, K$ and

$$\mathbf{q}_t - \mathbf{q}_{t-\tau} = \mathbf{a}_t - \mathbf{a}_{t-\tau} - \left(1 - e^{\beta_L \tau}\right)(\mathbf{q}_{t-\tau} - \mathbf{q}^*). \tag{12}$$

3 The econometric specification

Similar to McQuinn and Whelan (2007) labor augmenting technological progress in each region is assumed to follow a common deterministic trend. Its stochastic component is subject to region specific technology shocks that are spatially correlated and exhibit some persistence over time. Specifically it is postulated that

$$\begin{aligned} \mathbf{a}_t &= at\mathbf{e} + (\mathbf{I} - \phi\mathbf{W})^{-1}\mathbf{v}_t \\ \mathbf{v}_t &= \delta\mathbf{v}_{t-1} + \varepsilon_t, \end{aligned} \tag{13}$$

where ε_t is a vector of iid normal random variables and $|\delta| < 1$. As shown in the Appendix, it follows that

$$\begin{aligned} \mathbf{a}_{t-\tau} &= a(t-\tau)\mathbf{e} + (\mathbf{I} - \phi\mathbf{W})^{-1}_{t-\tau}\xi_{t-\tau} \\ \frac{1}{\tau}(\mathbf{a}_t - \mathbf{a}_{t-\tau}) &= a\mathbf{e} + (\mathbf{I} - \phi\mathbf{W})^{-1}_t\xi_t, \end{aligned}$$

where $\xi_t = \frac{1}{\tau}(\delta^\tau - 1)\mathbf{v}_{t-\tau} + \frac{1}{\tau}\sum_{m=0}^{\tau-1}\delta^m\varepsilon_{t-m}$ and $\xi_{t-\tau} = \delta^{t-\tau}\mathbf{v}_0 + \sum_{m=0}^{t-\tau-1}\delta^m\varepsilon_{t-\tau-m}$. The latter two terms denote time aggregated error terms that are independent across units i, but correlated across time for a given i. Their variances are denoted by σ_t^2 and $\sigma_{t-\tau}^2$, respectively, and their covariance is described by the parameter $\sigma_{t,t-\tau}$. It can immediately be seen that $\sigma_{t,t-\tau} \neq 0$, if technology shocks are persistent and $\delta \neq 0$. Note, in the empirical implementation the spatial correlation coefficient ϕ is allowed to differ between $t-\tau$ and t to account for different spatial correlation in the disturbances referring to the starting values. To simplify notation, we set $\tau = t$ to obtain the following empirical specification that is estimated below:

$$\begin{aligned} \mathbf{q}_0 &= \mathbf{X}_0\delta_0 + (\mathbf{I} - \phi_0\mathbf{W})^{-1}\xi_0 \\ \tfrac{1}{t}(\mathbf{q}_t - \mathbf{q}_0) &= \tfrac{1}{t}\left(\mathbf{I} - e^{\beta_l\mathbf{B}_l t}\right)\mathbf{q}_0 + \mathbf{X}_t\delta_t + (\mathbf{I} - \phi_t\mathbf{W})_t^{-1}\xi_t \\ & \xi_{it}, \xi_{i0} \sim N\left(\mathbf{0}, \begin{bmatrix} \sigma_0^2 & \sigma_{t0} \\ \sigma_{t0} & \sigma_t^2 \end{bmatrix}\right). \end{aligned} \tag{14}$$

For small absolute values of β_l one can use the following approximation to obtain a linear specification:

$$\tfrac{1}{t}\left(\mathbf{I} - e^{\beta_l\mathbf{B}_l t}\right) \approx \beta_l\mathbf{B}_l = \beta_l\mathbf{I} + \gamma_l\mathbf{W}, \tag{15}$$

where $\gamma_l = -\beta_l\theta_l$ and one expects $\beta_l < 0$ and $\gamma_l > 0$. For the general case, a non-linear series estimator can be applied that uses the first k elements of $e^{\beta_l\mathbf{B}_l t} = \mathbf{I} + \beta_l\mathbf{B}_l t + \beta_l^2\mathbf{B}_l^2\frac{t^2}{2!} + ...$(see Kelejian, Prucha and Yuzefovich, 2004).

In the presence of (restricted) factor mobility the econometric specification of convergence equations has to be re-established. Important explanatory variables considered in the literature, notably the investment to output ratio as a proxy of the savings rate and the population growth rate, are clearly endogenous under these more realistic assumptions.

The matrix $\mathbf{X}_0$ comprises the systematic determinants of the starting conditions as captured by $\mathbf{q}_0$ and $\mathbf{a}_0$. Below, these are modelled by a spatial trend defined by the longitude and latitude of a region's center (see, Haining, 2003) and country fixed effects. The design matrix of the convergence equation, $\mathbf{X}_t$, includes steady state determinants of $\mathbf{q}^*$ and that of the trend growth rate of technological progress represented by the systematic part of $\frac{1}{t}(\mathbf{a}_t - \mathbf{a}_0)$. Under conditional convergence it is assumed that country specific dummies capture all these systematic effects, implying for example country specific savings rates. The approximation error $\mathbf{\Delta g}_K$ is subsumed under the disturbances. In this case, one observes conditional convergence across countries, but unconditional convergence within countries. In the unconditional specification of the convergence equation includes a constant only. The corresponding likelihood of this system of equations is derived in the Appendix. It will be maximized numerically.

4 Data and estimation results

The regional income data come from Cambridge Econometrics and comprise information on 212 regions of the EU15 and the regions of Switzerland and Norway. The observations for 10 regions in the former German Democratic Republic are not available for 1980 and these regions are therefore excluded. Also, the Portuguese islands Azores have been skipped because of their very large distance from the European continent.

The proposed model of spatial convergence is applied to investigate the evolution of real income per capita of European NUTS II regions over the period 1980-2002. The dependent variable is the average log difference of real GDP per capita scaled by the factor 100, where a regions's working population is used in the denominator of this figure. The elements of the spatial weighting matrix $\mathbf{W}$ are given by $w_{ij} = e^{-d_{ij}/c} / \sum_{j=1}^{N} e^{-d_{ij}/c}$, where

d_{ij} denotes the distance between the centers of regions i and j and c defines the spatial decay. The decay parameter cannot be estimated and to assess the robustness of the estimation results, it takes the values 50, 100 and 150. The preferred model will be selected by comparing the estimated likelihood. Since the number of parameters is always the same for each model, this approach is equivalent to applying model selection criteria like AIC or BIC. Note this specification implies that $\mathbf{W}$ is row normalized.

Table 1 exhibits the estimation results of the linear specification of the system for the model with country specific effects in both estimated equations and, therefore, deals with conditional convergence. For the preferred model the decay parameter c takes the value 150. The country dummies are highly significant in both estimated equations of the system, supporting conditional convergence across countries, but unconditional convergence within countries. The presence of country dummies seems to wipe out spatial correlation in error term which has been found significant in many previous studies. The corresponding parameter estimates turn out insignificant in both the initial income equation and in the convergence equation under the systems specification with correlation of the disturbances across equations (see Attfield et al., 2000 for a similar result). In the initial income equation the regions' longitude is significant even after controlling for country effects, implying an increasing initial income gap when moving from east to west within the average country.

The estimated convergence parameter with exogenous initial income at $\rho = 0$ and $c = 100$ amounts to -0.012 indicating slow income convergence in the absence of knowledge spillovers. This estimate is similar to the available ones in the literature (see e.g. Armstrong, 1995; Neven and Guyette 1995; Carrington, 2003; López-Bazo et al., 2004; Le Gallo and Dall'erba, 2006 and Pfaffermayr, 2009. Abreu, De Groot, and Florax, 2005 and Fingleton and López-Bazo, 2006 provide a comprehensive survey on available studies). Assuming reasonable values for the generalized depreciation rate, $x + n + \delta$, this estimate implies an implausible high value of the capital share even when capital is interpreted in broad terms. For example, setting $x + \delta = 0.07$ (slightly lower than McQhinn and Whelan, 2007) and n to the sample average of 0.005 yield a capital share of 84 percent. However, this result also suggests that the possible endogeneity of the investment share as a measure of the savings rate and of population growth rate does not seem to be a major source of bias. The estimated speed of convergence is comparable to previous estimates that are based on the structural Solow model with exogenous initial income (López-Bazo et al., 2004, Fingleton and López-Bazo, 2006).

In line with previous work there are significant knowledge spillovers as indicated by the significant positive impact of the spatially lagged initial income. Regions with initially rich neighbors have more potential to learn from their neighbors and, therefore, tend to grow faster on average ceteris paribus.

*** Tables 1 and 2 ***

The estimation results in Table 1 furthermore suggest that the initial income levels are indeed best treated as endogenous variable in the convergence equation. The system estimates reveal a significant correlation of the error terms of the two estimated equations amounting to 0.82 in the preferred specification. The estimates of the convergence parameters broadly confirm the findings at the country level without knowledge spillovers reported by McQuinn and Whelan (2007). The persistence of technology shocks and the resulting endogeneity of initial income lead to a substantial underestimation of the convergence parameter when initial income is treated as an exogenous variable. The convergence parameter now amounts to -0.044 for the system estimates as compared to -0.012 when treating initial income exogenously. Under the parameter values assumed above this would imply a more plausible capital share of 41 percent. In addition, the spillover parameter also tends to be downward biased under exogenously treated initial income.

The findings for unconditional convergence in Table 2 are very similar, indicating a convergence parameter of -0.057, but also a much higher spillover parameter amounting to 0.048. Here the model with decay parameter $c = 100$ is the preferred one. With exogenous initial income the corresponding estimates are -0.011 and 0.006, respectively. The results for the non-linear series estimator in Table 3 confirm these results, although the convergence parameter for the specification with exogenous initial income turns out somewhat higher, while that with endogenous initial income is a bit lower. Overall, the linear approximation seems accurate enough for estimating spatial convergence equations.

*** Table 3 ***

As argued in Egger and Pfaffermayr (2006) and Pfaffermayr (2009) the implied speed of convergence is typically region specific in the spatial Solow model with knowledge spillovers. Furthermore, the convergence speed can only be inferred from an experiment of thought, since the steady state income level of the regions and, hence, the income gap, remain unobserved. For illustration let us assume that the initial gap in income per capita with respect to the steady state is given by $q_i^* = 1.2 mean(\mathbf{q}_0)$. The speed of convergence is measured as the share of the gap in income per capita that is closed within a year on average. Table 4 calculates this figure for the preferred models of Tables 1 and 2 according to the formula

$$\psi_l = \frac{1}{t} Diag[\mathbf{q}_{l0}^* - \mathbf{q}_{l0}]^{-1}[\left(\mathbf{I} - e^{\beta_l \mathbf{B}_l t}\right)(\mathbf{q}_{l0}^* - \mathbf{q}_{l0})] \tag{16}$$

at $t = 1, 10$ and 50.[6] Under regional knowledge spillovers a region's speed of convergence depends on the strength of knowledge spillovers and the initial income gap of its neighbors, besides the convergence coefficient β_l. The higher the absolute value of the initial income gap of a region's neighbors is, the more the region can learn from its neighbors and the higher are the spatial spillovers. If the income gaps are positive on average and convergence mostly occurs from below, ignoring regional knowledge spillovers leads to an overestimation of the convergence speed (see also Pfaffermayr, 2009).

For illustration first consider the case where the initial income gap is the same for all regions, i.e., $\mathbf{q}_{l0}^* - \mathbf{q}_{l0} = \psi \mathbf{e}$, where $\mathbf{e}$ is a vector of ones ψ is a constant. Under this assumption there is no need to refer to the hypothetical steady state defined above. One can easily show that in this case one obtains

$$\psi_l = \frac{1}{t}\left(\mathbf{I} - e^{\beta_l \mathbf{B}_l t}\right)\mathbf{e} = \frac{1}{t}\left(1 - e^{(\beta_l - \beta_l \theta_l)t}\right)\mathbf{e} \approx -\left(\beta_l - \beta_l \theta_l\right)\mathbf{e},$$

using the fact that under row normalization we have $\mathbf{We} = \mathbf{e}$. Remember, $\beta_l < 0$ and $\theta_l > 0$. The estimated value of $-(\beta_l - \beta\theta_l)$ is considerably biased downwards when treating initial income as exogenous as shown in the first column of Table 4. The second column calculates the estimated values of $\frac{1}{t}\left(1 - e^{(\beta_l - \beta\theta_l)t}\right)$ and gives similar results. However, it nicely illustrates that

[6] The matrix exponential is calculated using the Matlab procedure expm.

the convergence speed decreases over time and that it is higher for the rich regions that converge from above.[7] Lastly, when the initial gap is assumed to be heterogenous there is considerable variation of the convergence rates. Yet, the main conclusion, namely that the convergence rates are considerably higher on average when initial income is treated endogenously, also shows up here. Under conditional convergence the speed of convergence for the first year is 2.93 percent per year and it falls to 1.55 percent when averaged over 50 years. The standard deviations of these figures are 3.01 and 1.67, respectively. Taking initial income as an exogenous variable implies an average convergence speed in the first year as low as 0.24 in contrast. Inserting the estimated parameters of the models for unconditional convergence, still leads to an average convergence speed of 1.16 on average at $t = 1$ as opposed to 0.61 under exogenous initial income.

*** Table 4 ***

To sum up, the estimated speed of convergence among European regions is substantially underestimated under persistent technology shocks that lead to endogeneity of the level of initial income.

5 Conclusions

This paper reconsiders the spatial Solow model of regional growth under local knowledge spillovers. It argues that in the presence of factor mobility and persistent technology shocks the widely used spatially augmented Mankiw, Romer and Weil (1992) specification to estimate β-convergence is

[7] See Mathunjwa and Temple (2007) for the opposite finding for the Solow model without spillovers.

prone to endogeneity of its most important explanatory variables. First, under factor mobility the investment share as proxy of the savings rate and the growth rate of population growth are clearly endogenous. Instead of using these explanatory variables the present approach suggests including country fixed effects so that regional convergence is conditional across countries and unconditional within countries. Second, and more importantly, under persistent technology shocks the initial income is endogenous in the convergence equation.

Using a bivariate spatial systems approach, this paper shows for regional income data comprising 212 regions of the EU15 and the regions of Switzerland and Norway that initial income is indeed endogenous. The error terms of the estimated initial income equation and that of the convergence equation are highly and significantly correlated rendering a single equation approach to measure β-convergence biased. The system estimates suggest that the speed of convergence in income per capita is considerably higher under this more general approach than previously found.

References

Abreu, Maria, Henri L.F. De Groot, and Raymond J.G.M. Florax (2005), Space and Growth, Région et Développement 21, pp. 13-44.

Anselin, Luc (1988), Spatial Econometrics: Methods and Models, London and Dordrecht: Kluwer.

Anselin, Luc (2003), Spatial Externalities, Spatial Multipliers and Spatial Econometrics, International Regional Science Review 26 (2), pp. 153-166.

Armstrong, Harvey W. (1995), Convergence among Regions of the European Union, 1950-1990, Papers in Regional Science 74 (2) , pp. 143-52.

Attfield, Clifford, Edmund S. Cannon, D. Demery, and Nigel W. Duck (2000), Economic Growth and Geographic Proximity, Economics Letters, 68(1), pp. 109-112.

Barro, Robert J., Gregory N. Mankiw and Xavier Sala-i-Martin (1995), Capital Mobility in Neoclassical Models of Growth, American Economic Review 85(1), pp. 103-115.

Barro, Robert and Xavier Sala-i-Martin (2004), Economic Growth, 2nd ed., MIT-Press.

Braun, Juan (1993), Essays on Economic Growth and Migration, PhD-Thesis Harvard University.

Carrington, Anca (2003), A Divided Europe? Regional Convergence and Neighborhood Spillover Effects, Kyklos 56 (3), pp. 381-394.

Caselli, Francesco, Gerardo Esquivel and Fernando Lefort (1997), Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics, Journal of Economic Growth 1(3), pp. 363-389.

Cohen, Daniel and Jeffrey Sachs (1986), Growth and External Debt under Risk of Debt Repudiation, European Economic Review 30(3), pp. 529-560.

Chua, Hak.B. (1993), On Spillovers and Convergence, PhD-Thesis, Harvard University.

Escot, Lorenzo and Miguel-Angel Galindo (2000), International Capital Flows and Convergence in the Neoclassical Growth Model, International Advances in Economic Research 6 (3), pp. 451-460.

Egger Peter and Michael Pfaffermayr (2006), Spatial Convergence, Papers in Regional Science, Papers in Regional Science 85(2), pp. 199-215.

Ertur, Cem and Wilfried Koch (2007), Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence, Journal of Applied Econometrics 22(6), pp. 1033-1062.

Faini, Riccardo (1996), Increasing Returns, Migration and Convergence, Journal of Development Economics, 49 (1), pp. 121-136.

Fields, Gary S. (1979), Place-to-Place Migration: Some New Evidence, Review of Economics and Statistics 61(1), pp. 21-32.

Fingleton, Bernard and Enrique López-Bazo (2006), Empirical Growth Models with Spatial Effect, Papers in Regional Science 85(2), pp. 177-198.

Goetz, Stephan J. and Dayuan Hu (1996), Economic Growth and Human Capital Accumulation: Simultaneity and Expanded Convergence Tests, Economics Letters 51(3), pp. 355-362.

Haining, Robert (2003), Spatial Analysis: Theory and Practice, Cambridge University Press, Cambridge UK.

Horn, R. and C. Johnson, 1985. Matrix Analysis. Cambridge University Press, Cambridge.

Kelejian, Harry H., Ingmar R. Prucha and Yevgeny Yuzefovich (2004), Instrumental Variable Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances: Large and Small Sample Results, in James P. LeSage and R. Kelly Pace, Spatial and Spatiotemporal Econometrics, Advances in Econometrics 18.

Le Gallo Julie and Sandy Dall'erba (2006), Evaluating the Temporal and Spatial Heterogeneity of the European Convergence Process, 1980-1999, Journal of Regional Science 46(2), pp. 269-288.

López-Bazo, Enrique, Esther Vayá, and Manuel Artis (2004), Regional Externalities and Growth: Evidence from European Regions, Journal of Regional Science 44 (1), pp. 43-73.

Mankiw, N. Gregory, David Romer and David N. Weil (1992), A Contribution to the Empirics of Economic Growth, Quarterly Journal of Economics 107(2), pp. 407-437.

Mathunjwa Jochonia S. and Jonathan R. W. Temple (2007), Convergence Behavior in Exogenous Growth Models, University of Bristol working paper.

McQuinn, Kieran and Karl Whelan (2007), Conditional Convergence and the Dynamics of the Capital-Output Ratio, Journal of Economic Growth 12(2), pp. 159-184.

Moreno, Ramon and Bharat Trehan (1997), Location and the Growth of Nations, Journal of Economic Growth 2 (4), pp. 399-418

Neven, Damien and Gouyette Claudine (1995), Regional Convergence in the European Community, Journal of Common Market Studies 33 (1), pp. 47-65.

Obstfeld, Maurice and Kenneth Rogoff (1996), Foundations of International Macroeconomics, MIT-Press, Cambridge MA.

Pfaffermayr, Michael (2009), Conditional β and σ-Convergence in Space: A Maximum Likelihood Approach, Regional Science and Urban Economics 39 (1), pp. 63-78.

Tu, Pierre N.V. (1992), Dynamical Systems, 2nd ed. Springer, New York.

Vayá, Ester, López-Bazo, Enrique, Moreno, Rosina and Jordi Surinach (2004), Growth and Externalities Across Economies: An Empirical Analysis Using Spatial Econometrics, in Anselin, Luc, Raymond J.G.M. Florax and Sergio S. Rey (ed.), Advances in Spatial Econometrics, Springer, Berlin, Heidelberg New York.

Appendix

Derivation of the convergence equation: Under stationarity of $\mathbf{k}^*$ and under the migration model the solution to the system of differential equations is given by (Tu, 1992, p. 98)

$$\mathbf{k}_t - \mathbf{k}_t{}^* = e^{\beta_M \mathbf{B}_M t} \left(\mathbf{k}_0 - \mathbf{k}_0^* \right).$$

Under capital mobility, one obtains

$$\begin{aligned} \mathbf{k}_t - \bar{\mathbf{k}}_t &= e^{\beta_K \mathbf{B}_K t} \left(\mathbf{k}_0 - \bar{\mathbf{k}} \right) - \left(\mathbf{I} - e^{\beta_l \mathbf{B}_l t} \right) \beta_K^{-1} \mathbf{B}_K^{-1} \mathbf{\Delta g}_K \\ &\approx e^{\beta_K \mathbf{B}_K t} \left(\mathbf{k}_0 - \bar{\mathbf{k}} \right) + t \mathbf{\Delta g}_K \end{aligned}$$

Note $e^{\beta_l \mathbf{B}_l t} = \mathbf{I} + \beta_l \mathbf{B}_l t + \beta_l^2 \mathbf{B}_l^2 \frac{t^2}{2!} + ...,l = M, K$. In analogy to the univariate case, one can derive the following econometric specification. Using $e^{\mathbf{B}_l (t-\tau)} = e^{\mathbf{B}_l t} e^{-\mathbf{B}_l \tau}$ one gets for the migration model

$$\begin{aligned} \mathbf{k}_t - \mathbf{k}_{t-\tau} &= -\left(\mathbf{I} - e^{\beta_M \mathbf{B}_M \tau} \right) \left(\mathbf{k}_{t-\tau} - \mathbf{k}^* \right) \\ \mathbf{k}_{t-\tau} &= \mathbf{k}_0 - \left(\mathbf{I} - e^{\beta_M \mathbf{B}_M \tau} \right) \left(\mathbf{k}_0 - \mathbf{k}^* \right) \end{aligned}$$

and for the capital mobility model

$$\begin{aligned} \mathbf{k}_t - \mathbf{k}_{t-\tau} &\approx -\left(\mathbf{I} - e^{\beta_K \mathbf{B}_K \tau} \right) \left(\mathbf{k}_{t-\tau} - \bar{\mathbf{k}} \right) + \tau \mathbf{\Delta g}_K \\ \mathbf{k}_{t-\tau} &\approx \mathbf{k}_0 - \left(\mathbf{I} - e^{\beta_M \mathbf{B}_K \tau} \right) \left(\mathbf{k}_0 - \bar{\mathbf{k}} \right) + (t - \tau) \mathbf{\Delta g}_K. \end{aligned}$$

Since the stock of capital usually remains unobserved, one has to specify this system in terms of real income per capita, $q_i = \ln \frac{Y_i}{L_i}$. Using the natural log of the production function given by

$$\mathbf{q}_t - \mathbf{a}_t = \rho \mathbf{W} \mathbf{k}_t + \alpha \mathbf{k}_t$$

one obtains

$$\mathbf{k}_t = (\alpha \mathbf{I} + \rho \mathbf{W})^{-1} (\mathbf{q}_t - \mathbf{a}_t) = \mathbf{F}^{-1}(\mathbf{q}_t - \mathbf{a}_t),$$

where $\mathbf{a}_t$ denotes labor saving technical progress. One can use the following decomposition: $(\alpha \mathbf{I} + \rho \mathbf{W}) \left(\mathbf{I} - e^{\beta_l \mathbf{B}_l \tau}\right) (\alpha \mathbf{I} + \rho \mathbf{W})^{-1} = \mathbf{P} \left(Diag(\alpha + \rho\lambda_i)\right) \mathbf{P}^{-1} \cdot \mathbf{P} Diag(1 - e^{\beta_l (1-\theta_l \lambda_i)\tau}) \cdot \mathbf{P}\mathbf{P}^{-1} Diag\left(\frac{1}{\alpha+\rho\lambda_i}\right) \mathbf{P}^{-1} = \mathbf{P}(1 - e^{\beta_l(1-\theta_l\lambda_i)\tau})\mathbf{P}^{-1} = \mathbf{I} - e^{\beta_l \mathbf{B}_l \tau}$. Hence, the convergence equation in GDP per capita under migration is given by

$$\mathbf{F} (\mathbf{k}_t - \mathbf{k}_{t-\tau}) = -\mathbf{F} \left(\mathbf{I} - e^{\beta_M \mathbf{B}_M \tau}\right) \mathbf{F}^{-1} (\mathbf{q}_{t-\tau} - \mathbf{a}_{t-\tau} - \mathbf{q}^* + \mathbf{a}_{t-\tau})$$

or

$$\mathbf{q}_t - \mathbf{q}_{t-\tau} = \mathbf{a}_t - \mathbf{a}_{t-\tau} - \left(\mathbf{I} - e^{\beta_M \mathbf{B}_M \tau}\right) \mathbf{q}_{t-\tau} + \left(\mathbf{I} - e^{\beta_M \mathbf{B}_M \tau}\right) \mathbf{q}^*,$$

and under capital mobility

$$\mathbf{q}_t - \mathbf{q}_{t-\tau} \approx \mathbf{a}_t - \mathbf{a}_{t-\tau} - \left(\mathbf{I} - e^{\beta_K \mathbf{B}_K \tau}\right) \mathbf{q}_{t-\tau} + \left(\mathbf{I} - e^{\beta_K \mathbf{B}_K \tau}\right) \mathbf{q}^* + \tau \mathbf{F} \boldsymbol{\Delta} \mathbf{g}_K.$$

Without knowledge spillovers, the possibility of labor migration or capital mobility, at $\rho = 0$, $\mathbf{B}_l = \mathbf{I}$ and $\beta_l = -(1-\alpha)(x+n+\delta)$, the spatial convergence equation collapses to the well known β-convergence equation as derived by e.g. McQuinn and Whelan (2007).

$$\mathbf{q}_t - \mathbf{q}_{t-\tau} = \mathbf{a}_t - \mathbf{a}_{t-\tau} - \left(1 - e^{\beta_l \tau}\right) \mathbf{q}_{t-\tau} + \left(1 - e^{\beta_l \tau}\right) \mathbf{q}^*.$$

Stability analysis: The stability of the system of differential equations describing the convergence precess is best analyzed in terms of k_i. The normalized spatial weighting matrix is decomposed as $\mathbf{W} = \mathbf{W}_1\mathbf{W}_2$, where $\mathbf{W}_1$ has full rank and is symmetric (Hermitian). The elements of $\mathbf{W}_1$ are given by $w_{1,ij} = e^{-d_{ij}/c}$, $\mathbf{W}_2 = Diag(1/d_i, .., 1/d_i)$, $d_i = \sum_{j=1}^{N} w_{1,ij}$ is the normalization matrix which has full rank and is positive definite. Theorem 7.6.3 of Horn and Johnson (1985, p. 465) implies that all eigenvalues of $\mathbf{W}$ are real. We denote the diagonal matrix comprising the eigenvalues of $\mathbf{W}$ by $\mathbf{\Lambda}$ and the corresponding matrix of eigenvectors by $\mathbf{P}$.

Now consider the characteristic roots of $\mathbf{B}_l = \beta_l(\mathbf{I}-\theta_l\mathbf{W})$. Since $\mathbf{W}$ is normalized by $\mathbf{W}_2$, we know from Gershgorin's Theorem that $|\lambda_i| \leq 1$ (see Theorem 6.1.1 in Horn and Johnson, 1985, p. 344). Furthermore, $\beta_l\mathbf{I}-\beta_l\theta_l\mathbf{P}^{-1}\mathbf{WP} = \beta_l\mathbf{I}-\beta_l\theta_l\mathbf{\Lambda}$. Therefore, the eigenvalues of $\beta_l\mathbf{B}_l$ are given by $\beta_l + \beta_l\theta_l\lambda_i$. Since $|\lambda_i| \leq 1$ and $-\beta_l\theta_l > 0$, we have $\beta_l - \beta_l\theta_l\lambda_i \leq \beta_l\,(1-\theta_l) < 0$ if $\theta_l < 1$, and the eigenvalues of $\mathbf{B}_l$ are all real and negative. In the migration model this implies the parameter restriction $g\rho + d < (1-\alpha)g + d$ or $1 - \alpha - \rho > 0$. Under the capital mobility model a similar assumption is required.

We conclude that under the maintained assumptions the system of differential equations implied by the spatial Solow model is Liapunov stable (Tu, 1992, p. 100). In particular, this implies that the system converges to a unique steady state, when starting in its neighborhood.

Stochastic specification of the system: With respect to the stochastic specification of the error term it is assumed that

$$\begin{aligned}
\mathbf{v}_t &= \delta\mathbf{v}_{t-1} + \varepsilon_t \\
\mathbf{a}_{t-\tau} &= a\,(t-\tau)\,\mathbf{e} + (\mathbf{I}-\phi\mathbf{W})^{-1}\,\mathbf{v}_{t-\tau} \\
\mathbf{a}_t &= at\mathbf{e} + (\mathbf{I}-\phi\mathbf{W})^{-1}\,\mathbf{v}_t \\
\mathbf{v}_t &= \delta^{\tau}\mathbf{v}_{t-\tau} + \sum_{m=0}^{\tau-1}\delta^m\varepsilon_{t-m}
\end{aligned}$$

where ε_{it} are normal iid random variables and $|\delta| < 1$. Using $\mathbf{v}_t - \mathbf{v}_{t-\tau} = \delta^{\tau}\mathbf{v}_{t-\tau} + \sum_{m=0}^{\tau-1}\delta^m\varepsilon_{t-m}\ \mathbf{v}_t - \mathbf{v}_{t-\tau} = (\delta^{\tau}-1)\mathbf{v}_{t-\tau} + \sum_{m=0}^{\tau-1}\delta^m\varepsilon_{t-m}$, it follows that

$$\begin{aligned}
\mathbf{a}_{t-\tau} &= a\,(t-\tau)\,\mathbf{e}{+}(\mathbf{I}-\phi\mathbf{W})^{-1}\xi_{t-\tau} \\
\tfrac{1}{\tau}\,(\mathbf{a}_t - \mathbf{a}_{t-\tau}) &= a\mathbf{e} + \tfrac{1}{\tau}(\mathbf{I}-\phi\mathbf{W})^{-1}\,(\mathbf{v}_t - \mathbf{v}_{t-\tau}) \\
&= a\mathbf{e} + (\mathbf{I}-\phi\mathbf{W})^{-1}\left[\tfrac{1}{\tau}(\delta^{\tau}-1)\mathbf{v}_{t-\tau} + \tfrac{1}{\tau}\sum_{m=0}^{\tau-1}\delta^m\varepsilon_{t-m}\right] \\
&= a\mathbf{e}{+}(\mathbf{I}-\phi\mathbf{W})_t^{-1}\xi_t
\end{aligned}$$

where $\xi_{t-\tau} = \delta^{t-\tau}\mathbf{v}_0 + \sum_{m=0}^{t-\tau-1}\delta^m\varepsilon_{t-\tau-m}$, and $\xi_t = (\delta^{\tau}-1)\xi_{t-\tau} + \sum_{m=0}^{\tau-1}\delta^m\varepsilon_{t-m}$. Note, the correlation of the error terms is

$$\begin{aligned}
E[\xi_{i(t-\tau)}\xi_{it}] &= E[(\delta^{\tau}-1)v_{i(t-\tau)}^2] = \\
&= (\delta^{2(t-\tau)}\sigma_0^2 {+} \tfrac{1-\delta^{2(t-\tau)}}{1-\delta^2}\sigma_{\varepsilon}^2)(\delta^{\tau}-1) \neq 0
\end{aligned}$$

In the empirical implementation the spatial correlation coefficient is allowed to differ between $t-\tau$ and t to capture the possibly different spatial correlation of the starting values.

Derivation of the likelihood: Following Anselin (1988) it is useful to introduce the following matrices:

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_0^2 & \sigma_{0t} \\ \sigma_{0t} & \sigma_t^2 \end{bmatrix}, \ \mathbf{\Sigma}^{-1} = \frac{1}{\sigma_0^2\sigma_t^2 - \sigma_{0t}^2} \begin{bmatrix} \sigma_t^2 & -\sigma_{0t} \\ -\sigma_{0t} & \sigma_0^2 \end{bmatrix}$$

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_0 & 0 \\ 0 & \mathbf{G}_t \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{2N} - \begin{pmatrix} \phi_0 \mathbf{W} & 0 \\ 0 & \phi_t \mathbf{W} \end{pmatrix} \end{bmatrix}$$

Stacking the two equations in a system with $2N$ rows yields

$$\begin{aligned} \mathbf{\Omega} &= E\left[\begin{pmatrix} \mathbf{G}_0 \xi_0 \\ \mathbf{G}_t \xi_t \end{pmatrix} (\mathbf{G}_0\xi_0, \mathbf{G}_t\xi_t)' \right] \\ &= \mathbf{G}^{-1}(\mathbf{\Sigma} \otimes \mathbf{I}_N)\mathbf{G}'^{-1} \\ \mathbf{\Omega}^{-1} &= \mathbf{G}'(\mathbf{\Sigma}^{-1} \otimes \mathbf{I}_N)\mathbf{G} \end{aligned}$$

and

$$\tfrac{1}{2} \ln \det \mathbf{\Omega} = \tfrac{N}{2} \ln(\sigma_t^2 \sigma_0^2 - \sigma_{0t}^2) - \ln \det \mathbf{G}_0 - \ln \det \mathbf{G}_t$$

For the structural form define

$$\begin{aligned} \mathbf{Y} &= \begin{bmatrix} \mathbf{q}_0 \\ \frac{1}{t}(\mathbf{q}_t - \mathbf{q}_0) \end{bmatrix}, \ \mathbf{\Gamma}_l = \begin{bmatrix} \mathbf{I}_N & -(\beta_l \mathbf{I} + \gamma_l \mathbf{W}) \\ 0 & \mathbf{I}_N \end{bmatrix}, \\ \mathbf{X} &= \begin{bmatrix} \mathbf{X}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_t \end{bmatrix}, \ \delta = \begin{bmatrix} \delta_0 \\ \delta_t \end{bmatrix} \end{aligned}$$

so that

$$\begin{aligned}
\xi_l &= (\mathbf{Y}-\mathbf{\Gamma}_l^{-1}\mathbf{X}\delta) = \mathbf{\Gamma}_l^{-1}(\mathbf{\Gamma}_l\mathbf{Y}-\mathbf{X}\delta) \\
\mathbf{\Omega}_\xi &= \mathbf{\Gamma}_l^{-1}\mathbf{\Omega}\mathbf{\Gamma}_l^{\prime -1} \\
\mathbf{\Omega}_\xi^{-1} &= \mathbf{\Gamma}_l^{\prime}\mathbf{\Omega}^{-1}\mathbf{\Gamma}_l \\
\det \mathbf{\Gamma}_l &= \det \begin{bmatrix} \mathbf{I}_N & \mathbf{0} \\ -(\beta_l\mathbf{I}+\gamma_l\mathbf{W}) & \mathbf{I}_N \end{bmatrix} = \det \mathbf{I}_N \det(\mathbf{I}_N + (\beta_l\mathbf{I}+\gamma_l\mathbf{W})\,\mathbf{I}_N\mathbf{0}) = 1.
\end{aligned}$$

$$\begin{aligned}
\xi_l^{\prime}\mathbf{\Omega}_\xi^{-1}\xi_l &= (\mathbf{\Gamma}_l\mathbf{Y}-\mathbf{X}\delta)^{\prime}\mathbf{G}^{\prime}\left(\begin{bmatrix} \sigma_0^2 & \sigma_{0t} \\ \sigma_{0t} & \sigma_t^2 \end{bmatrix}^{-1} \otimes \mathbf{I}_N\right)\mathbf{G}(\mathbf{\Gamma}_l\mathbf{Y}-\mathbf{X}\delta) \\
&= \tfrac{1}{\sigma_t^2\sigma_0^2(1-\rho^2)}(\xi_{l0}^{\prime},\xi_{lt}^{\prime}) \begin{bmatrix} \sigma_t^2\mathbf{I}_N & -\sigma_{0t}\mathbf{I}_N \\ -\sigma_{0t}\mathbf{I}_N & \sigma_0^2\mathbf{I}_N \end{bmatrix} (\xi_{l0}^{\prime},\xi_{lt}^{\prime})^{\prime} \\
&= \tfrac{1}{1-\rho^2}\left(\tfrac{\xi_{l0}^{\prime}\xi_{l0}}{\sigma_0^2} - \tfrac{2\rho\xi_{lt}^{\prime}\xi_{l0}}{\sigma_0\sigma_t} + \tfrac{\xi_{lt}^{\prime}\xi_{lt}}{\sigma_t^2}\right),
\end{aligned}$$

using $\sigma_{0t} = \rho\sigma_t\sigma_0$, $\sigma_t^2\sigma_0^2 - \sigma_{0t}^2 = \sigma_t^2\sigma_0^2(1-\rho^2)$ and

$$\begin{bmatrix} \xi_{l0} \\ \xi_{lt} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_0\,(\mathbf{q}_0 - \mathbf{X}_0\delta_0) \\ \mathbf{G}_t(\tfrac{1}{t}\,(\mathbf{q}_t-\mathbf{q}_0) - (\beta_l\mathbf{I}+\gamma_l\mathbf{W})\,\mathbf{q}_0 - \mathbf{X}_t\delta_t) \end{bmatrix}.$$

The log likelihood function is then given by

$$\begin{aligned}
&\ln L_l = c - \tfrac{N}{2}\ln(\sigma_t^2\sigma_0^2(1-\rho^2)) + \ln\det\mathbf{G}_0 + \ln\det\mathbf{G}_t \\
&\qquad - \tfrac{1}{2(1-\rho^2)}\left(\tfrac{\xi_{l0}^{\prime}\xi_{l0}}{\sigma_0^2} - \tfrac{2\rho\xi_{lt}^{\prime}\xi_{l0}}{\sigma_0\sigma_t} + \tfrac{\xi_{lt}^{\prime}\xi_{lt}}{\sigma_t^2}\right).
\end{aligned}$$

Table 1: Conditional β-convergence among European regions: 1980-2002, country fixed effects in both equations

	c=50, ρ=0		c=100, ρ=0		c=150, ρ=0		c=50, ρ≠0		c=100, ρ≠0		c=150, ρ≠0	
	b	z	b	z	b	z	b	z	b	z	b	z
Initial GDP per capita equation												
Latitude	-0.023	-0.1	0.044	0.1	0.095	0.2	-0.190	-0.5	-0.102	-0.2	-0.093	-0.2
Longitude	1.465	2.9 ***	1.485	2.9 ***	1.443	2.6 ***	1.327	3.0 ***	1.446	3.0 ***	1.180	2.5 ***
$\sigma_{\varepsilon 0}$	0.209		0.212		0.221		0.209		0.209		0.209	
ϕ_0	0.260	2.1 **	0.510	3.4 ***	0.752	6.2 ***	0.108	0.9	0.155	0.9	-0.264	-0.9
Growth of real income per capita equation												
Initial real income per capita	-0.011	-6.7 ***	-0.012	-7.0 ***	-0.012	-6.9 ***	-0.036	-58.7 ***	-0.036	-56.7 ***	-0.044	-79.6 ***
Initial real income per capita - spatially weighted	0.005	1.9 *	0.010	3.0 ***	0.013	3.1 ***	0.006	5.2 ***	0.013	7.1 ***	0.014	4.9 ***
$\sigma_{\varepsilon T}$	0.005		0.005		0.005	0.0	0.007		0.007		0.008	
$\sigma_{\varepsilon 0T}$	-		-		-		0.001	8.6 ***	0.001	8.6 ***	0.001	9.2 ***
ρ	-		-		-		0.734		0.727		0.817	
ϕ_T	0.189	1.4	0.098	0.5	-0.028	-0.1	0.170	1.5 #	0.089	0.5	0.038	0.2
LR-Test: Country fixed effects (32)	188.4 ***		188.3 ***		207.1 ***		196.2 ***		196.2 ***		215.6 ***	
Likelihood	478.9		480.6		480.1		482.8		485.0		485.2	

Note: Country dummies and the constant are not reported. The sample includes 212 European NUTSII regions. *** significant at 1%; ** significant at 5%; * significant at 10%; # significant at 15%.

Table 2: Conditional β-convergence among European regions: 1980-2002, country fixed effects in the initial GDP per capita equation

	c=50, ρ=0		c=100, ρ=0		c=150, ρ=0		c=50, ρ≠0		c=100, ρ≠0		c=150, ρ≠0	
	b	z	b	z	b	z	b	z	b	z	b	z
Initial GDP per capita equation												
Latitude	-0.023	-0.1	0.044	0.1	0.095	0.2	0.004	0.0	0.251	0.3	0.353	0.2
Longitude	1.465	2.9 ***	1.485	2.9 ***	1.443	2.6 ***	1.328	1.5 #	0.971	0.9	1.226	0.7
$\sigma_{\varepsilon 0}$	0.209		0.212		0.221		0.308		0.402		0.729	
ϕ_0	0.260	2.1 ***	0.510	3.4 ***	0.752	6.2 ***	0.723	13.9 ***	0.914	25.9 ***	0.963	42.1 ***
Growth of real income per capita equation												
Initial real income per capita	-0.011	-5.6 ***	-0.012	-6.2 ***	-0.013	-6.5 ***	-0.042	-57.4 ***	-0.057	-49.7 ***	-0.049	-20.6 ***
Initial real income per capita - spatially weighted	0.006	2.3 **	0.008	2.8 ***	0.010	3.2 ***	0.027	31.5 ***	0.048	31.8 ***	0.047	12.6 ***
$\sigma_{\varepsilon T}$	0.007		0.007		0.007		0.011		0.013		0.012	
$\sigma_{\varepsilon 0T}$	-		-		-		0.003	9.2 ***	0.004	8.6 ***	0.003	5.4 ***
ρ	-		-		-		0.826		0.732		0.403	
ϕ_T	0.535	3.7 ***	0.633	2.8 ***	0.688	2.3 ***	0.417	3.7 ***	0.431	2.3 **	0.451	1.5 #
LR-Test: Country fixed effects (16)	84.020 ***		79.340 ***		96.360 ***		88.560 ***		106.300 ***		119.980 ***	
Likelihood	426.710		426.180		424.690		428.980		440.010		437.390	

Note: Country dummies and the constant are not reported. The sample includes 212 European NUTSII regions. *** significant at 1%; ** significant at 5%; * significant at 10%; # significant at 15%.

Table 3: β-convergence among European regions: 1980-2002, series estimation

	Conditional convergence, country fixed effects in both equations				Unconditional convergence, country fixed effects only in the convergence equation			
	c=100, ρ=0		c=150, ρ≠0		c=50, ρ=0		c=100, ρ≠0	
	b	z	b	z	b	z	b	z
Initial GDP per capita equation								
Latitude	0.044	0.1	-0.092	-0.4	-0.023	-0.1	0.000	
Longitude	1.485	3.2 ***	1.261	2.2 ***	1.465	2.7 ***	1.060	2.1 **
$\sigma_{\varepsilon 0}$	0.212		0.209		0.209		0.383	
ϕ_0	0.510	2.1 ***	-0.192	-0.4	0.260	1.8 *	0.902	16.5 ***
Growth of real income per capita equation								
Initial real income per capita	-0.023	-11.2 ***	-0.039	-11.7 ***	-0.023	-10.1 ***	-0.041	-18.6 ***
Initial real income per capita - spatially weighted	0.022	4.5 ***	0.004	1.4 #	0.016	3.5 ***	0.007	2.6 ***
$\sigma_{\varepsilon T}$	0.006		0.025		0.010		0.037	
$\sigma_{\varepsilon 0T}$	-		0.001		-		0.010	
ρ	-		0.243		-		0.693	
ϕ_T	0.135	0.5	0.008	0.0	0.540	5.9 ***	0.714	5.3 ***
Likelihood	479.970		485.250		427.270		437.290	

Note: Country dummies and the constant are not reported. The standard errors are based on the numerically derived Hessian. The sample includes 212 European NUTSII regions. *** significant at 1%; ** significant at 5%; * significant at 10%; # significant at 15%.

Table 4: Speed of convergence under spatial spillovers, share of the income gap closed within a year on average

		Reduction in percent					
		Equal initial gap		initial gap = 1.2mean(y0)-y0			
Model	T	App.	Exact	Mean	Std	richest	poorest
Conditional convergence, c=100, ρ=0	1	0.19	0.19	0.24	0.69	1.98	0.55
	10		0.19	0.24	0.66	1.87	0.53
	50		0.18	0.22	0.54	1.49	0.46
Conditional convergence, c=150, ρ≠0	1	2.98	2.94	3.01	3.61	5.36	3.43
	10		2.58	2.63	3.06	4.33	2.93
	50		1.55	1.56	1.67	1.97	1.64
Unconditional convergence, c=50, ρ=0	1	0.58	0.58	0.61	0.85	1.56	0.78
	10		0.57	0.59	0.82	1.48	0.75
	50		0.51	0.53	0.69	1.16	0.64
Unconditional convergence, c=100, ρ≠0	1	0.92	0.92	1.16	3.22	9.22	2.60
	10		0.88	1.08	2.63	7.21	2.20
LR-Test: Country fixed effects (32)	50		0.74	0.84	1.43	2.98	1.28

Note: The gap is defined as 1.2mean(y_0)-y_0 and amounts to 0.59 on average. Its standard deviation is 0.43. The gap of the richest region is -0.42 and that of the poorest is 1.90. The averages figures are based on 50 years.