University of Innsbruck

Working Papers in Economics and Statistics

Conditional beta- and sigma-convergence in space: A maximum-likelihood approach

Michael Pfaffermayr

2007-17

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Economics (Institut für Wirtschaftstheorie, Wirtschaftspolitik und Wirtschaftsgeschichte)
- Department of Public Finance (Institut für Finanzwissenschaft)
- Department of Statistics (Institut für Statistik)

Contact Address:
University of Innsbruck
Department of Public Finance
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: +43 512 507 7151
Fax: +43 512 507 2970
e-mail: finanzwissenschaft@uibk.ac.at

The most recent version of all working papers can be downloaded at
http://www.uibk.ac.at/fakultaeten/volkswirtschaft_und_statistik/forschung/wopec

For a list of recent papers see the backpages of this paper.

Conditional β- and σ-Convergence in Space: A Maximum Likelihood Approach

Michael Pfaffermayr*

7th August 2007

Abstract

Empirical work on regional growth under spatial spillovers uses two workhorse models: the spatial Solow model and Verdoorn's model. This paper contrasts these two views on regional growth processes and demonstrates that in a spatial setting the speed of convergence is heterogenous in both considered models, depending on the remoteness and the income gap of all regions. Furthermore, the paper introduces Wald tests for conditional spatial σ-convergence based on a spatial maximum likelihood approach. Empirical estimates for 212 European regions covering the period 1980-2002 reveal a slow speed of convergence of about 0.7 percent per year under both models. However, pronounced heterogeneity in the convergence speed is evident. The Wald tests indicate significant conditional spatial σ-convergence of about 2 percent per year under the spatial Solow model. Verdoorn's specification points to a smaller and insignificant variance reduction during the considered period.

Keywords: Conditional spatial β- and σ-convergence; Spatial Solow model; Verdoorn's model; Spatial maximum likelihood estimates; European regions

JEL: R11; C21; O47

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

1 Introduction

The large empirical literature on regional income convergence under spatial externalities uses two workhorse models to motivate the estimated econometric specifications. The spatial Solow model assumes constant returns to scale and explains convergence by decreasing marginal returns of capital. Under the second model the growth process follows Verdoorn's law, which relates growth in GDP per capita to output growth and emphasizes increasing return to scale. In both cases economic growth is associated with convergence and spatial knowledge spillovers.

Empirical research on regional growth processes cannot treat regions as independent units. Regions interact in terms of knowledge spillovers, forward and backward linkages, factor mobility and trade. The large literature on spatial income convergence so far mainly concentrates on regional knowledge spillovers and accounts for them by a spatial weighting matrix to capture the stylized fact that regional spillovers usually decline in distance.

There is ample empirical evidence confirming the Solow view of regional convergence under regional externalities.[1] In particular, this literature provides evidence on spatial correlation in income growth rates, β-convergence (typically found to be slow) and emphasizes the importance of spatial knowledge spillovers that diminish in distance.

Under Verdoorn's law the growth rate of income per capita as a proxy of labor productivity is found to be positively related to output growth. Fingleton (2001) argues that the Verdoorn model provides a more realistic description of regional growth processes as it is compatible with certain endogenous growth models as well as with models of economic geography that allow for agglomeration of economic activity. Evidence on Verdoorn's model comes from Fingleton and

[1] see Armstrong, 1995; Chatterji and Dewhurst, 1996; Ades and Chua, 1997; Moreno and Trehan, 1997; Rey and Montouri, 1999; Attfield et al., 2000; Conley and Ligon, 2002; Carrington, 2003; López-Bazo et al., 2004; Vayáet al., 2004; Ertur, Le Gallo and Baumont, 2006.

McCombie (1998), Fingleton (2001) and Fingleton and López-Bazo (2006) to mention the most important studies. Islam (2003) and Abreu et al. (2005) provide comprehensive surveys on this issue. The latter authors in addition assess the empirical evidence of regional growth under spatial externalities in a meta analysis.

So far, most empirical evidence on income convergence at the regional level is confined to conditional β-convergence based on cross-section data. However, it is well established that evidence on conditional β-convergence does not imply that incomes indeed have become more equal over time in the sense of a shrinking income distribution. β-convergence is only a necessary condition, but not a sufficient one. Even if β -convergence and thus mean reversion is evident, conditional σ-convergence may be absent due the intra distribution dynamics as captured by the error term of the convergence equation (see Quah, 1993; Hart, 1995; Furceri, 2005; Rey and Janikas, 2005, Egger and Pfaffermayr, 2006a and 2006b; and Wodon, Yitzhaki, 2006). Evidence on (conditional) σ-convergence under spatial spillovers is scarce, however. Le Gallo and Dall'erba (2006) analyze unconditional σ-convergence among European regions and find mixed evidence based on the LR-test of Carree and Klomp (1997), which ignores spatial spillovers and is only suited to test unconditional σ-convergence. At a descriptive level, Rey and Dev (2006) report a reduction in the variance of US regional income over time. They show that the estimates of the variance of regional income can be substantially biased, if spatial correlation is ignored.

The main contribution of this paper is threefold. First, the theoretical basis of β and σ-convergence under spatial spillovers is at issue. It is demonstrated that in a spatial setting where distance matters the spatial spillovers lead to heterogeneity in both the speed of adjustment and the steady state growth rate, even if all model parameters and the initial income positions are identical among regions. The proposed spatial Solow model model differs from that one analyzed in Ertur and Koch (2007), since it incorporates knowledge spillovers in the way as originally proposed by Chua (1993). This specification has been widely used in the literature. In addition, the econometric specification is based on the solutions of the implied system of linear differential equations as derived in Egger and Pfaffermayr (2006b).

Second, new Wald tests for conditional σ-convergence under local and global spatial spillovers in the disturbances are derived for both the spatial Solow model and the spatial Verdoorn model. With these tests the descriptive evidence on σ-convergence mentioned above can be investigated in more detail by (i) allowing for spatial dependence and conditional convergence and (ii) by providing a concise econometric test. These Wald tests are based on the spatial maximum likelihood estimator introduced by Anselin (1988), which has been extensively applied in the spatial convergence literature. Once maximum likelihood estimates of the spatial convergence models are available, the proposed Wald tests are easy to calculate, since they only use the output provided by any numerical maximum likelihood maximization procedure. Monte Carlo experiments suggest that the proposed Wald tests are properly sized in medium to large samples and that they are powerful enough to detect conditional σ-convergence in a spatial setting if it indeed took place.

Third, the proposed approach to test spatial conditional β- and σ-convergence is illustrated in an empirical analysis of the growth process of European regions during the period $1980-2002$. In line with the previous findings, there is evidence of conditional β-convergence with an average convergence speed of about 0.7 percent per year under both approaches. While the spatial Solow model implies significant conditional σ-convergence with a reduction of the average predicted variance in income per capita of about 2 percent per year, the findings for Verdoorn's model are less clear cut. The estimates suggest a smaller average variance reduction (approximately 1.4 percent per year), and under global error spillovers it is insignificant. In this latter case, β-conditional convergence does not imply conditional σ-convergence and one may erroneously conclude that convergence occurs (Gallon's fallacy). Unfortunately, these two models are not nested and cannot be tested against each other. However, model selection criteria marginally prefer the spatial Solow model.

The paper is organized as follows. Section 2 presents the spatial Solow and Verdoorn's model upon which the econometric estimates are based. Section 3 introduces the econometric specifications as linear approximations of the theoretical models, derives the implied speed of convergence and introduces the concept of spatial σ-convergence and tests thereof. Section 4 presents estimates as well as the proposed Wald test for European regions. The last section summarizes the main findings.

2 Two views on regional growth

The first view of regional convergence in income per capita is based on the spatial Solow model with a constant return to scale production function, exhibiting a diminishing marginal product of capital and spatial knowledge spillovers (see Egger and Pfaffermayr, 2006b; Ertur and Koch, 2007; López-Bazo et al., 2004, and Vayá et al., 2004).[2] Formally, under regional knowledge spillovers the aggregate production function of region i in intensive form (indicated by a "hat") is given by:

$$\widehat{y}_i = \frac{Y_i}{L_i A} = C(\Pi_{j \neq i} \widehat{k}_j^{\rho w_{ij}}) \widehat{k}_i^{\alpha}. \tag{1}$$

$\widehat{y}_i$ denotes output per efficiency unit of labor and $\widehat{k}_i$ the stock of capital per efficiency unit of labor ($\frac{K_i}{L_i A}$). Spatial externalities are reflected in the total factor productivity (TFP) of a region that is given by $C \Pi_{j \neq i} \widehat{k}_j^{\rho w_{ij}}$ with $0 < \rho < 1 - \alpha$ and C a normalizing constant. This specification is based on the work of Chua (1993) and captures regional spillovers such as learning effects in course of

[2] Obstfeld and Rogoff (1996) demonstrate for the Solow model without spillovers that this model also covers the case of mobile capital as long as one factor (e.g., human capital) is immobile and accumulated.

capital accumulation (see also Moreno and Trehan, 1997). However, it differs from Ertur and Koch (2007) who assume spatial spillovers in total factor productivity rather than capital per efficiency unit of labor. The proposed production function implies that the total factor productivity of a specific region is positively related to the level of development of its surrounding regions as measured by their spatially weighted ratio of capital to efficiency units of labor. The strength of regional spillovers depends on their overall intensity $\rho > 0$, the spatial weights w_{ij} which are inversely related to distance for $i \neq j$ and the regions' initial level of $\widehat{k}_i$. Note, $w_{ij} = 0$ for $i = j$. Hence, the spatial spillovers between any two regions decrease with the distance between them and at a given distance they are the higher the more advanced neighboring regions are. Under a constant savings rate s, the system of first order non-linear differential equations describing the growth process is given by:

$$\frac{d \ln \widehat{k}_i}{dt} = sC(\Pi_{j \neq i} \widehat{k}_j^{\rho w_{ij}}) \widehat{k}_i^{\alpha - 1} - (x + n + \delta), \; i = 1, ..., N. \tag{2}$$

x denotes the exogenously given growth rate of TFP represented by C. n is the population growth rate and δ stands for the rate of depreciation of capital.

At $\frac{d \ln \widehat{k}_i}{dt} = 0$ for all i, this system of first order differential equations implies the steady state values $\widehat{k}_i^* = \left(\frac{x+n+\delta}{sC\Pi_{j \neq i} \widehat{k}_j^{*\rho w_{ij}}} \right)^{\frac{1}{a-1}}$. Denoting the steady state values $\ln(\widehat{k}_i^*)$ by k_i^* and stacking them into a vector results in $\mathbf{k}^* = \frac{1}{1-\alpha}[(\ln \frac{sC}{x+n+\delta})\mathbf{e} + \rho \mathbf{W} \mathbf{k}^*]$, where $\mathbf{W}$ is the row-normalized spatial weighting matrix with elements w_{ij} and $\mathbf{e}$ a vector of ones. Note bold letters denote matrices or vectors. Hence, the steady state levels of capital per efficiency unit of labor are given by $\mathbf{k}^* = \frac{1}{1-\alpha}$ $(\ln \frac{sC}{x+n+\delta})(\mathbf{I} - \frac{\rho}{1-\alpha}\mathbf{W})^{-1}\mathbf{e}$, where the inverse of $(\mathbf{I} - \frac{\rho}{1-\alpha}\mathbf{W})$ is assumed to exist (see the conditions below).

Under $\rho = 0$ the derived steady state is identical to that derived for the traditional Solow model. However, with spatial spillovers ($\rho > 0$) the steady state values are augmented by a global spatial multiplier. Therefore, the steady levels $\mathbf{k}^*$ are heterogenous across regions and depend on their location via $\mathbf{W}$, even if all model parameters are identical. According to the typology of spatial econometric models (Anselin, 2003) the spillovers are global as one region affects all others via the spatial multiplier $(\mathbf{I} - \frac{\rho}{1-\alpha}\mathbf{W})^{-1}$. Since all diagonal entries in the spatial multiplier matrix are 1 and the off-diagonal entries are positive, the steady state ratios of capital to efficiency units of labor and, therefore, income per capita, is enhanced by the spatial multiplier in the steady state.

The corresponding linearly approximated spatial Solow growth equation can be easily derived (see Egger and Pfaffermayr, 2006b):

$$\dot{k}_i \simeq (\alpha - 1)(x + n + \delta)(k_i - k_i^*) + \rho(x + n + \delta) \sum_{j \neq i} w_{ij}(k_j - k_j^*) \tag{3}$$

Inserting in the production function, one can show that the same law of motion holds for the log of output per unit of labor (denoted by $\mathbf{q}$ with components $q_i = \ln(\frac{Y_i}{L_i})$). In vector form the corresponding system of first order differential equations is given by:

$$\begin{aligned} \dot{\mathbf{q}} - \dot{\mathbf{q}}^* &= \beta(\mathbf{I} - \tfrac{\rho}{1-\alpha}\mathbf{W})(\mathbf{q} - \mathbf{q}^*) := \mathbf{B}(\mathbf{q} - \mathbf{q}^*) \\ \text{with } \beta &= -(1-\alpha)(x + n + \delta). \end{aligned} \tag{4}$$

A steady state and, hence, a balanced growth path exists, if $\left(\mathbf{I} - \frac{\rho}{1-\alpha}\mathbf{W}\right)$ is invertible and if $\beta\left(\mathbf{I} - \frac{\rho}{1-\alpha}\mathbf{W}\right)$ has real and negative eigenvalues. That this is indeed the case under the present assumptions is formally demonstrated in Egger and Pfaffermayr (2006b). This system of first order differential equations possesses the following solutions which form the basis of the empirical specification to be estimated below (see Tu, 1992, p.98) :

$$\mathbf{q}_t - \mathbf{q}_t^* = e^{\mathbf{B}t}(\mathbf{q}_0 - \mathbf{q}_0^*), \tag{5}$$

where $e^{\mathbf{B}t} = \mathbf{I} + \mathbf{B}t + \mathbf{B}^2\frac{t^2}{2!} + ...$, $\mathbf{B} = \beta(\mathbf{I} - \frac{\rho}{1-\alpha}\mathbf{W})$. Reformulating the system in log-differences of GDP per capita for a period of given length T yields the convergence equation which will be estimated in a reparameterized, approximated form below:

$$\Delta\widetilde{\mathbf{q}}_T - \Delta\widetilde{\mathbf{q}}_T^* = -\frac{1}{T}\left(\mathbf{I} - e^{\mathbf{B}t}\right)(\mathbf{q}_0 - \mathbf{q}_0^*)\,, \tag{6}$$

with $\Delta\widetilde{\mathbf{q}}_T = \frac{1}{T}(\mathbf{q}_T - \mathbf{q}_0)$ and $\Delta\widetilde{\mathbf{q}}_T^*$ analogously defined. This specification of the convergence equation differs from that one used in the previous literature (e.g. Ertur and Koch, 2007; Fingleton and López-Bazo, 2006; Le Gallo and Dall'erba (2006); López-Bazo et al., 2004 and Vayá et al., 2004) as it is based on the solution of the linearized system of differential equations that defines the growth path of the regions.

Fingleton and McCombie (1998), Fingleton (2001, 2003) and Fingleton and López-Bazo (2006) introduce a different view on regional growth based on Verdoorn's law which states that the growth rate of income per capita is positively related to the growth rate of real income due to economies of scale. An increase in output thus leads to a higher labor productivity. Fingleton (2001) shows that, as a reduced form, this approach is consistent with models of economic geography and with certain endogenous growth models both based on economies of scale and spatial knowledge spillovers. The basic equation describing Verdoorn's law (Fingleton, 2001, p. 126) likewise assumes an aggregate Cobb-Douglas production function. However, Fingleton (2001) introduces non-tradeable, differentiated intermediate inputs and allows for economies of scale based on the approach of Ciccone and Hall (1996) and as a result agglomeration effects. He shows that such a model leads to the following reduced form Verdoorn-equation,

$$\ln(\widehat{y}_{it}) = \ln(\frac{Y_{it}}{A_{it}L_{it}}) = \ln D + \gamma \ln Y_{it},\ t = 1, T \tag{7}$$

where γ is the Verdoorn parameter and $\ln D$ summarizes the remaining structural parameters of the model. The evolution of labor augmenting technical progress A_{it} over time is explained by further variables collected in $\mathbf{Z}_i$ as well as regional spillovers in productivity growth and a convergence term. The latter rests on the hypothesis that learning is more efficient and catching up is faster the higher the initial income gap is (i.e. difference to the leader or to the steady state, see Fingleton, 2001, p. 127 and p. 133 as well as Fingleton and López-Bazo, 2006). In sum, $\ln\frac{A_i(T)}{A_i(0)} = \delta_1 T + \Delta\mathbf{Z}_i\delta_2 + \rho\Sigma_{i=1}^{N} w_{ij}\ln\frac{A_j(T)}{A_j(0)} + \beta T\left(q_{i0} - q_{i0}^{*}\right)$, where $\rho > 0$ and $\beta < 0$ is hypothesized. Using matrix notation one gets $\Delta\ln\mathbf{A} = \delta_1 T\mathbf{e} + \Delta\mathbf{Z}\delta_2 + \rho\mathbf{W}\Delta\ln\mathbf{A} + \beta T\left(\mathbf{q}_0 - \mathbf{q}_0^{*}\right)$. This gives an econometric specification similar to that estimated in Fingleton and McCombie (1998, p. 95) and Fingleton (2001, p. 128):

$$\Delta\widetilde{\mathbf{q}}_T = \gamma\Delta\widetilde{\mathbf{Y}}_T + (\mathbf{I} - \rho\mathbf{W})^{-1}(\delta_1\mathbf{e} + \Delta\widetilde{\mathbf{Z}}\delta_2 + \beta\left(\mathbf{q}_0 - \mathbf{q}_0^{*}\right)) \tag{8}$$

It has to be emphasized that this specification does not close the model and treats output growth as an exogenous variable. Specifying the system in continuous time and subtracting the constant steady state growth rate given by $\dot{\mathbf{q}}^{*} = \gamma\ln\dot{\mathbf{Y}} + (\mathbf{I} - \rho\mathbf{W})^{-1}(\delta_1\mathbf{e} + \delta_2\dot{\mathbf{Z}})$, results in

$$(\dot{\mathbf{q}} - \dot{\mathbf{q}}^{*}) = \beta(\mathbf{I} - \rho\mathbf{W})^{-1}\left(\mathbf{q} - \mathbf{q}^{*}\right) := \mathbf{C}\left(\mathbf{q} - \mathbf{q}^{*}\right). \tag{9}$$

Verdoorn's model likewise comprises global knowledge spillovers interrelating the growth rates of all regions. The solution of this system of differential equations, upon which an empirical specification of Verdoorn's model can be based, reads as follows

$$\mathbf{q}_t - \mathbf{q}_t^{*} = e^{\mathbf{C}t}\left(\mathbf{q}_0 - \mathbf{q}_0^{*}\right), \tag{10}$$

where $\mathbf{C} = \beta(\mathbf{I} - \rho\mathbf{W})^{-1}$ and $e^{\mathbf{C}t} = \mathbf{I} + \mathbf{C}t + \mathbf{C}^2\frac{t^2}{2!}\ldots$. The inverse $\beta(\mathbf{I} - \rho\mathbf{W})^{-1} = \beta\sum_{k=0}^{\infty}\rho^k\mathbf{W}^k = \beta\mathbf{P}^{-1}\left(\sum_{k=0}^{\infty}\rho^k\mathbf{\Lambda}^k\right)\mathbf{P}$ exists for $\rho < 1$ since $\mathbf{W}$ is row normalized. The real eigenvalues of $\beta(\mathbf{I} - \rho\mathbf{W})^{-1}$ are $\widetilde{\widetilde{\Lambda}} = Diag[\frac{\beta}{1-\rho\lambda_i}]$ so that $e^{\mathbf{C}t} = \mathbf{P}^{-1}e^{\widetilde{\widetilde{\mathbf{\Lambda}}}t}\mathbf{P}$. Since $|\rho\lambda_i| < |\rho|\,|\lambda_i| < 1$ and $\beta < 0$ all eigenvalues of $\mathbf{C}$ are real and negative and one can conclude that this system is Liapunov stable (Tu, 1992, p. 100).

The main difference between the two approaches lies in the set of variables explaining the steady state and in the nature of spillovers. In the spatial Solow model exogenous technical progress along with investment in physical capital drives the growth process, while diminishing returns lead to convergence. In the Verdoorn specification output growth is exogenously given and it translates into productivity growth, since employment grows at a lower rate under economies of scale. The convergence process is based on learning, but formally quite similar to the spatial Solow model. In both models there are regional spillovers which decline with distance. In the Solow model these are local and depend on differences in initial capital labor ratios. Over time they spread out in course of the growth process and become global. In Verdoorn's model, in contrast, regional spillovers rest on a region's ability to learn the productivity improvements of its neighbors as captured by the global spatial multiplier. Unfortunately, these two specifications are not nested and cannot be tested against each other. However, a comparison in terms of model selection criteria is possible.[3]

3 The econometric specification, spatial maximum likelihood estimation and the Wald test on conditional σ-convergence

Econometric model: The empirical counterpart of the non-linear convergence equations of the spatial Solow model and the spatial Verdoorn model has to rely on a linear approximation, which is derived in Appendix A. Specifically, in this Appendix conditions are given, which guarantee that the first order term in the

[3]In a different setting, Fingleton (2006) proposes to test two non-nested models in an artificially nesting model which encompasses both. This approach cannot be applied here, since the artificially nesting model would include log income growth and log output population growth which together make up the dependent variable.

approximation dominates, while the higher order terms tend to zero fast. Hence, the approximation error remains small under the proposed approximation under a plausible set of parameters.

For the spatial Solow-model one can use the approximation $-\frac{1}{T}\left(\mathbf{I}-e^{\mathbf{B}t}\right) \approx -\frac{1}{T}(1-e^{\beta T})\mathbf{I}+\widetilde{\rho}e^{\beta T}\mathbf{W} := b\mathbf{I}+r\mathbf{W}$ with $\widetilde{\rho} = -\frac{\beta\rho}{1-\alpha}$. Therefore, one may interpret the estimated coefficients as $b = -\frac{1}{T}(1-e^{\beta T})$, resulting in $\beta = \frac{1}{T}\ln(1+bT)$, and $r = \widetilde{\rho}e^{\beta T} = -\frac{\beta\rho}{1-\alpha}e^{\beta T}$ so that $\rho = -\frac{r(1-\alpha)T}{(1+bT)\ln(1+bT)}$. In the absence of spillovers at $r=\rho=0$ the convergence term reverts to that derived in Barro and Sala-i-Martin (2004). The linear specification of the spatial Solow model is then given by

$$\begin{aligned}\mathbf{q}_0 &= \mathbf{X}_0\delta_0+\mathbf{u}_0 \\ \Delta\widetilde{\mathbf{q}}_T &= \Delta\tfrac{1}{T}\mathbf{q}_T = b\mathbf{q}_0+r\mathbf{W}\mathbf{y}_0+\mathbf{X}_T\delta_T+\mathbf{u}_T.\end{aligned} \tag{11}$$

The parametrization of Verdoorn's model can be based on the approximation $-\frac{1}{T}\left(\mathbf{I}-e^{\mathbf{C}T}\right) \approx \beta(\mathbf{I}-\rho\mathbf{W})^{-1}$. Hence, the specification to be estimated is given by

$$\Delta\widetilde{\mathbf{q}}_T = \beta\left(\mathbf{q}_0-\mathbf{q}_0^*\right)+\rho\mathbf{W}\Delta\widetilde{\mathbf{q}}_T+\gamma(\mathbf{I}-\rho\mathbf{W})\Delta\widetilde{\mathbf{Y}}_T+\delta_1+\Delta\widetilde{\mathbf{Z}}\delta_2. \tag{12}$$

using $(\mathbf{I}-\rho\mathbf{W})\Delta\widetilde{\mathbf{q}}_T^* = \gamma(\mathbf{I}-\rho\mathbf{W})\Delta\widetilde{\mathbf{Y}}_T+\delta_1+\Delta\widetilde{\mathbf{Z}}\delta_2$. The system to be estimated reads

$$\begin{aligned}\mathbf{q}_0 &= \mathbf{X}_0\delta_0+\mathbf{u}_0 \\ \Delta\widetilde{\mathbf{q}}_T &= b\mathbf{q}_0+r\mathbf{W}\Delta\widetilde{\mathbf{q}}_T+\gamma\Delta\widetilde{\mathbf{Y}}_T+\varphi\mathbf{W}\Delta\widetilde{\mathbf{Y}}_T+\mathbf{X}_T\delta_T+\mathbf{u}_T.\end{aligned} \tag{13}$$

In this approximation the parameters can be estimated directly, i.e. $b=\beta$ and $r=\rho$. This specification is usually referred to as Durbin representation, implying the parameter restriction $\varphi = -\gamma\rho$. The restrictions on the spatial error term are not considered here, since the error term inter alia captures the approximation error (see Fingleton and López-Bazo, 2006).

As shown below, the Wald tests on conditional σ-convergence are based on a non-linear restriction that involves the convergence parameter, the spillover parameter as well as the predicted variances of the final and the initial income (after conditioning on the exogenous explanatory variables). Therefore, an equation that explains initial income per capita has to be estimated in addition to the convergence equation.

For both models, it is likely that there are unobserved spillovers across regions stemming from other sources (random shocks) than that captured by the systematic part of the econometric model. Indeed, many studies on income convergence among European regions have found spatially correlated error terms based on diagnostic tests (see e.g. Fingleton and López-Bazo ,2006; Le Gallo and Dall'erba, 2006; López-Bazo et. al 2004, and Vayá, et al. 2004 to mention a few). Generally, the spatial correlation error terms may either be local or global. In both cases, inference on β- and σ-convergence will be misleading if these are ignored.

With local spillovers in the error process one assumes

$$\mathbf{u}_t = (\mathbf{I} - \phi_t \mathbf{W})\varepsilon_t := \mathbf{L}_t \varepsilon_t, \; t = 0, T \tag{14}$$

while the assumption of global spillovers in the error maintains:

$$\mathbf{u}_t = (\mathbf{I} - \phi_t \mathbf{W})^{-1} \varepsilon_t := \mathbf{G}_t \varepsilon_t, \; t = 0, T. \tag{15}$$

In both cases, $|\phi_0| < 1$ and $|\phi_T| < 1$ is assumed so that the corresponding spatial multipliers exist.

The specification of Verdoorn's model uses a different weighting matrix for the error term (see below for the details), since this model includes both a spatially lagged endogenous variable and the spatially correlated error term. It is well known that the likelihood can be rather flat under identical spatial weighting matrices so that the corresponding parameters are hard to identify (see also Anselin, 1988).[4]

Speed of convergence: The speed of convergence of a single region i is defined as the share $1-\psi_i$ of the initial gap in GDP per capita, which has been closed after T years. Following Egger and Pfaffermayr (2006b) the income gap at time t is denoted by $\mathbf{v}_t = \mathbf{q}_t - \mathbf{q}_t^*$, $t = 0, T$ and assumed to be negative for all regions. Then, the share of the gap still present after T periods under the spatial Solow model is given by

$$\psi = Diag[\mathbf{v}_0]^{-1} e^{\mathbf{B}T} \mathbf{v}_0. \tag{16}$$

Using the estimated parameters of the econometric model one obtains the following approximation for the average speed of convergence:

$$\begin{aligned} \psi &\approx Diag[\mathbf{v}_0]^{-1}[(1+Tb)\mathbf{I}+Tr\mathbf{W}]\mathbf{v}_0 \qquad (17) \\ &= (1+Tb)\mathbf{e} + TrDiag[\mathbf{v}_0]^{-1}\mathbf{W}\mathbf{v}_0 \text{ or} \\ \frac{1-\psi_i}{T} &\approx -b - r\sum_{j=1}^{N} w_{ij}\frac{v_{0j}}{v_{0i}}, \ i = 1, ..., N. \end{aligned}$$

Note $\frac{1-\psi_i}{T}$ is the share of the income gap which is closed within a year on average. This approximation is valid if the initial income gap is negative for all regions, i.e., there is convergence from below and the income gap is not too small in absolute value. A sufficient condition is $v_{0i} < -\frac{r}{b}(\mathbf{W}\mathbf{v}_0)_i < 0$ for all i. This result illustrates that the speed of convergence varies across regions and in addition to convergence parameter b depends on strength of spillovers. The lower the absolute value of the gap in initial income of a region's neighbors, the more the region can learn and the higher the spatial spillovers. If the income gaps are

[4]In fact, in the empirical analysis below the spatial spillover parameters could not be estimated with reasonable precision under identical spatial weighting schemes for the endogenous spatial lag and the spatially correlated errors.

negative and convergence occurs from below for all regions, we see that ignoring regional spillovers leads to an overestimation of the convergence speed. In the absence of regional spillovers, at $r = \rho = 0$, the speed of convergence collapses to that calculated e.g. by Barro and Sala-i-Martin (2004): $\psi_i = 1 + bT = e^{\beta T}$ and $\beta = \frac{1}{T}\ln\,(1 + Tb)$.

Under Verdoorn's law the speed of convergence can be calculated as

$$
\begin{aligned}
\psi &= Diag[\mathbf{v}_0]^{-1} e^{\mathbf{C}T}\mathbf{v}_0 && (18)\\
&\approx \mathbf{e} + Diag[\mathbf{v}_0]^{-1}[bT(\mathbf{I} - r\mathbf{W})^{-1}]\mathbf{v}_0 \text{ or} \\
\frac{1}{T}(\mathbf{e} - \psi) &\approx -bDiag[\mathbf{v}_0]^{-1}[(\mathbf{I} - r\mathbf{W})^{-1}]\mathbf{v}_0 && (19)
\end{aligned}
$$

In this case, there is no need to introduce restrictions to guarantee that the convergence speed is positive. Again, there is heterogeneity in the speed of convergence across regions. However, in contrast to the spatial Solow model the regional spillover parameter increases the convergence speed at negative income gaps.[5]

With the estimated parameters at hand the implied speed of convergence for both models can be calculated easily, if one is willing to make assumptions on the initial income gap in an experiment of thought. Since the steady state gap remains unidentified in the econometric specification, the direct estimation of the speed of convergence based on the estimated parameters only is not possible, however.

σ-convergence: Conditional σ-convergence is observed if the income distribution shrinks over time, controlling for the exogenous variables. This implies a decreasing variance of predicted GDP per capita. In a spatial setting, the variance of predicted GDP per capita is heterogenous across regions due to the spatial

[5]This can be seen by differentiating $(\mathbf{I} - r\mathbf{W})^{-1}$ in (18): $\frac{\partial(\mathbf{I}-r\mathbf{W})^{-1}}{\partial r} = (\mathbf{I} - r\mathbf{W})^{-1}\mathbf{W}(\mathbf{I} - r\mathbf{W})^{-1}$. Since all elements of the matrix of derivatives are positive, one can conclude that an increase in r indeed accelerates convergence.

spillovers. Therefore, the conditional σ-convergence is defined as the variance of $q_{it}, i = 1, ..N$ being smaller in period T than in period 0 *on average.* This hypothesis can be tested by a Wald test of a non-linear restriction on the second moments of $\mathbf{q}_T$ and $\mathbf{q}_0$, conditioning on the explanatory variables. Under global regional spillovers in the error terms the variance covariance matrix of $\mathbf{q}_0$ and $\Delta\widetilde{\mathbf{q}}_T$ under the spatial Solow model is given by

$$\begin{aligned}
\mathbf{\Sigma}_0 &= E\left[\mathbf{G}_0\varepsilon_0\varepsilon_0'\mathbf{G}_0'\right] = \sigma^2_{\varepsilon_0}\mathbf{G}_0\mathbf{G}_0'. \qquad (20)\\
\mathbf{\Sigma}_T &= \left((1+Tb\mathbf{I})+Tr\mathbf{W}\right)\mathbf{G}_0\mathbf{G}_0'\left((1+Tb\mathbf{I})+Tr\mathbf{W}\right)'\sigma^2_{\varepsilon_0} + T^2\mathbf{G}_T\mathbf{G}_T'\sigma_{\varepsilon_T},
\end{aligned}$$

using $\mathbf{q}_T - E\left[\mathbf{q}_T\right] = T\left(b\mathbf{I}+r\mathbf{W}\right)\mathbf{G}_0\varepsilon_0 + T\mathbf{G}_T\varepsilon_T$. Under H_0 the Wald test for conditional σ-convergence considers the non-linear restriction:

$$H_0: \ \varphi^{GS}(\theta) = \frac{1}{N}tr(\mathbf{\Sigma}_T - \ \mathbf{\Sigma}_0) = 0 \text{ versus } H_1: \ \varphi(\theta)^{GS} < 0, \qquad (21)$$

where the parameter vector is $\theta = (\delta_0, b, r, \delta_T, \sigma^2_{\varepsilon_0}, \sigma^2_{\varepsilon_T}, \phi_0, \phi_T)$. The available tests for conditional convergence without spatial spillovers (see Egger and Pfaffermayr, 2006a) result as a special case and illustrate this approach. Setting $r = \phi_0 = \phi_T = 0$ and $T = 1$, one obtains $\varphi^{GS}(\ \theta) = (1+b)^2\sigma^2_{\varepsilon_0} + \sigma^2_{\varepsilon T} - \sigma^2_{\varepsilon_0}$ and the test simplifies to H_0: $(1+b)^2 = \frac{\sigma^2_{\varepsilon 0} - \sigma^2_{\varepsilon T}}{\sigma^2_{\varepsilon_0}}$ vs. H_0: $(1+b)^2 < \frac{\sigma^2_{\varepsilon 0} - \sigma^2_{\varepsilon T}}{\sigma^2_{\varepsilon_0}}$. This again illustrates that β-convergence, or the presence of mean reversion, is not sufficient to guarantee conditional σ-convergence (see Hart, 1995; Furceri, 2005; Rey and Janikas, 2005; Egger and Pfaffermayr, 2006b; and Wodon and Yitzhaki, 2006). It is perfectly possible that the variance remains constant over time or even increases despite finding $b < 0$. Rather, for a reduction of the variance of GDP per capita the variance induced by the intra-distribution dynamics as captured by $\sigma^2_{\varepsilon_T}$ has to be sufficiently small as compared to the variance reduction resulting from Galtonian mean reversion, $b^2\sigma^2_{\varepsilon_0}$. This implicitly assumes that the variance of initial GDP per capita is larger that its steady state counterpart, otherwise one would observe conditional σ-divergence. For this simple case, Egger and Pfaffermayr (2006a) demonstrate that one can set $T = 1$ without loss of generality, since average growth rates are considered. In addition, they show that the Wald test possesses maximum power in this case. Therefore, the Wald test for spatial, conditional σ-convergence is based on $T = 1$ and the restriction to test simplifies to

$$\begin{aligned} \varphi^{GS}(\theta) &= \tfrac{1}{N}[b(2+b)tr(\mathbf{G}_0\mathbf{G}_0')+2r(1+b)tr(\mathbf{W}\mathbf{G}_0\mathbf{G}_0') \\ &\quad +r^2 tr(\mathbf{W}\mathbf{G}_0\mathbf{G}_0'\mathbf{W}')]\sigma^2_{\varepsilon_0}+\tfrac{1}{N}tr(\mathbf{G}_T\mathbf{G}_T')\sigma^2_{\varepsilon_T}. \end{aligned} \tag{22}$$

The vector of first derivatives of this restriction is denoted by $\mathbf{R}^{SG} = \frac{\partial\varphi(\theta)^{GS}}{\partial\theta}$ and reads

$$\mathbf{R}^{SG} = \begin{bmatrix} \mathbf{0} \\ 2\left[(1+b)\frac{tr(\mathbf{G}_0\mathbf{G}_0')}{N}+r\frac{tr(\mathbf{W}\mathbf{G}_0\mathbf{G}_0')}{N})\right]\sigma^2_{\varepsilon_0} \\ 2\left[(1+b)\frac{tr(\mathbf{W}\mathbf{G}_0\mathbf{G}_0')}{N}+r\frac{tr(\mathbf{W}\mathbf{G}_0\mathbf{G}_0'\mathbf{W}')}{N}\right]\sigma^2_{\varepsilon_0} \\ b(2+b)\frac{tr(\mathbf{G}_0\mathbf{G}_0')}{N}+r(1+b)\frac{2tr(\mathbf{W}\mathbf{G}_0\mathbf{G}_0')}{N}+r^2\frac{tr(\mathbf{W}'\mathbf{W}\mathbf{G}_0\mathbf{G}_0')}{N} \\ \frac{tr(\mathbf{G}_T\mathbf{G}_T')}{N} \\ \left[b(2+b)\frac{tr(\boldsymbol{\Upsilon}_0^{SG})}{N}+2r(1+b)\frac{tr(\mathbf{W}\boldsymbol{\Upsilon}_0^{SG})}{N}+r^2\frac{tr(\mathbf{W}'\mathbf{W}\boldsymbol{\Upsilon}_0^{SG})}{N}\right]\sigma^2_{\varepsilon_0} \\ \frac{tr(\boldsymbol{\Upsilon}_T^{SG})}{N}\sigma^2_{\varepsilon_T} \end{bmatrix}, \tag{23}$$

where one can use $dtr(\mathbf{X}) = tr(d\mathbf{X})$, $d\mathbf{G}_t = -\mathbf{G}_t d(\mathbf{I}-\phi_t\mathbf{W})\mathbf{G}_t = \mathbf{G}_t\mathbf{W}\mathbf{G}_t d\phi_t$ and $d(\mathbf{G}_0\mathbf{G}_0') = (d\mathbf{G}_0)\mathbf{G}_0' + \mathbf{G}_0(d\mathbf{G}_0')$ (see Magnus and Neudecker, 1999, pp. 177 and 183) so that $\boldsymbol{\Upsilon}_t^{SG} = \frac{\partial\mathbf{G}_t\mathbf{G}_t'}{\partial\phi_t} = \mathbf{G}_t(\mathbf{W}\mathbf{G}_t+\mathbf{G}_t'\mathbf{W}')\mathbf{G}_t'$, $t = 0, T$. Furthermore, $\frac{\partial\mathbf{W}\mathbf{G}_0\mathbf{G}_0'}{\partial\phi_0} = \mathbf{W}\boldsymbol{\Upsilon}_0$, $\frac{\partial\mathbf{W}\mathbf{G}_0\mathbf{G}_0'\mathbf{W}'}{\partial\phi_0} = \mathbf{W}\boldsymbol{\Upsilon}_0\mathbf{W}'$.

Under local error spillovers the variance-covariance matrix of the predicted values of GDP per capita at time 0 and T can be derived as

$$\begin{aligned} \boldsymbol{\Sigma}_0 &= E\left[\mathbf{L}_0\varepsilon_0\varepsilon_0'\mathbf{L}_0'\right] = \sigma^2_{\varepsilon_0}\mathbf{L}_0\mathbf{L}_0' \\ \boldsymbol{\Sigma}_T &= \left(\mathbf{I}+T\left(b\mathbf{I}+r\mathbf{W}\right)\right)\mathbf{L}_0\mathbf{L}_0'\left(\mathbf{I}+T\left(b\mathbf{I}+r\mathbf{W}\right)\right)'\sigma^2_{\varepsilon_0}+T^2\mathbf{L}_T\mathbf{L}_T'\sigma_{\varepsilon_T}, \end{aligned} \tag{24}$$

using $\mathbf{q}_T - E[\mathbf{q}_T] = T(b\mathbf{I}+r\mathbf{W})\mathbf{L}_0\varepsilon_0 + T\mathbf{L}_T\varepsilon_T$. The non-linear restriction to test has index SL and at $T=1$ it is given by

$$\begin{aligned}\varphi^{SL}(\theta) &= \left[b(2+b)\tfrac{tr(\mathbf{L}_0\mathbf{L}_0')}{N} + r(1+b)\tfrac{2tr(\mathbf{W}\mathbf{L}_0\mathbf{L}_0')}{N} + r^2\tfrac{tr(\mathbf{W}\mathbf{L}_0\mathbf{L}_0'\mathbf{W}')}{N}\right]\sigma_{\varepsilon_0}^2 \\ &\quad + \tfrac{tr(\mathbf{L}_T\mathbf{L}_T')}{N}\sigma_{\varepsilon_T}^2. \end{aligned} \tag{25}$$

The derivative of this restriction as well as that for the Verdoorn's model are derived in the Appendix.

Under the Verdoorn's model and global spillovers in the error term (with index VG), the variance covariance matrix of $\mathbf{q}_T$ is based on $\Delta\widetilde{\mathbf{q}}_T - E(\Delta\widetilde{\mathbf{q}}_T) = (\mathbf{I}-r\mathbf{W})^{-1}(b(\mathbf{q}_0 - E[\mathbf{q}_0]) + \mathbf{u}_T) = \mathbf{H}(b(\mathbf{q}_0 - E[\mathbf{q}_0]) + \mathbf{u}_T) := \mathbf{H}(b\mathbf{G}_0\varepsilon_0 + \mathbf{G}_T\varepsilon_T)$, using $\mathbf{q}_T - E(\mathbf{q}_T) = T(\Delta\widetilde{\mathbf{q}}_T - E(\Delta\widetilde{\mathbf{q}}_T)) + (\mathbf{q}_0 - E(\mathbf{q}_0)) = T\mathbf{H}(b\mathbf{G}_0\varepsilon_0\mathbf{G}_T\varepsilon_T) + \mathbf{G}_0\varepsilon_0 = (\mathbf{I}+Tb\mathbf{H})\mathbf{G}_0\varepsilon_0 + T\mathbf{H}\mathbf{G}_T\,\varepsilon_T$. At $T=1$ this results in

$$\mathbf{\Sigma}_T = (\mathbf{I}+b\mathbf{H})\mathbf{G}_0\mathbf{G}_0'(\mathbf{I}+b\mathbf{H})'\sigma_{\varepsilon_0}^2 + \mathbf{H}\mathbf{G}_T\mathbf{G}_T'\mathbf{H}'\sigma_{\varepsilon_T}. \tag{26}$$

and the restriction to test in case Verdoorn's model is the true one reads

$$\varphi^{VG}(\theta) = \left[2btr(\mathbf{H}\mathbf{G}_0\mathbf{G}_0') + b^2tr(\mathbf{H}\mathbf{G}_0\mathbf{G}_0'\mathbf{H}')\right]\sigma_{\varepsilon_0}^2 + tr(\mathbf{H}\mathbf{G}_T\mathbf{G}_T'\mathbf{H}')\sigma_{\varepsilon_T}^2. \tag{27}$$

Lastly, for the test of the Verdoorn's model under local error spillovers the variance covariance matrix of the forecasted GDP per capita is given by

$$\mathbf{\Sigma}_T = (\mathbf{I}+Tb\mathbf{H})\mathbf{L}_0\mathbf{L}_0'(\mathbf{I}+Tb\mathbf{H})'\sigma_{\varepsilon_0}^2 + T^2\mathbf{L}_T\mathbf{L}_T'\sigma_{\varepsilon_T}, \tag{28}$$

and

$$\varphi(\theta)^{VL} = \left[2btr(\mathbf{H}\mathbf{L}_0\mathbf{L}_0') + b^2tr(\mathbf{H}\mathbf{L}_0\mathbf{L}_0'\mathbf{H}')\right]\sigma_{\varepsilon_0}^2 + tr(\mathbf{H}\mathbf{L}_T\mathbf{L}_T'\mathbf{H})\sigma_{\varepsilon_T}^2 \tag{29}$$

For all four models the Wald test statistic is defined in the usual way as

$$\mathcal{W}_l = \varphi^l(\widehat{\theta})^2 \left[\mathbf{R}_l'(\widehat{\theta})\widehat{\mathbf{\Omega}}_l(\widehat{\theta})\mathbf{R}_l(\widehat{\theta})\right]^{-1}, \; l = SG, VG, SL, VL \tag{30}$$

and it is distributed as $\chi^2(1)$. It is important to emphasize that the alternative hypothesis is one-sided. Hence, the appropriate critical χ^2-value at significance level α is given at 2α of the χ^2-tabulation. For example, at $\alpha = 0.05$ one has too look at the critical level $\chi^2_{0.1}(1)$ for negative realized values of the restriction under test.

Appendix D provides evidence from Monte Carlo simulations on the performance of the proposed Wald tests. In all experiments the Wald tests perform quite well. Almost all tests are properly sized and exhibit sufficient power, even in moderate sample sizes. Larger samples evidently increase the power of the Wald tests. Compared to other available tests for σ-convergence that ignore spatial correlation in the error term (Carree and Klomp, 1997; Egger and Pfaffermayr, 2006a), the power functions of the proposed maximum likelihood based Wald tests are quite similar. Hence, the power of these tests should be large enough to detect conditional spatial σ-convergence in typical growth studies with about 200 observations if it indeed took place.

Maximum likelihood estimation: In order to derive the Wald test on conditional σ-convergence one has estimate a system that includes an equation that explains the initial values of income per capita and the convergence equation.[6] With respect to the first equation two approaches are pursued here: The first approach argues that under a stationary data generating process the initial val-

[6] In order to test for σ-convergence the variance-covariance matrix of *all* parameters of the econometric model is needed. The current available GMM-estimators (see e.g. Kelejian and Prucha, 2007) and the related tests do not provide the full variance-covariance matrix. Typically, they derive estimates for the slope parameters and the spatial spillover parameters and their variance-covariance matrix. The latter usually excludes the variance estimates, however. This is the main reason why a ML-approach is appropriate here.

ues can be considered as a random deviation from the steady state. The second one ignores the history of the process and assumes that there are no systematic influences present, i.e., $y_{i_0} = \mu + u_{i_0}$. This provides upper and lower bounds of the change in the variance of the predicted values of GDP per capita between period 0 and T.

Following the convergence literature, the initial values of GDP per capita are treated as predetermined, i.e., it is assumed that these do not correlate with the errors of the convergence equation. Since a spatial maximum likelihood approach is applied to estimate the parameters of interest, the paper follows Anselin (1988) and assumes that the error terms of the two equations for a region are identically and independently distributed as bivariate normal:

$$\begin{bmatrix} \varepsilon_{iT} \\ \varepsilon_{i0} \end{bmatrix} \sim iid\ N\left(\mathbf{0}, \begin{bmatrix} \sigma_T^2 & 0 \\ 0 & \sigma_0^2 \end{bmatrix}\right). \tag{31}$$

The likelihood for each considered model is derived in Appendix B. Based on Lee (2004) one gets the asymptotic distribution of the maximum likelihood estimates of the parameters.[7] Since the first order conditions for the ML-estimates cannot be solved analytically, standard numerical optimization procedures are employed to derive the parameters of interest. Estimates of the variance covariance matrix of the parameters are readily available, since numerical estimation procedures usually provide the Hessian matrix. With these estimates at hand standard Wald tests can be applied to test (non-linear) restrictions on the parameters and, in particular, whether conditional σ-convergence has taken place.

[7]More precisely, Lee (2004) derives the asymptotic distribution of a SAR model without spatial spillovers in the error term. His results can be readily extended to the spatial Solow model by imposing a common factor restriction on the parameters to obtain and model with SAR-errors but without an endogenous spatial lag, see Anselin (2003). For the Verdoorn specification asymptotic results seem not to be available.

4 Spatial β- and σ-convergence among European regions

The proposed concept of conditional spatial β- and σ-convergence is applied to investigate the evolution of income per capita (in PPP's) of European NUTS II regions over the period 1980-2002. The dependent variable is the average log difference of real value added per head, where the size of a regions's working population is used in the denominator of this figure.

The set of explanatory variables describing the steady state level of regional GDP per capita and its growth rate follows the literature. The spatial Solow model suggests including the regional investment to GDP ratio as proxy of the savings rate and the growth rate of the labor force as explanatory variables besides initial income and its spatial lag. Lastly, the share of people working in the agricultural sector and the share of employees of the market service sector enter as additional control variables to account for structural differences between regions (see López-Bazo et al., 2004). The same variables explain the initial level of GDP per capita, where the size of the working populations enters instead of its growth rate. With the exception of the investment rate, which is included as period average, the explanatory variables in levels refer to the initial period. The specification of Verdoorn's model is similar. However, it includes the growth rate of real value added instead of the growth rate of the working population, and it substitutes an endogenous spatial lag for the spatially lagged initial income.

The data are provided by Cambridge Econometrics and comprise information on 212 regions of the EU15 as well as the regions of Switzerland and Norway. The observations of the 10 regions in the former German Democratic Republic are not available for 1980 and are not included in the cross-section. In addition, the Portuguese islands Azores have been skipped because of the very large distance from the European continent. Table A1 in the Appendix reports the descriptive statistics of all variables.

The elements of $\mathbf{W}$ are given by $w_{ij} = e^{-d_{ij}/c} / \sum_{j=1}^{N} e^{-d_{ij}/c}$, where c defines the spatial decay. In the basic specification c takes the value 50. However, in Appendix the robustness of the estimation results with respect to this crucial parameter is checked setting $c = 25$ and $c = 75$, respectively (see Tables A2 and A3 in the Appendix).[8] As mentioned above, in case of Verdoorn's model a different weighting scheme is used for disturbances since identification under a common spatial weighting matrix is known to be difficult in small and medium samples. It is a contiguity matrix with $w_{ij} = 1$, if the raw flying distance between regions i and j is smaller than 500 kilometers and zero otherwise. Thus, this spatial weighting scheme does not depend on c.

The density plot in Figure 1 compares the distribution of real income per capita in the initial period (1980) with that in the final period (2002). The figure indicates convergence to a single peaked distribution, but only weak signs of σ-convergence. The per capita income distribution has become less flat, but more skewed towards the right over time.

Figure 1

Table 1 reports the estimates of the β-convergence equation for the four considered models along with estimates of the equation explaining the initial income. Estimations for fixed initial incomes are provided in Table A4 in the Appendix. The error terms of both the initial income equation and the convergence equation in all four estimated models exhibit significant spatial correlation as indicated by the rejection of the corresponding z-tests. This supports the proposed econometric specifications and rules out the application of available tests for conditional

[8]To illustrate this weighting scheme, consider a region with two neighbors. One is located 50 and the other one is located 100 kilometers away. For $c = 50$ the first region gets a spatial weight of 73 percent and the second 27 percent. Under $c = 25$ these figures change to 88 percent and the 12 percent, respectively, while under $c = 75$ the weights are 66 percent and the 34 percent. Hence, the higher c, the smaller the spatial decay and the more weight get the more distant regions.

σ-convergence, which ignore spatial dependence in the error term.

Since all four estimated models exhibit the same number of parameters, the comparison by model selection criteria like AIC and BIC is equivalent to comparing the corresponding likelihoods. As can be seen from Table 1, the models with globally correlated errors outperform those with locally correlated ones. So the subsequent interpretation of the estimation results can be based on the former. Furthermore, this criterion slightly favors the spatial Solow model over Verdoorn's model.

Table 1

The parameter estimates of the spatial Solow model are in line with theory and highly significant. The investment/gross value added ratio as a proxy of the savings rate enters significantly positive. As expected, population growth and the share of agriculture (although not significantly) affect growth negatively, while the share of market services exhibits a positive sign. This finding is in line with López et al. (2004) and suggests that a regions's initial economic structure significantly determines its subsequent growth performance. The estimates for Verdoorn's model indicate a highly significant elasticity of GDP per capita growth with respect to output growth of about 0.5, which is also in line with previous findings (e.g. Fingleton and McCombie, 1998; Fingleton and López-Bazo, 2006). However, its spatial lag turns out to only marginally insignificant. Interestingly, in this model the share of employed people in agriculture is significant, but not that of the service sector.

Similar to many other studies, we observe relatively slow conditional convergence in real income per capita among European regions during the period $1980 - 2002$. For the spatial Solow model the convergence coefficient amounts to -0.015, which implies a speed of convergence of 1.8 percent per year, if spatial spillovers are ignored. This estimate is slightly below the available ones in the literature (see Armstrong, 1995; Neven and Guyette 1995; Carrington, 2003;

López-Bazo et al., 2004; Le Gallo and Dall'erba, 2006). Under the Verdoorn specification convergence is even slower with an estimated convergence parameter of -0.005 or 0.5 percent per year in the absence of regional spillovers. However, spatial spillovers do matter for convergence. In the spatial Solow model a significant spillover parameter of 0.007 has been estimated, while for Verdoorn's model the spillover parameter is 0.26. Note the parameter restriction implied by Verdoorn's model is not rejected in any of the considered models.

As shown above, in a spatial setting the speed of a region's convergence depends on the initial income gap of all other regions as well as on its distance to them. Since the steady state value of initial income per capita remains unidentified in econometric convergence models, the speed of convergence under spatial spillovers can only be calculated in an experiment of thought. For illustration the initial gap in income per capita is assumed to amount to $v_{i0} = q_{i0} - gq_0^{\max}$, $g = 1.5$ and 1.2, respectively. Table 2 calculates the corresponding speed of convergence according to approximations given in (17) and (18) and exhibits a breakdown according to the quartiles of the deviations of initial income from its mean. Also the minimum, maximum and mean are reported for each column.

The spatial Solow model implies that 0.77 percent of the income gap could have been closed per year on average under $g = 1.5$. As expected, the 25 percent quartile with low initial income per capita but a high growth rate exhibits the fastest convergence. Under this model there is considerable variation across regions. At $g = 1.5$, the fastest converging region would have been able to close 0.89 percent of the gap, while the slowest one would achieve a convergence rate of 0.48 percent only. Verdoorn's model implies that 0.72 percent are closed per year. Here, the richest regions get the largest spillovers and converge somewhat faster. However, the heterogeneity is less pronounced; at $g = 1.5$ the difference between the fastest and lowest convergence speed is 0.1 percentage points. The heterogeneity in the speed of convergence, is even more pronounced if the initial gap is smaller on average ($g = 1.2$), ranging from 0.10 percent to 0.97 percent in case of the spatial Solow model and from 0.67 percent to 0.89 under Verdoorn's model.

Overall, this finding suggests that ignoring spatial spillovers may lead to a substantial overestimation of the convergence speed if the spatial Solow model is the true description of the regional growth process. The convergence speed of Verdoorn's model is comparable in size, but differs with respect to its distribution. Under this view, regions take advantage of productivity growth of their neighbors, rather than of their initial level of (human) capital. As a result, the richest regions gain most from knowledge spillovers, since their neighbors are catching up and experience faster productivity growth. This formulation does not seem not to be in line with the stylized facts reported in the knowledge spillover literature (see Keller, 2002).

Table 2

Figure 2 illustrates these findings in a map. It turns out that under the spatial Solow model the regions receiving the lowest growth spillovers tend to be located in the periphery. These regions on the one hand exhibit a large initial income gap, but on the other hand they lack rich advanced neighboring regions. In contrast, the regions with the highest speed of convergence are clumped together in agglomerations with a rich, but slow growing center. Verdoorn's model provides a contrasting picture. The centers of agglomerations benefit most from regional spillovers. Overall, one can conclude that regional spillovers significantly shape the process of convergence among European regions.

Figure 2

Lastly, σ-convergence is at issue. Table 3 summarizes the findings of the Wald test for all estimated models. Under the spatial Solow model, there has been significant conditional σ-convergence with a variance reduction of about 2 percent per year. This finding suggests that the mobility within the distribution of income per capita as captured by $\sigma^2_{\varepsilon_T}$ has been more than outweighed by the decrease in variance due to β-convergence. This becomes evident from the decline in $\frac{1}{N} tr \Sigma_T$. Since the income distribution is single-peaked in 2002 and club-convergence does not seem to be that relevant in our sample of regions, one can conclude that β-percentage convergence has been sufficient to reduce income inequality among the European regions during the eighties and nineties. These findings are robust with respect to variations in the spatial decay in the spatial weighting matrix as captured by the parameter c. Note an increase in c reduces the spatial decay and broadens the regional scope of knowledge spillovers. Using fixed initial incomes rather than conditioning on explanatory variables does not change the conclusion either.

These findings are hardly comparable to the existing literature which mainly investigates β-convergence. Le Gallo and Dall'erba (2006) report evidence on unconditional σ-convergence of labor productivity among European regions based on the test of Carree and Klomp (1997), who consider unconditional convergence and ignore spatial spillovers. Le Gallo and Dall'erba (2006) also calculate the unconditional variances of income per capita taking into account spatial correlation following Rey and Dev (2006) and show that the variance in income per capita decreased between 1975 and 2000, although at a very low rate (see their Figure 1).

The estimates of Verdoorn's model under global error spillovers indicate much higher variances in real income per capita, but a lower reduction in the predicted variance of GDP per capita (1.4 percent per year for $c = 50$) than that found for the spatial Solow model. According to the Wald tests this reduction is not significant and the robust checks reveal a significant reduction of the same magnitude only for a lower spatial decay at $c = 0.25$. This means that mean reversion or β-convergence may be weak under Verdoorn's model and Galton's fallacy can be relevant. However, this model leaves output growth unexplained, while it is endogenously determined in the spatial Solow model. Rather Verdoorn's model conditions on output growth. This implies that we do not observe σ-convergence and, therefore, a shrinking income distribution if all regions experience the same growth rate of output. However, this is hardly a plausible assumption. So an important task for future research is to derive a more general representation of Verdoorn's model, which also explains convergence in output growth.

5 Conclusions

Two views on regional income convergence are prevalent in the literature. The spatial Solow model assumes constant returns to scale and introduces spatial spillovers in terms of spatially weighted initial regional incomes. Verdoorn's model allows for increasing returns to scale and it is shown to be compatible with certain endogenous growth models as well as with models of economic geography. In this model, knowledge spillovers occur due to spatial externalities in productivity growth. While both approaches lead to regional heterogeneity in the implied speed of convergence, their impact on income convergence differs substantially. Under the spatial Solow model regions located near rich centers are catching up faster. Under Verdoorn's model it is the other way round. Regions located near rich centers experience high growth rates as they catch up, and the centers of agglomerations take out most advantages of the spatial spillovers in productivity growth.

For both models convergence equations are estimated for 212 European regions observed over the period 1908-2002. Since the speed of convergence depends on the remoteness of the considered regions and their initial income gaps, which remain unobserved, the speed of convergence of a region can only be inferred in an experiment of thought based on the estimated parameters. This simulation exercise suggests that the average speed of convergence amounts to about 0.7 percent per year, but exhibits considerable variation across regions.

In a spatial setting, conditional σ-convergence has to be redefined, since the predicted variance in income per capita is not uniform across regions due to the spatial spillovers. Following Egger and Pfaffermayr (2006b), conditional σ-convergence is defined as average variance reduction of the predicted income per capita between the initial and the final period of observation. This paper takes a spatial maximum likelihood approach and introduces Wald tests for the hypothesis that the average predicted variance implied by one of the two models remains constant between the initial and the final period. These Wald tests are easily calculated from the output of any numerical optimization procedure, and Monte-Carlo experiments suggest that these tests are properly sized and exhibit enough power in sample sizes commonly used in convergence studies. The application of the Wald tests indicates significant spatial σ-convergence of about 2 percent per year under the spatial Solow specification, once it is conditioned on the explanatory variables usually included in β-convergence regressions. However, under Verdoorn's model the variance reduction is smaller and insignificant.

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.

Ades, Alberto and Hak B. Chua (1997), Thy Neighbors Curse, Regional Instability and Economic Growth, Journal of Economic Growth 2 (3), pp. 279-304.

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

Carree, Martin and Luuk Klomp (1997), Testing the Convergence Hypothesis: A Comment, Review of Economics and Statistics 79 (4), pp. 683-686.

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

Chatterji, Monojit and John H. Dewhurst (1996), Convergence Clubs and Relative Economic Performance in Great Britain: 1977-1991, Regional Studies 30 (1), pp. 31-39.

Ciccone, Antonio and Robert E. Hall (1996), Productivity and the Density of Economic Activity, American Economic Review 86(1), pp. 54-70.

Conley, Timothy G. and Ethan Ligon (2002), Economic Distance and Cross-Country Spillovers, Journal of Economic Growth 7 (2), pp. 157-187.

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

Egger Peter and Michael Pfaffermayr (2006a), On Testing Conditional Sigma-Convergence, Oxford Bulletin of Economics and Statistics, forthcoming.

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

Ertur, Cem, Julie Le Gallo, and Catherine Baumont (2006), The European Regional Convergence Process, 1980-1995: Do Spatial Regimes and Spatial Dependence Matter?, International Regional Science Review 29 (1), pp. 3-34.

Ertur, Cem and Wilfried Koch (2007), Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence, forthcoming in Journal of Applied Econometrics.

Fingleton, Bernard (2001), Theoretical Economic Geography and Spatial Econometrics: Dynamic Perspectives, Journal of Economic Geography 1 (2), pp. 201-225.

Fingleton, Bernard (2003), Externalities, Economic Geography, and Spatial Econometrics: Conceptual and Modelling Developments, International Regional Science Review 26 (2), pp. 197-207.

Fingleton, Bernard (2006), The new economic geography versus urban economics : an evaluation using local wage rates in Great Britain, Oxford Economic Papers 58(3), pp.501-530.

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

Fingleton, Bernard and John S. L. McCombie (1998), Increasing Returns and Economic Growth: Some Evidence for Manufacturing from the European Union Regions, Oxford Economic Papers 50 (1), pp. 89-105.

Furceri, Davide (2005), β-and σ-Convergence: A Mathematical Relation of Causality, Economics Letters 89 (2), pp. 212-215.

Hart, Peter E. (1995), Galtonian Regression across Countries and the Convergence of Productivity, Oxford Bulletin of Economics and Statistics 57 (3), pp. 287-293.

Howitt, Peter (2004), Endogenous Growth and Cross-Country Income Differences, mimeo.

Islam, Nazrul, (2003), What Have We Learnt from the Convergence Debate?, Journal of Economic Surveys 17 (3), pp. 309-362.

Kelejian, Harry H. and Ingmar R. Prucha (2007), Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances, Department of Economics, University of Maryland.

Keller, Wolfgang (2002), Geographic Localization of International Technology Diffusion, American Economic Review 92 (1), pp. 120-142.

Lee, Lung-fei (2004), Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Econometric Models, Econometrica 72(7) pp. 1899-1926.

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.

Magnus Jan R. and Heinz Neudecker (1999), Matrix Differential Calculus with Applications to Statistics and Econometrics, John Wiley, Chechester and New York.

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.

Quah, Danny (1993), Galton's Fallacy and Tests of the Convergence Hypothesis, Scandinavian Journal of Economics 95 (4), pp. 427-443.

Rey, Sergio.J. and Mark V. Janikas (2005), Regional Convergence, Inequality and Space, Journal of Economic Geography 5(2), pp. 155-176.

Rey, Sergio.J. and Boris Dev (2006). Sigma Convergence in the Presence of Spatial Effects, Papers in Regional Science, forthcoming.

Rey, Segio J. and Brett D. Montouri (1999), US Regional Income Convergence: A Spatial Econometric Perspective, Regional Studies 33 (2), pp. 143-156.

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

Wodon, Quentin and Shlomo Yitzhaki (2006), Convergence Forward and Backward?, Economics Letters 92 (1), pp. 47-51.

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

A The linear approximation of the non-linear convergence equations

Regarding the linear approximation of the spatial Solow model let $e^{\mathbf{B}t} = \mathbf{P}^{-1}\widetilde{\mathbf{\Lambda}}^t\mathbf{P}$ with $\widetilde{\mathbf{\Lambda}} = Diag(\frac{\beta(1-\alpha-\rho\lambda_i)}{1-\alpha}) = Diag(\beta + \widetilde{\rho}\lambda_i)$, $\widetilde{\rho} = -\frac{\beta\rho}{1-\alpha}$. $\mathbf{P}$ denotes the eigenvectors of $\mathbf{W}$ and the diagonal matrix $\widetilde{\mathbf{\Lambda}}$ comprises the eigenvalues of $\mathbf{B}$. Lastly, λ_i denote the eigenvalues of $\mathbf{W}$. One can approximate $e^{\beta T + \widetilde{\rho}\lambda_i T} \approx e^{\beta T}(1 + \widetilde{\rho}\lambda_i T)$. This approximation is viable provided that $|\widetilde{\rho}\lambda_i T| \leq |\widetilde{\rho}|\,|\lambda_i|\,T < |\widetilde{\rho}|\,T < 1$ or $T < \frac{1}{|\widetilde{\rho}|}$, using $|\lambda_i| < 1$ in case of row normalized spatial weighting matrices (see Egger and Pfaffermayr, 2006b for a similar approach). Under this condition the higher order terms tend to zero fast and the first order term dominates. Whether this condition holds empirically can easily be checked using plausible values of β, ρ and α. In the empirical analysis below it turns out that $\rho = 0.007$ and $\beta = -0.015$. Assuming $\alpha = .33$ implies $|\widetilde{\rho}| = 0.0002$, so this restriction is fulfilled for any plausible value of T. Inserting gives $-\frac{1}{T}\left(\mathbf{I} - e^{\mathbf{B}t}\right) = -\frac{1}{T}\left(\mathbf{I} - \mathbf{P}^{-1}e^{\widetilde{\mathbf{\Lambda}}t}\mathbf{P}\right) \approx -\frac{1}{T}\left(\mathbf{I} - \mathbf{P}^{-1}Diag[e^{\beta T}(1 + \widetilde{\rho}\lambda_i T)]\mathbf{P}\right) = -\frac{1}{T}(1 - e^{\beta T})\mathbf{I} + \widetilde{\rho}e^{\beta T}\mathbf{P}^{-1}\mathbf{\Lambda}\mathbf{P}] = -\frac{1}{T}(1 - e^{\beta T})\mathbf{I} + \widetilde{\rho}e^{\beta T}\mathbf{W} := b\mathbf{I} + r\mathbf{W}$, using $\mathbf{P}^{-1}\mathbf{\Lambda}\mathbf{P} = \mathbf{W}$. Therefore, one may interpret the estimated coefficients as $b = -\frac{1}{T}(1 - e^{\beta T})$, resulting in $\beta = \frac{1}{T}\ln(1 + bT)$, and $r = \widetilde{\rho}e^{\beta T} = -\frac{\beta\rho}{1-\alpha}e^{\beta T}$ so that $\rho = -\frac{r(1-\alpha)T}{(1+bT)\ln(1+bT)}$. In the absence of spillovers at $r = \rho = 0$ the convergence term reverts to that derived in Barro and Sala-i-Martin (2004). In sum, the linear approximation to the spatial Solow model is given by

$$
\begin{aligned}
\mathbf{q}_0 &= \mathbf{X}_0\delta_0 + \mathbf{u}_0 && (32)\\
\Delta\widetilde{\mathbf{q}}_T &= \Delta\tfrac{1}{T}\mathbf{q}_T = b\mathbf{q}_0 + r\mathbf{W}\mathbf{y}_0 + \mathbf{X}_T\delta_T + \mathbf{u}_T.
\end{aligned}
$$

The parametrization of Verdoorn's model can be based on the approximation of $e^{\frac{\beta T}{1-\rho\lambda_i}}$ around $\beta = 0$: $e^{\frac{\beta T}{1-\rho\lambda_i}} \approx 1 + \frac{\beta T}{1-\rho\lambda_i} = 1 + \beta T \sum_{k=0}^{\infty} \rho^k \lambda_i^k$. For this approximation to be viable with a dominating first order term (i.e. $\left|\frac{\delta T}{1-\rho\lambda_i}\right| < 1$), it must hold that $1 > -\frac{\beta T}{1-\rho} > -\frac{\beta T}{1-\rho\lambda_i}$ or $T < -\frac{1-\rho}{\beta}$ using $\beta < 0$, $\rho > 0$, $1 - \rho\lambda_i > 1 - \rho|\lambda_i| > 1 - \rho$ and $|\lambda_i| < 1$. Again, for plausible values of ρ and β this restriction is fulfilled for values of T usually used in empirical applications. Inserting gives $-\frac{1}{T}\left(\mathbf{I} - e^{\mathbf{C}T}\right) = -\frac{1}{T}\mathbf{P}^{-1}\left(Diag[1 - e^{\frac{\beta T}{1-\rho\lambda_i}}]\right)\mathbf{P} \approx \frac{1}{T}\mathbf{P}^{-1}\left(Diag(\beta T \sum_{k=0}^{\infty} \rho^k \lambda_i^k)\right)\mathbf{P} = \beta\left(\sum_{k=0}^{\infty} \rho^k \mathbf{P}^{-1}\mathbf{\Lambda}^k\mathbf{P}\right) = \beta(\mathbf{I} - \rho\mathbf{W})^{-1}$, using $\mathbf{P}^{-1}\mathbf{\Lambda}^k\mathbf{P} = \mathbf{W}^k$. Hence, the specification to be estimated is given by

$$
\begin{aligned}
\Delta\widetilde{\mathbf{q}}_T &= \beta(\mathbf{I} - \rho\mathbf{W})^{-1}(\widetilde{\mathbf{q}}_0 - \widetilde{\mathbf{q}}_0^*) + \Delta\widetilde{\mathbf{q}}_T^* \quad \text{or} \\
\Delta\widetilde{\mathbf{q}}_T &= \beta(\mathbf{q}_0 - \mathbf{q}_0^*) + \rho\mathbf{W}\Delta\widetilde{\mathbf{q}}_T + \gamma(\mathbf{I} - \rho\mathbf{W})\Delta\widetilde{\mathbf{Y}}_T + \delta_1 + \Delta\widetilde{\mathbf{Z}}\delta_2.
\end{aligned} \tag{33}
$$

using $(\mathbf{I} - \rho\mathbf{W})\Delta\widetilde{\mathbf{q}}_T^* = \gamma(\mathbf{I} - \rho\mathbf{W})\Delta\widetilde{\mathbf{Y}}_T + \delta_1 + \Delta\widetilde{\mathbf{Z}}\delta_2$. The system to be estimated reads

$$
\begin{aligned}
\mathbf{q}_0 &= \mathbf{X}_0\delta_0 + \mathbf{u}_0 \\
\Delta\widetilde{\mathbf{q}}_T &= b\mathbf{q}_0 + r\mathbf{W}\Delta\widetilde{\mathbf{q}}_T + \gamma\Delta\widetilde{\mathbf{Y}}_T + \phi\mathbf{W}\Delta\widetilde{\mathbf{Y}}_T + \mathbf{X}_T\delta_T + \mathbf{u}_T.
\end{aligned} \tag{34}
$$

Based on this approximation the parameters can be estimated directly, i.e. $b = \beta$ and $r = \rho$.

B Derivations of the likelihood functions

B.1 Spatial Solow model - global error spillovers

1. Observe that $E[\mathbf{u}_0\mathbf{u}_0'] = \sigma_0^2\mathbf{G}_0\mathbf{G}_0'$ and $E[\mathbf{u}_T\mathbf{u}_T'] = \sigma_T^2\mathbf{G}_T\mathbf{G}_T'$ so that

$$
\mathbf{\Omega}_G = E\left[\begin{pmatrix} \mathbf{u}_0 \\ \mathbf{u}_T \end{pmatrix}(\mathbf{u}_0, \mathbf{u}_T)\right] = \begin{bmatrix} \sigma_0^2\mathbf{G}_0\mathbf{G}_0' & \mathbf{0} \\ \mathbf{0} & \sigma_T^2\mathbf{G}_T\mathbf{G}_T' \end{bmatrix}
$$

2. It is useful to introduce the following matrices (see Anselin, 1988):

$$\mathbf{\Phi} = \begin{bmatrix} \phi_0 & 0 \\ 0 & \phi_T \end{bmatrix}, \mathbf{\Sigma} = \begin{bmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_T^2 \end{bmatrix}$$

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

$$\begin{aligned}
\mathbf{G} &= (\mathbf{I}_{2N} - (\mathbf{\Phi} \otimes \mathbf{W}))^{-1} \\
\mathbf{L} &= (\mathbf{I}_{2N} - (\mathbf{\Phi} \otimes \mathbf{W})) \\
\mathbf{\Omega}_G &= \mathbf{G}(\mathbf{\Sigma} \otimes \mathbf{I}_N)\mathbf{G}' \\
\mathbf{\Omega}_G^{-1} &= \mathbf{L}(\mathbf{\Sigma}^{-1} \otimes \mathbf{I}_N)\mathbf{L}' \\
-\tfrac{1}{2} \ln \det \mathbf{\Omega}_G &= -\tfrac{N}{2} \ln(\det \Sigma) + \det(\mathbf{L}) \\
&= -\tfrac{N}{2} \ln(\sigma_T^2 \sigma_0^2) + \ln \det \mathbf{L}_0 + \ln \det \mathbf{L}_T
\end{aligned}$$

3. For the structural form let

$$\begin{aligned}
\mathbf{Y} &= \begin{bmatrix} \mathbf{q}_0 \\ \Delta\widetilde{\mathbf{q}}_T \end{bmatrix}, \ \mathbf{\Gamma}^S = \begin{bmatrix} \mathbf{I}_N & \mathbf{0} \\ -b\mathbf{I}_N - r\mathbf{W} & \mathbf{I}_N \end{bmatrix}, \ \mathbf{X} = \begin{bmatrix} \mathbf{X}_0 & 0 \\ \mathbf{0} & \mathbf{X}_T \end{bmatrix} \\
\varepsilon' \mathbf{\Omega}_G^{-1} \varepsilon &= (\mathbf{\Gamma}^S \mathbf{Y} - \mathbf{X}\beta)' \mathbf{L}' \left(\begin{bmatrix} \sigma_T^2 & 0 \\ 0 & \sigma_0^2 \end{bmatrix}^{-1} \otimes \mathbf{I}_N \right) \mathbf{L}(\mathbf{\Gamma}^S \mathbf{Y} - \mathbf{X}\beta) \\
&= (\xi_{0,G}', \xi_{T,SG}') \begin{bmatrix} \sigma_0^2 \mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & \sigma_T^2 \mathbf{I}_N \end{bmatrix} (\xi_{0,G}', \xi_{T,SG}')' \\
&= [\tfrac{\xi_{0,G}' \xi_{0,G}}{\sigma_0^2} + \tfrac{\xi_{T,SG}' \xi_{T,SG}}{\sigma_T^2}],
\end{aligned}$$

where $\xi_{0,G} = (\mathbf{I}_N - \phi_0 \mathbf{W})(\mathbf{q}_0 - \mathbf{X}_0 \beta_0)$ and $\xi_{T,SG} = (\mathbf{I}_N - \phi_T \mathbf{W})(\Delta\widetilde{\mathbf{q}}_T - b\mathbf{q}_0 - r\mathbf{W}\mathbf{q}_0 -$

$\mathbf{X}_T\beta_T)$. The log likelihood function is then given by

$$\begin{aligned}
\ln L^{SG} &= -\tfrac{1}{2}\ln\det(\mathbf{\Omega}_G) - \tfrac{1}{2}(\mathbf{\Gamma}^S\mathbf{Y} - \mathbf{X}\beta)'\mathbf{\Omega}_G^{-1}(\mathbf{\Gamma}^S\mathbf{Y} - \mathbf{X}\beta) \\
&= -\tfrac{N}{2}\ln(\sigma_u^2\sigma_v^2) + \ln\det(\mathbf{I}_N - \phi_T\mathbf{W}) + \ln\det(\mathbf{I}_N - \phi_0\mathbf{W}) \\
&\quad -\tfrac{1}{2}[\tfrac{\xi_{0,G}'\xi_{0,G}}{\sigma_0^2} + \tfrac{\xi_{T,SG}'\xi_{T,SG}}{\sigma_T^2}].
\end{aligned}$$

B.2 Spatial Solow model - local error spillovers

Under local error spillovers the variance covariance matrix of the error term is given by

$$\begin{aligned}
\mathbf{\Omega}_L &= \mathbf{L}(\mathbf{\Sigma} \otimes \mathbf{I}_N)\mathbf{L}' \\
\mathbf{\Omega}_L^{-1} &= \mathbf{G}(\mathbf{\Sigma}^{-1} \otimes \mathbf{I}_N)\mathbf{G}' \\
-\tfrac{1}{2}\ln\det\mathbf{\Omega}_L &= -\tfrac{N}{2}\ln(\det\Sigma) + \det(\mathbf{G}) \\
&= -\tfrac{N}{2}\ln(\sigma_T^2\sigma_0^2) - \ln\det\mathbf{G}_0 - \ln\det\mathbf{G}_T
\end{aligned}$$

and

$$\begin{aligned}
\ln L^{SL} &= -\tfrac{1}{2}\ln\det(\mathbf{\Omega}_L) - \tfrac{1}{2}(\mathbf{\Gamma}^S\mathbf{Y} - \mathbf{X}\beta)'\mathbf{\Omega}_L^{-1}(\mathbf{\Gamma}^S\mathbf{Y} - \mathbf{X}\beta) \\
&= -\tfrac{N}{2}\ln(\sigma_u^2\sigma_v^2) - \ln\det(\mathbf{I}_N - \phi_T\mathbf{W}) - \ln\det(\mathbf{I}_N - \phi_0\mathbf{W}) \\
&\quad -\tfrac{1}{2}[\tfrac{\xi_{0,L}'\xi_{0,L}}{\sigma_0^2} + \tfrac{\xi_{T,SL}'\xi_{T,SL}}{\sigma_T^2}].
\end{aligned}$$

where $\xi_{0,L} = (\mathbf{I}_N - \phi_0\mathbf{W})^{-1}(\mathbf{q}_0 - \mathbf{X}_0\beta_0)$ and $\xi_{T,VG} = (\mathbf{I}_N - \phi_T\mathbf{W})^{-1}(\Delta\widetilde{\mathbf{q}}_T - b\mathbf{q}_0 - r\mathbf{W}\Delta\mathbf{q}_0 - \mathbf{X}_T\beta_T)$.

B.3 Verdoorn's model - global error spillovers

The structural model for Verdoorn's law uses

$$\begin{aligned}
\mathbf{\Gamma}^V &= \begin{bmatrix} \mathbf{I}_N & \mathbf{0} \\ -b\mathbf{I}_N & \mathbf{I}_N - r\mathbf{W} \end{bmatrix} \\
\varepsilon'\mathbf{\Omega}_V^{-1}\varepsilon &= (\mathbf{\Gamma}^V\mathbf{Y} - \mathbf{X}\beta)'\mathbf{L}'\left(\begin{bmatrix} \sigma_T^2 & \mathbf{0} \\ \mathbf{0} & \sigma_0^2 \end{bmatrix} \otimes \mathbf{I}_N\right)\mathbf{L}(\mathbf{\Gamma}^V\mathbf{Y} - \mathbf{X}\beta).
\end{aligned}$$

Note the only difference to the Solow model lies in definition of $\mathbf{\Gamma}$ and $\mathbf{X}$ and the same likelihood as above applies using $\xi_{0,G} = (\mathbf{I}_N - \phi_0\mathbf{W})(\mathbf{q}_0 - \mathbf{X}_0\beta_0)$ and $\xi_{T,VG} = (\mathbf{I}_N - \phi_T\mathbf{W})(\Delta\widetilde{\mathbf{q}}_T - b\mathbf{q}_0 - r\mathbf{W}\Delta\mathbf{q}_0 - \mathbf{X}_T\beta_T)$.

$$\begin{aligned}
\ln L^{VG} &= -\tfrac{N}{2}\ln(\sigma_u^2\sigma_v^2) + \ln\det(\mathbf{I}_N - \phi_T\mathbf{W}) + \ln\det(\mathbf{I}_N - \phi_0\mathbf{W}) \\
&\quad + \ln\det(\mathbf{I}_N - r\mathbf{W}) - \tfrac{1}{2}[\tfrac{\xi_{0,G}'\xi_{0,G}}{\sigma_0^2} + \tfrac{\xi_{T,VG}'\xi_{T,VG}}{\sigma_T^2}].
\end{aligned}$$

B.4 Verdoorn's model - local error spillovers

Lastly under the Verdoorn model with local spillovers in the error term, the likelihood function is given by

$$\begin{aligned}
\ln L^{VL} &= -\tfrac{N}{2}\ln(\sigma_u^2\sigma_v^2) - \ln\det(\mathbf{I}_N - \phi_T\mathbf{W}) - \ln\det(\mathbf{I}_N - \phi_0\mathbf{W}) \\
&\quad + \ln\det(\mathbf{I}_N - r\mathbf{W}) - \tfrac{1}{2}[\tfrac{\xi_{0,L}'\xi_{0,L}}{\sigma_0^2} + \tfrac{\xi_{T,VL}'\xi_{T,VL}}{\sigma_T^2}].
\end{aligned}$$

where $\xi_{0,L} = (\mathbf{I}_N - \phi_0\mathbf{W})^{-1}(\mathbf{q}_0 - \mathbf{X}_0\beta_0)$ and $\xi_{T,VL} = (\mathbf{I}_N - \phi_T\mathbf{W})^{-1}(\Delta\widetilde{\mathbf{q}}_T - b\mathbf{q}_0 - r\mathbf{W}\Delta\mathbf{q}_0 - \mathbf{X}_T\beta_T)$.

C The gradients of the restrictions of the Wald test for conditional σ-convergence

C.1 Spatial Solow model - local error spillovers

The restriction to test is

$$\begin{aligned}\varphi(\theta)^{SL} &= \left[b(2+b)\frac{tr(\mathbf{L}_0\mathbf{L}_0')}{N}+r(1+b)\frac{2tr(\mathbf{W}\mathbf{L}_0\mathbf{L}_0')}{N}+r^2\frac{tr(\mathbf{W}\mathbf{L}_0\mathbf{L}_0'\mathbf{W}')}{N}\right]\sigma^2_{\varepsilon_0}\\ &\quad+\frac{tr(\mathbf{L}_T\mathbf{L}_T')}{N}\sigma^2_{\varepsilon_T},\end{aligned}$$

with gradient

$$\mathbf{R}^{SL}=\begin{bmatrix}\mathbf{0}\\ 2\left[(1+b)\frac{tr(\mathbf{L}_0\mathbf{L}_0')}{N}+r\frac{tr(\mathbf{W}\mathbf{L}_0\mathbf{L}_0')}{N})\right]\sigma^2_{\varepsilon_0}\\ 2\left[(1+b)\frac{tr(\mathbf{W}\mathbf{L}_0\mathbf{L}_0')}{N}+r\frac{tr(\mathbf{W}\mathbf{L}_0\mathbf{L}_0'\mathbf{W}')}{N}\right]\sigma^2_{\varepsilon_0}\\ b(2+b)\frac{tr(\mathbf{L}_0\mathbf{L}_0')}{N}+r(1+b)\frac{2tr(\mathbf{W}\mathbf{L}_0\mathbf{L}_0')}{N}+r^2\frac{tr(\mathbf{W}\mathbf{L}_0\mathbf{L}_0'\mathbf{W}')}{N}\\ \frac{tr(\mathbf{L}_T\mathbf{L}_T')}{N}\\ \left[b(2+b)\frac{tr(\boldsymbol{\Upsilon}_0^{SL})}{N}+r(1+b)\frac{tr(\mathbf{W}\boldsymbol{\Upsilon}_0^{SL})}{N}+r^2\frac{tr(\boldsymbol{\Upsilon}_0^{SL}\mathbf{W}'\mathbf{W})}{N}\right]\sigma^2_{\varepsilon_0}\\ \frac{tr(\boldsymbol{\Upsilon}_T^{SL})}{N}\sigma^2_{\varepsilon_T}\end{bmatrix},$$

using $\boldsymbol{\Upsilon}_t^{SL}=\frac{\partial\mathbf{L}_t\mathbf{L}_t'}{\partial\phi_t}=-(\mathbf{W}+\mathbf{W}')+2\phi_t\mathbf{W}\mathbf{W}'$ and $\frac{\partial(\mathbf{L}_t\mathbf{L}_t'\mathbf{A})}{\partial\phi_t}=\boldsymbol{\Upsilon}_t^{SL}\mathbf{A}$ where $\mathbf{A}$ is either $\mathbf{W}$ or $\mathbf{W}'\mathbf{W}$.

C.2 Verdoorn's model - global error spillovers

For convenience the non-linear restriction to be tested is restated here:

$$\varphi(\theta)^{VG}=\left[2btr(\mathbf{H}\mathbf{G}_0\mathbf{G}_0')+b^2tr(\mathbf{H}\mathbf{G}_0\mathbf{G}_0'\mathbf{H}')\right]\sigma^2_{\varepsilon_0}+tr(\mathbf{H}\mathbf{G}_T\mathbf{G}_T'\mathbf{H}')\sigma^2_{\varepsilon_T}$$

Its gradient is given by

$$\mathbf{R}^{VG} = \begin{bmatrix} \mathbf{0} \\ 2\left[\frac{tr(\mathbf{HG}_0\mathbf{G}_0'))}{N} + b\frac{tr(\mathbf{HG}_0\mathbf{G}_0'\mathbf{H}')}{N})\right]\sigma^2_{\varepsilon_0} \\ \left[2b\frac{tr(\mathbf{\Gamma}^{VG}\mathbf{G}_0\mathbf{G}_0')}{N} + b^2\frac{tr(\mathbf{\Psi}^{VG}\mathbf{G}_0\mathbf{G}_0')}{N}\right]\sigma^2_{\varepsilon_0} + \frac{tr(\mathbf{\Psi}^{VG}\mathbf{G}_T\mathbf{G}_T')}{N}\sigma^2_{\varepsilon_T} \\ 2b\frac{tr(\mathbf{HG}_0\mathbf{G}_0')}{N} + b^2\frac{tr(\mathbf{HG}_0\mathbf{G}_0'\mathbf{H}')}{N} \\ \frac{tr(\mathbf{HG}_T\mathbf{G}_T'\mathbf{H})}{N} \\ \left[2b\frac{tr(\mathbf{H\Upsilon}_0^{VG})}{N} + b^2\frac{tr(\mathbf{H\Upsilon}_0^{VG}\mathbf{H}')}{N}\right]\sigma^2_{\varepsilon_0} \\ \frac{tr(\mathbf{\Upsilon}_T^{VG})}{N}\sigma^2_{\varepsilon_T} \end{bmatrix},$$

using $\mathbf{\Gamma}^{VG} = \frac{\partial \mathbf{H}}{\partial r} = \mathbf{HWH}$, $\Psi^{VG} = \frac{\partial \mathbf{H}'\mathbf{H}}{\partial r} = \mathbf{HWHH}' + \mathbf{HH}'\mathbf{W}'\mathbf{H}' = \mathbf{H}(\mathbf{WH}+\mathbf{H}'\mathbf{W}')\mathbf{H}'$ and $\Upsilon_t^{VG} = \frac{\partial \mathbf{G}_t\mathbf{G}_t'}{\partial \phi_t} = \mathbf{G}_t(\mathbf{W}'\mathbf{G}_t+\mathbf{G}_t'\mathbf{W})\mathbf{G}_t'$ so that $\frac{\partial \mathbf{HG}_0\mathbf{G}_0'}{\partial r} = \Gamma^{VG}\mathbf{G}_0\mathbf{G}_0'$ and $\frac{\partial \mathbf{HG}_t\mathbf{G}_t'\mathbf{H}'}{\partial}r = \Psi^{VG}\mathbf{G}_t\mathbf{G}_t'$.

C.3 Verdoorn's model - local error spillovers

The restriction to test is

$$\varphi(\theta)^{VL} = \left[2btr(\mathbf{HL}_0\mathbf{L}_0') + b^2 tr(\mathbf{HL}_0\mathbf{L}_0'\mathbf{H}')\right]\sigma^2_{\varepsilon_0} + tr(\mathbf{HL}_T\mathbf{L}_T'\mathbf{H}')\sigma^2_{\varepsilon_T}$$

and

$$\mathbf{R}^{VL} = \begin{bmatrix} \mathbf{0} \\ 2\left[\frac{tr(\mathbf{HL}_0\mathbf{L}_0'))}{N} + b\frac{tr(\mathbf{HL}_0\mathbf{L}_0'\mathbf{H}')}{N})\right]\sigma^2_{\varepsilon_0} \\ \left[2b\frac{tr(\mathbf{\Gamma}^{VL}\mathbf{L}_0\mathbf{L}_0')}{N} + b^2\frac{tr(\mathbf{\Psi}^{VL}\mathbf{L}_0\mathbf{L}_0')}{N}\right]\sigma^2_{\varepsilon_0} + \frac{tr(\mathbf{\Psi}^{VL}\mathbf{L}_T\mathbf{L}_T')}{N}\sigma^2_{\varepsilon_T} \\ 2b\frac{tr(\mathbf{HL}_0\mathbf{L}_0')}{N} + b^2\frac{tr(\mathbf{HL}_0\mathbf{L}_0'\mathbf{H}')}{N} \\ \frac{tr(\mathbf{L}_T\mathbf{L}_T')}{N} \\ \left[2b\frac{tr(\mathbf{H\Upsilon}_0^{VL})}{N} + b^2\frac{tr(\mathbf{H\Upsilon}_0^{VL}\mathbf{H}')}{N}\right]\sigma^2_{\varepsilon_0} \\ \frac{tr(\mathbf{\Upsilon}_T^{VG})}{N}\sigma^2_{\varepsilon_T} \end{bmatrix},$$

using $\mathbf{\Gamma}^{VL} = \frac{\partial \mathbf{H}}{\partial r} = \mathbf{HWH}$, $\mathbf{\Psi}^{VL} = \frac{\partial \mathbf{H}'\mathbf{H}}{\partial r} = \mathbf{H}'\mathbf{W}'\mathbf{H}'\mathbf{H}+ \mathbf{H}'\mathbf{HWH} = \mathbf{H}'(\mathbf{W}'\mathbf{H}'+ \mathbf{HW})\mathbf{H}$ and $\Upsilon_t^{VG} = \frac{\partial \mathbf{L}_t\mathbf{L}_t'}{\partial \phi_t} = -(\mathbf{W}+\mathbf{W}') + 2\phi_t\mathbf{WW}'$ so that $\frac{\partial \mathbf{HL}_0\mathbf{L}_0'}{\partial r} = \Gamma^{VG}\mathbf{L}_0\mathbf{L}_0'$ and $\frac{\partial \mathbf{H}'\mathbf{HL}_t\mathbf{L}_t'}{\partial r} = \Psi^{VG}\mathbf{L}_t\mathbf{L}_t'$.

D Descriptive statistics and the robustness of the estimation results

Table A1 reports the descriptive statistics of the variables.

Tables A1

Tables A2 and A2 assess the robustness of the estimated convergence equations with respect to the spatial decay as captured by the parameter c. Stetting $c = 25$ (faster decay with a lower reach of regional spillovers) and $c = 75$ (slower decay with a broader reach of regional spillovers), shows that the estimation results do not change much. However, for the spatial Solow model, we observe that the spatial correlation of the error term increases in c as expected.

Tables A2 and A3

Lastly, Table A4 reports the estimation results for the basic models under fixed starting values.

Tables A4

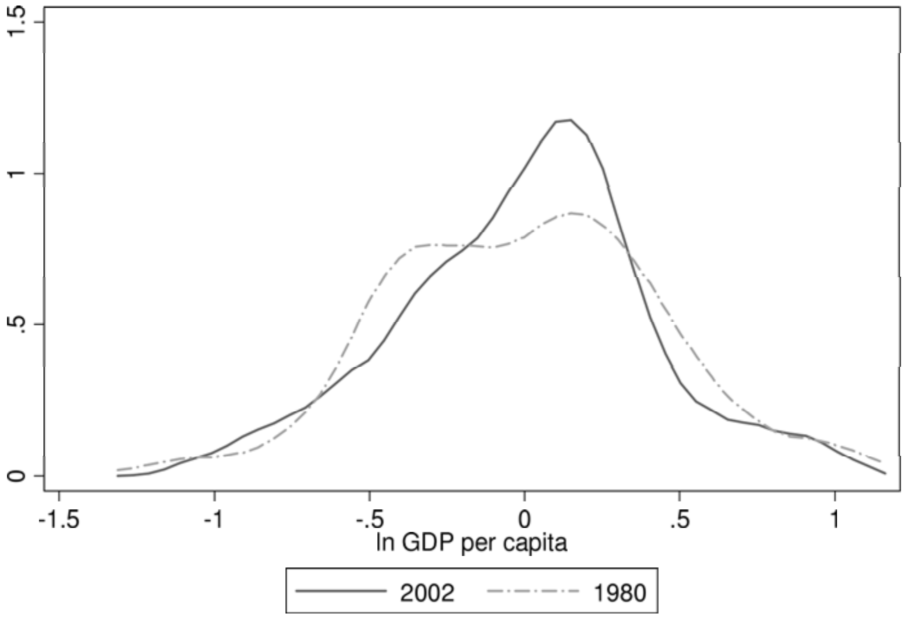
E Monte Carlo Simulation

The size and power of the proposed Wald tests for conditional spatial σ-convergence in small samples is examined in a Monte-Carlo experiment. For parameter the combinations $\{N, b, \phi_0, \phi_T\}$ 1000 replications are performed, where $N \epsilon \{250, 500\}$ and $\phi_0, \phi_0 \epsilon \{0, 0.2, 0.4\}$. The power functions are simulated at a size of 5 percent over a range of b-values that slightly changes from experiment to experiment. The chosen parameter values for the remaining parameters are roughly in line with the empirical studies on regional convergence: $r = 0.025$ under the spatial Solow model and $r = 0.3$ under Verdoorn's model, $\sigma_0^2 = 0.1$, $\sigma_T^2 = 0.0025$. For example, at $\phi_0 = 0$, $\phi_T = 0.2$ and $N = 250$ the null holds true at $b = -.012723$. The spatial weighting matrix is derived by allocating observations randomly on a grid of $0.5 * N$ squares of length 10. So the number of squares in the grid grows larger if the number of observations N increases. This guarantees decreasing spatial correlation at an increasing sample size. Each square includes only one observation and the probability that an observation is located on a particular coordinate is equal for all squares. The procedure guarantees that each data point possesses unique coordinates. The spatial weights are calculated as $e^{d_{ij}/10} / \sum_{j=1}^{N} e^{d_{ij}/10}$, where d_{ij} denotes the distance between observation i and j. Note, this spatial weighting matrix is row-normalized. Under Verdoorn's model the spatial weights used in the error term are different and based on a row normalized Queens design.

Table A5

In all experiments of Table A5, the Wald test performs quite well. Almost all tests are properly sized and exhibit sufficient power, even in the smaller sample at $N = 250$. For instance, under the spatial Solow model with globally correlated errors a deviation of -0.03 from the value at which H_0 holds true leads to a rejection rate of 0.63 for $N = 250$ and $\phi_0 = \phi_T = 0.2$), which reaches 0.86 at $N = 500$. The corresponding rejection rates the other models are similar. A larger sample size evidently increases the power of the Wald tests.

Figure 1: Convergence in income per capita in PPPs among European regions, 1980-2002, deviations from means



Source: Cambridge Econometrics

Figure 2a: Speed of convergence, spatial Solow model (share of the gap closed in 20 years)

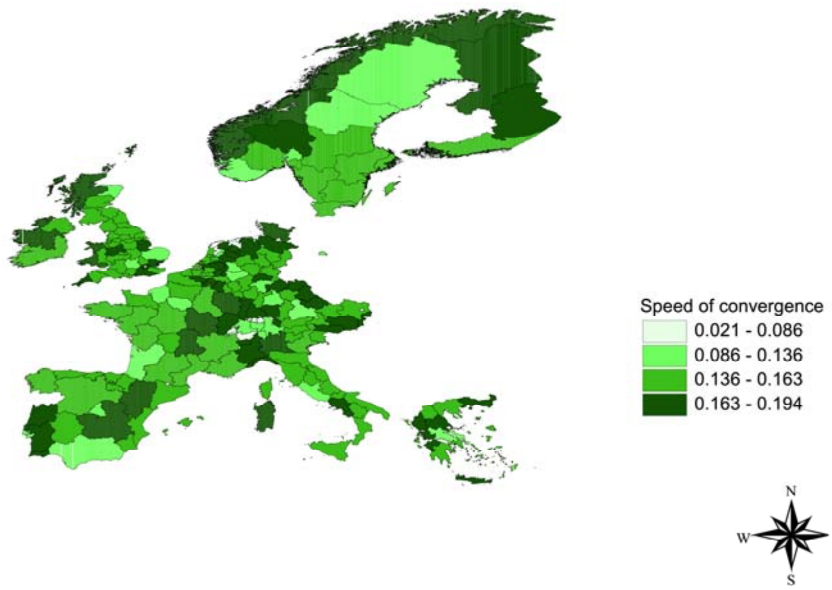


Figure 2b: Speed of convergence, spatial Verdoorn model (share of the gap closed in 20 years)

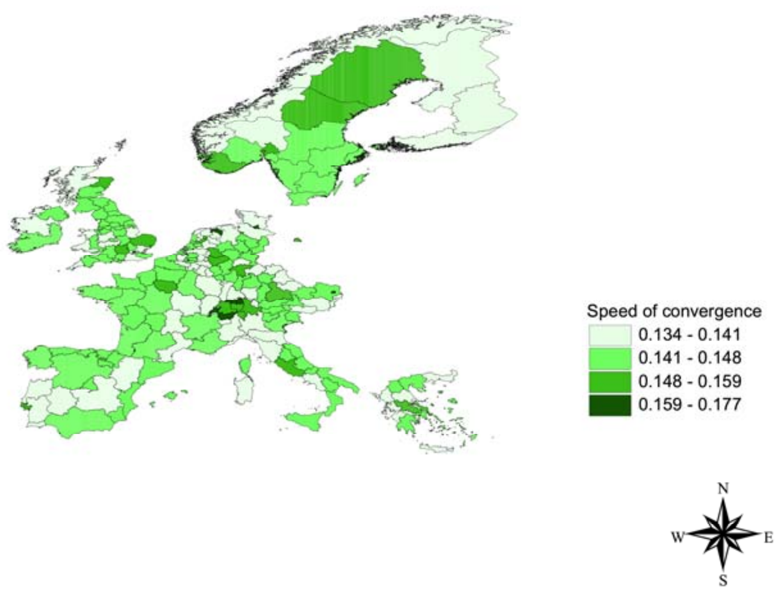


Table 1: β-convergence among European regions: 1980-2002, c=50

	Spatial Solow model, global error spillovers		Spatial Solow model, local error spillovers		Verdoorn's model, global error spillovers		Verdoorn's model, local error spillovers	
	b	z	b	z	b	z	b	z
Initial GDP per capita equation								
Initial working population	0.019	1.2	0.000	0.0	0.001	0.1	-0.014	-0.4
Initial investment rate	-0.117	-0.4	0.226	0.8	-0.026	-1.2	0.137	0.3
Initial share of agriculture	-0.353	-1.4	-1.130	-4.3 ***	-0.397	-1.9	-1.130	-1.5 #
Initial share of market services	1.190	4.6 ***	0.657	2.2 **	0.986	3.0 ***	0.655	1.8 *
Constant	2.394	10.6 ***	2.793	13.6 ***	1.290	1.1	2.885	9.0 ***
$\sigma_{\varepsilon 0}$	0.215		0.283		0.281		0.307	
ϕ_0	0.884	25.6 ***	-0.963	-13.5 ***	0.974	49.1 ***	-0.990	-1.2
Growth of real income per capita equation								
Initial real income per capita	-0.015	-9.5 ***	-0.015	-9.6 ***	-0.005	-5.5 ***	-0.006	-5.3 **
Initial real income per capita - spatially weighted	0.007	3.0 ***	0.008	3.6 ***	-	-	-	-
Growth of working population	-0.277	-4.57 ***	-0.264	-4.16 ***	-	-	-	-
Initial investment rate	0.050	4.0 ***	0.049	3.6 ***	-	-	-	-
Growth inreal income per capita - spatially weighted	-	-	-	-	0.255	2.6 ***	0.250	2.0 **
Growth of real income	-	-	-	-	0.534	14.8 ***	0.533	13.5 ***
Growth of real income - spatially weighted	-	-	-	-	-0.135	-1.5 #	-0.130	-1.0
Initial share of agriculture	-0.002	-0.4	-0.006	-1.0	-0.010	-2.5 **	-0.010	-2.1 **
Initial share of market services	0.041	5.7 ***	0.040	5.6 ***	0.003	0.6	0.004	0.6
Constant	0.015	2.0 **	0.015	2.2 **	0.019	4.2 ***	0.019	4.4 ***
$\sigma_{\varepsilon T}$	0.006		0.006		0.005		0.005	
ϕ_T	0.566	6.4 ***	-0.659	-4.7 ***	0.508	2.2 **	-0.990	-2.0 **
Likelihood	-950.1		-984.0		-963.3		-974.5	

Note: *** significant at 1%; ** significant at 5%; * significant at 10%; # significant at 15%.

Table 2: Speed of convergence under spatial spillovers, share of the income gap closed within a year

			Gap_1		Gap_2	
	Income growth rate	Initial Income	Spatial Solow model	Spatial Verdoorn model	Spatial Solow model	Spatial Verdoorn model
25 % Quartil	0.83	-26.80	0.71	0.69	0.67	0.67
50% Quartil	-1.00	4.36	0.60	0.69	0.48	0.67
75 % Quartil	-1.34	33.56	0.48	0.71	0.10	0.70
Min	-1.34	-69.90	0.48	0.69	0.10	0.67
Max	5.17	175.49	0.89	0.79	0.97	0.89
Mean	1.47	9.37	0.77	0.72	0.76	0.72

Note: Gap_1 is defined as y_0-1.5max(y_0) and gap_2 = y_0-1.2max(y_0). The breakdown by quartiles refers to intial income, which is defined as deviation from the mean in percent.

Table 3: Conditional spatial sigma convergence among European regions, 1980-2002

	$tr\ \Sigma_{y0}/N$	$tr\ \Sigma_{yT}/N$	Percentage change	$tr\ \Sigma_{yT}/N = tr\ \Sigma_{y0}/N$ *vs.* $tr\ \Sigma_{yT}/N < tr\ \Sigma_{y0}/N$ Wald test	p-value
Starting values with explantory variables, c=50					
Spatial Solow model, global error spillovers	0.146	0.143	-1.95	10.9	0.00
Spatial Solow model, local error spillovers	0.089	0.087	-2.62	47.5	0.00
Spatial Verdoorn model, global error spillovers	0.746	0.736	-1.39	0.5	0.24
Spatial Verdoorn model, local error spillovers	0.096	0.095	-1.13	28.3	0.00
Starting values with explantory variables, c=25					
Spatial Solow model, global error spillovers	0.128	0.126	-2.05	18.8	0.00
Spatial Solow model, local error spillovers	0.101	0.099	-2.43	53.1	0.00
Spatial Verdoorn model, global error spillovers	0.746	0.737	-1.30	11.6	0.00
Spatial Verdoorn model, local error spillovers	0.096	0.095	-1.14	28.4	0.00
Starting values with explantory variables, c=75					
Spatial Solow model, global error spillovers	0.218	0.215	-1.69	3.3	0.03
Spatial Solow model, local error spillovers	0.080	0.077	-2.78	49.0	0.00
Spatial Verdoorn model, global error spillovers	0.746	0.736	-1.39	0.5	0.24
Spatial Verdoorn model, local error spillovers	0.096	0.095	-1.13	28.3	0.00
Fixed starting values, c=50					
Spatial Solow model, global error spillovers	0.169	0.165	-1.98	12.1	0.00
Spatial Solow model, local error spillovers	0.111	0.108	-2.62	18.5	0.00
Spatial Verdoorn model, global error spillovers	1.038	1.023	-1.40	0.4	0.26
Spatial Verdoorn model, local error spillovers	0.113	0.112	-1.13	28.7	0.00

Table A1: Descriptive statistics: Averages 1980-2002

	Mean	Std	Min	Max
Growth of real income per capita (log differences)	0.015	0.008	-0.013	0.052
Log initial real income per capita (log)	2.947	0.427	1.747	3.960
Log initial real income per capita- spatially weighted	2.960	0.356	1.920	3.527
Growth of real income (log differences)	0.020	0.009	-0.009	0.060
Growth of working population (log differences)	0.005	0.006	-0.005	0.072
Log initial working population	6.603	0.907	2.674	8.804
Initial investment rate	0.209	0.033	0.163	0.453
Initial share of agriculture (employment)	0.109	0.114	0.001	0.557
Initial share of markte services (employment)	0.336	0.073	0.107	0.543

Source: Cambridge Econometrics. Note 11 regions have been excluded because of missing data.

Table A2: β-convergence among European regions: 1980-2002, c=25

	Spatial Solow model, global error spillovers		Spatial Solow model, local error spillovers		Verdoorn's model, global error spillovers		Verdoorn's model, local error spillovers	
	b	z	b	z	b	z	b	z
Initial GDP per capita equation								
Initial working population	0.011	0.6	-0.010	-0.6	0.001	0.1	-0.014	-0.6
Initial investment rate	0.167	0.6	0.520	1.6 #	-0.026	-0.6	0.137	0.3
Initial share of agriculture	-0.570	-2.1 **	-1.324	-4.9 ***	-0.397	-1.5 #	-1.130	-1.5
Initial share of market services	1.099	4.1 ***	0.583	1.9 *	0.986	3.2 ***	0.655	1.7 *
Constant	2.483	13.2 ***	2.844	14.1 ***	1.290	1.8 *	2.885	9.8 ***
$\sigma_{\varepsilon 0}$	0.228		0.302		0.281		0.307	
ϕ_0	0.757	18.5 ***	-0.634	-11.8 ***	0.974	288.7 ***	-0.990	-1.2
Growth of real income per capita equation								
Initial real income per capita	-0.014	-9.5 ***	-0.014	-10.2 ***	-0.006	-5.5 ***	-0.006	-5.5 **
Initial real income per capita - spatially weighted	0.006	3.2 ***	0.006	3.5 ***	-	-		
Growth of working population	-0.285	-4.7 ***	-0.270	-4.5 ***	-	-		
Initial investment rate	0.054	3.8 ***	0.052	4.2 ***	-	-		
Growth inreal income per capita - spatially weighted	-	-	-	-	0.166	2.4 **	0.169	2.0 *
Growth of real income	-	-	-	-	0.528	13.3 ***	0.529	12.9 ***
Growth of real income - spatially weighted	-	-	-	-	-0.059	-1.1 *	-0.060	-0.9
Initial share of agriculture	-0.007	-1.2	-0.010	-1.83 *	-0.010	-2.4 **	-0.010	-1.7 *
Initial share of market services	0.039	5.3 ***	0.037	5.2 ***	0.003	0.7	0.004	0.6
Constant	0.018	2.6 ***	0.019	3.5 **	0.019	5.2 ***	0.019	5.2 ***
$\sigma_{\varepsilon T}$	0.006		0.006		0.005		0.005	
ϕ_T	0.432	6.2 ***	-0.390	-5.21 ***	0.562	678.0 ***	-0.990	-1.0
Likelihood	-967.7		-999.0		-963.1		-974.2	

Note: *** significant at 1%; ** significant at 5%; * significant at 10%; # significant at 15%.

Table A3: β-convergence among European regions: 1980-2002, c=75

	Spatial Solow model, global error spillovers		Spatial Solow model, local error spillovers		Verdoorn's model, global error spillovers		Verdoorn's model, local error spillovers	
	b	z	b	z	b	z	b	z
Initial GDP per capita equation								
Initial working population	0.019	1.2	-0.002	-0.1	0.001	0.1	-0.014	-0.4
Initial investment rate	-0.223	-0.9	0.075	0.4	-0.026	-1.2	0.137	0.3
Initial share of agriculture	-0.305	-1.2 **	-1.166	-1.4	-0.397	-1.9 *	-1.130	-1.5 #
Initial share of market services	1.176	4.7 ***	0.725	3.5 ***	0.986	3.0 ***	0.655	1.8 *
Constant	2.305	7.0 ***	2.820	11.0 ***	1.290	1.1	2.885	9.0 ***
$\sigma_{\varepsilon 0}$	0.215		0.273		0.281		0.307	
ϕ_0	0.942	30.3 ***	-0.990	-1.0 ***	0.974	49.1 ***	-0.990	-1.2
Growth of real income per capita equation								
Initial real income per capita	-0.015	-9.6 ***	-0.016	-9.9 ***	-0.005	-5.5 ***	-0.006	-5.3 **
Initial real income per capita - spatially weighted	0.008	3.0 ***	0.009	3.8 ***	-	-	-	-
Growth of working population	-0.269	-4.4 ***	-0.260	-3.3 ***	-	-	-	-
Initial investment rate	0.049	4.3 ***	0.047	3.07 ***	-	-	-	-
Growth inreal income per capita - spatially weighted	-	-	-	-	0.255	2.6 ***	0.250	2.0 **
Growth of real income	-	-	-	-	0.534	14.8 ***	0.533	13.5 **
Growth of real income - spatially weighted	-	-	-	-	-0.135	-1.5 #	-0.130	-1.0
Initial share of agriculture	0.000	0.0	-0.002	-0.38 *	-0.010	-2.5 **	-0.010	-2.1 **
Initial share of market services	0.040	5.3 ***	0.042	5.5 ***	0.003	0.6	0.004	0.6
Constant	0.012	1.5 ***	0.011	1.4 **	0.019	4.2 ***	0.019	4.4 ***
$\sigma_{\varepsilon T}$	0.006		0.006		0.005		0.005	
ϕ_T	0.608	5.3 ***	-0.969	-4.15 ***	0.508	2.2 *	-0.990	-2.0 **
Likelihood	-949.7		-980.4		-963.3		-974.5	

Note: *** significant at 1%; ** significant at 5%; * significant at 10%; # significant at 15%.

Table A4: β-convergence among European regions: 1980-2002, c=50, fixed starting values

	Spatial Solow model, global error spillovers			Spatial Solow model, local error spillovers			Verdoorn's model, global error spillovers			Verdoorn's model, local error spillovers		
	b	z		b	z		b	z		b	z	
Initial GDP per capita equation												
Constant	2.856	22.8	***	2.936	69.6	***	1.154	0.7		2.917	60.9	***
$\sigma_{\varepsilon 0}$	0.236			0.315			0.294			0.333		
ϕ_0	0.878	24.3	***	-0.990	-1.1	***	0.977	52.3	***	-0.990	-1.1	
Growth of real income per capita equation												
Initial real income per capita	-0.015	-9.3	***	-0.015	-9.9	***	-0.005	-4.2	***	-0.006	-5.4	***
Initial real income per capita - spatially weighted	0.007	3.1	***	0.008	3.7	***	-	-		-	-	
Growth of working population	-0.277	-4.7	***	-0.264	-4.0	***	-	-		-	-	
Initial investment rate	0.050	3.6	***	0.049	4.1	***	-	-		-	-	
Growth inreal income per capita - spatially weighted	-	-		-	-		0.255	1.9	*	0.250	1.8	*
Growth of real income	-	-		-	-		0.534	13.2	***	0.533	12.7	***
Growth of real income - spatially weighted	-	-		-	-		-0.135	-1.1		-0.130	-1.0	
Initial share of agriculture	-0.002	-0.4		-0.006	-1.1		-0.010	-2.1	**	-0.010	-2.1	**
Initial share of market services	0.041	5.5	***	0.040	6.0	***	0.003	0.5		0.004	0.6	
Constant	0.015	1.9	*	0.015	2.3	**	0.019	4.1	***	0.019	4.5	***
$\sigma_{\varepsilon T}$	0.006			0.006			0.005			0.005		
ϕ_T	0.566	5.9	***	-0.659	-5.2	***	0.508	1.6	#	-0.990	-1.3	
Likelihood	-970.3			-1006.2			-973.0			-991.3		

Note: *** significant at 1%; ** significant at 5%; * significant at 10%; # significant at 15%.

Table A5: Monte Carlo simulations of the Wald test for conditional spatial σ-convergence

		N=250				N=500			
		$b^{a)}$				$b^{a)}$			
ϕ_0	ϕ_T	-0.03	-0.02	-0.01	0.00	-0.03	-0.02	-0.01	0.00
Spatial Solow model with global error spillovers									
0.0	0.0	0.60	0.26	**0.05**	0.00	0.86	0.39	**0.04**	0.00
0.0	0.2	0.60	0.25	**0.06**	0.00	0.88	0.41	**0.04**	0.00
0.0	0.4	0.58	0.23	**0.05**	0.01	0.85	0.38	**0.05**	0.00
0.2	0.0	0.58	0.26	**0.04**	0.00	0.87	0.38	**0.05**	0.00
0.2	0.2	0.63	0.26	**0.05**	0.00	0.86	0.40	**0.05**	0.00
0.2	0.4	0.59	0.24	**0.05**	0.01	0.85	0.37	**0.05**	0.00
0.4	0.0	0.63	0.26	**0.05**	0.00	0.88	0.39	**0.05**	0.00
0.4	0.2	0.61	0.26	**0.06**	0.00	0.85	0.39	**0.05**	0.00
0.4	0.4	0.59	0.23	**0.05**	0.01	0.88	0.42	**0.05**	0.00
Spatial Solow model with locall error spillovers									
0.0	0.0	0.62	0.26	**0.05**	0.00	0.86	0.41	**0.04**	0.00
0.0	0.2	0.59	0.25	**0.06**	0.01	0.86	0.40	**0.06**	0.00
0.0	0.4	0.57	0.23	**0.05**	0.01	0.85	0.39	**0.05**	0.00
0.2	0.0	0.62	0.29	**0.06**	0.01	0.87	0.38	**0.05**	0.00
0.2	0.2	0.63	0.24	**0.05**	0.01	0.85	0.40	**0.05**	0.00
0.2	0.4	0.58	0.24	**0.05**	0.00	0.84	0.39	**0.05**	0.00
0.4	0.0	0.64	0.26	**0.04**	0.01	0.87	0.42	**0.05**	0.00
0.4	0.2	0.59	0.25	**0.05**	0.00	0.87	0.38	**0.05**	0.00
0.4	0.4	0.60	0.22	**0.06**	0.01	0.85	0.39	**0.05**	0.00
Verdoorn's model with global error spillovers									
0.0	0.0	0.62	0.28	**0.06**	0.01	0.86	0.43	**0.06**	0.01
0.0	0.2	0.61	0.26	**0.06**	0.01	0.84	0.38	**0.05**	0.00
0.0	0.4	0.60	0.25	**0.06**	0.00	0.88	0.38	**0.06**	0.00
0.2	0.0	0.60	0.29	**0.07**	0.02	0.87	0.44	**0.06**	0.00
0.2	0.2	0.59	0.27	**0.06**	0.01	0.85	0.40	**0.05**	0.00
0.2	0.4	0.59	0.23	**0.05**	0.01	0.83	0.39	**0.04**	0.00
0.4	0.0	0.67	0.30	**0.05**	0.01	0.91	0.46	**0.06**	0.00
0.4	0.2	0.64	0.28	**0.05**	0.01	0.87	0.43	**0.05**	0.00
0.4	0.4	0.58	0.24	**0.04**	0.00	0.86	0.37	**0.05**	0.00
Verdoorn's model with locall error spillovers									
0.0	0.0	0.64	0.27	**0.06**	0.01	0.86	0.38	**0.05**	0.00
0.0	0.2	0.70	0.31	**0.07**	0.01	0.90	0.48	**0.06**	0.00
0.0	0.4	0.83	0.48	**0.12**	0.02	0.95	0.64	**0.12**	0.00
0.2	0.0	0.58	0.25	**0.05**	0.01	0.86	0.37	**0.04**	0.00
0.2	0.2	0.63	0.29	**0.06**	0.01	0.88	0.42	**0.05**	0.00
0.2	0.4	0.76	0.39	**0.10**	0.01	0.94	0.56	**0.10**	0.00
0.4	0.0	0.57	0.21	**0.03**	0.00	0.81	0.31	**0.02**	0.00
0.4	0.2	0.60	0.24	**0.04**	0.01	0.84	0.32	**0.03**	0.00
0.4	0.4	0.68	0.30	**0.06**	0.01	0.88	0.42	**0.05**	0.00

a) Deviation form b at which H_0 is true.

University of Innsbruck – Working Papers in Economics and Statistics
Recent papers

2007-17 **Michael Pfaffermayr:** Conditional beta- and sigma-convergence in space: A maximum likelihood approach.

2007-16 **Anita Gantner:** Bargaining, search, and outside options. *Revised version forthcoming in: Games and Economic Behavior.*

2007-15 **Sergio Currarini and Francesco Feri:** Bilateral information sharing in oligopoly.

2007-14 **Francesco Feri:** Network formation with endogenous decay.

2007-13 **James B. Davies, Martin Kocher and Matthias Sutter:** Economics research in Canada: A long-run assessment of journal publications. *Revised version forthcoming in: Canadian Journal of Economics.*

2007-12 **Wolfgang Luhan, Martin Kocher and Matthias Sutter:** Group polarization in the team dictator game reconsidered. *Revised version forthcoming in: Experimental Economics.*

2007-11 **Onno Hoffmeister and Reimund Schwarze:** The winding road to industrial safety. Evidence on the effects of environmental liability on accident prevention in Germany.

2007-10 **Jesus Crespo Cuaresma and Tomas Slacik:** An “almost-too-late” warning mechanism for currency crises.

2007-09 **Jesus Crespo Cuaresma, Neil Foster and Johann Scharler:** Barriers to technology adoption, international R&D spillovers and growth.

2007-08 **Andreas Brezger and Stefan Lang:** Simultaneous probability statements for Bayesian P-splines.

2007-07 **Georg Meran and Reimund Schwarze:** Can minimum prices assure the quality of professional services?.

2007-06 **Michal Brzoza-Brzezina and Jesus Crespo Cuaresma:** Mr. Wicksell and the global economy: What drives real interest rates?.

2007-05 **Paul Raschky:** Estimating the effects of risk transfer mechanisms against floods in Europe and U.S.A.: A dynamic panel approach.

2007-04 **Paul Raschky and Hannelore Weck-Hannemann:** Charity hazard - A real hazard to natural disaster insurance.

2007-03 **Paul Raschky:** The overprotective parent - Bureaucratic agencies and natural hazard management.

2007-02 **Martin Kocher, Todd Cherry, Stephan Kroll, Robert J. Netzer and Matthias Sutter:** Conditional cooperation on three continents.

2007-01 **Martin Kocher, Matthias Sutter and Florian Wakolbinger:** The impact of naïve advice and observational learning in beauty-contest games.

University of Innsbruck

Working Papers in Economics and Statistics

2007-17

Michael Pfaffermayr

Conditional beta- and sigma-convergence in space: A maximum likelihood approach

Abstract

Empirical work on regional growth under spatial spillovers uses two workhorse models: the spatial Solow model and Verdoorn's model. This paper contrasts these two views on regional growth processes and demonstrates that in a spatial setting the speed of convergence is heterogenous in both considered models, depending on the remoteness and the income gap of all regions. Furthermore, the paper introduces Wald tests for conditional spatial sigma-convergence based on a spatial maximum likelihood approach. Empirical estimates for 212 European regions covering the period 1980-2002 reveal a slow speed of convergence of about 0.7 percent per year under both models. However, pronounced heterogeneity in the convergence speed is evident. The Wald tests indicate significant conditional spatial sigma-convergence of about 2 percent per year under the spatial Solow model. Verdoorn's specification points to a smaller and insignificant variance reduction during the considered period.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)