Working paper
2019-14
Statistics and Econometrics
ISSN 2387-0303

Growth with heterogenous interdependence

K. Miranda, M. Manjón, O. Martínez

uc3m | Universidad Carlos III de Madrid

Growth with heterogeneous technological interdependence*

K. Miranda[a,b], M. Manjón-Antolín [b], and O. Martínez-Ibañez [b]

[a]Department of Statistics, University Carlos III of Madrid

[b]QURE-CREIP Department of Economics, Rovira i Virgili University

Abstract

We present a growth model with spatial interdependencies in the heterogeneous technological progress and the stock of knowledge that, under certain conditions, yields a growth-initial equation that can be taken to the data. We then use data on EU-NUTS2 regions and a correlated random effects specification to estimate the resulting spatial Durbin dynamic panel model with spatially weighted individual effects. QML estimates support our model against simpler alternatives that impose a homogeneous technology. Also, our results indicate that rich regions tend to have higher (unobserved) productivity and are likely to stay rich because of the strong time and spatial dependence of the GDP per capita. Poor regions, on the other hand, tend to enjoy productivity spillovers but are likely to stay poor unless they increase their saving rates.

Keywords: correlated random effects, Durbin model, economic growth, spatial panel data

JEL Classification: C23, O47

*This research was funded by grants ECO2014-55553-P and ECO2016-78652 (Ministerio de Economía y Competitividad), ECO2018-88888-P (AEI/FEDER, UE) and 2014FI_B00301 (Agaur, Generalitat de Catalunya). We thank participants at the 59th ERSA Congress (Lyon), 21st INFER Annual Conference (Vrije Universiteit Brussel), UCLouvain Saint-Louis and the 6^{th} Workshop in Industrial and Public Economics (Universitat Rovira i Virgili) for useful comments. Usual caveats apply.

1 Introduction

Historically, the empirical economic growth literature consisted mostly of "aspatial empirical analyses that have ignored the influence of spatial location on the process of growth" (Fingleton and López-Bazo, 2006, p. 178). In the last two decades, however, a number of studies seek to incorporate "spatial effects" in the standard (i.e., non-spatial) economic growth models. In particular, the idea that the spatial location of an economy may drive its economic growth has been developed using models of absolute location, which account for the location of one economy in the geographical space, and models of relative location, which account for the location of one economy with respect to the others. Econometrically, these two types of model are closely related to the concepts of spatial heterogeneity and spatial dependence (Abreu et al., 2005).

Although spatial heterogeneity is usually associated with parameter heterogeneity (see e.g. Ertur and Koch, 2007; Basile, 2008), the most common approach in the literature is to allow for unobserved differences using panel data (Islam, 1995; Elhorst et al., 2010). Also, knowledge spillovers are the main mechanism used to incorporate interactions between economies into the Solow-Swan neoclassical growth model (López-Bazo et al., 2004; Egger and Pfaffermayr, 2006; Ertur and Koch, 2007; Pfaffermayr, 2009). It is interesting to note, however, that these two streams of the literature have developed quite separately. Notable exceptions include Elhorst et al. (2010), who consider the extension of the model proposed by Ertur and Koch (2007) to panel data; Ho et al. (2013), who consider an ad-hoc extension of the model proposed by Mankiw et al. (1992) that includes a spatial autoregressive term and a spatial time lag term; and Yu and Lee (2012), who use a simplified version of the technology assumed by Ertur and Koch (2007) to derive a growth model with spatial externalities based on the model of Mankiw et al. (1992). This paper aims to contribute to this limited literature by considering a growth model with spatial heterogeneity and spatial externalities that nests the models introduced by Islam (1995) and Ertur and Koch (2007).

To be precise, we present a growth model with interdependencies in the (heterogeneous) technological progress and the stock of knowledge. The basic framework is similar to that of Ertur and Koch (2007), which means assuming that the stock of knowledge depends on one's own stock of physical capital and the stock of knowledge of the other economies. However, we consider an alternative source of spatial externalities in the stock of knowledge, namely that arising from the technological progress of the other economies (Klenow and Rodriguez-Clare, 2005). To make this possible, we do not assume a common exogenous technological progress but introduce heterogeneity in the initial level of technology (Prescott, 1998), which is interpreted here as a proxy for total factor productivity (Islam, 1995) and, consistent with the empirical evidence, assumed to be spatially correlated (LeSage and Fischer, 2012; Miranda et al., 2017b).[1]

Having presented our model, we then derive the steady-state equation and a growth-initial equation that can be taken to the data. It is worth noting, however, that although we closely follow the approach proposed by Ertur and Koch (2007), we use weaker assumptions and provide conditions for such assumptions to be fulfilled. Also, the resulting econometric specification is a spatial Durbin model similar to theirs, but with a dynamic component (see also Elhorst et al., 2010; Yu and Lee, 2012; Ho et al., 2013) and spatially weighted individual-specific effects (arising from the heterogeneous technological progress). Thus, given the obvious interest in distinguishing individual effects from their spatial spillovers, we resort to a correlated random effects specification (Miranda et al., 2017a,b). In particular, we estimate our growth-initial equation by Quasi-Maximum Likelihood (see also Lee and Yu, 2016) using EU-NUTS2 regional data from Cambridge Econometrics. We use regional data because, as López-Bazo et al. (2004, p. 43) argue, once it is accepted that "[e]conomies interact with each other (...), linkages are [likely] to be stronger [between close-by regions] than across heterogeneous countries".

We find evidence of technological interdependence in the output per capita of the EU regions, that is, a positive and significant impact of the level of technology of neighbouring regions. However, there is also evidence of "unobserved" technological interdependence in the EU regions (i.e., local spatial contagion of "unobserved productivity"), which supports our assumed technology. In contrast, our estimates do not support the role that Ertur and Koch

[1] The model can be further extended by considering the stock of capital of the other economies as an additional source of spatial externalities in the stock of knowledge (López-Bazo et al., 2004; Egger and Pfaffermayr, 2006; Pfaffermayr, 2009). However, it can be shown that the resulting econometric specification becomes then overspecified and requires imposing constraints on the parameters and/or using an alternative specification (e.g. using lags and/or different weighting matrices) to identify some of the parameters. We leave this extension for future research.

(2007) assume capital plays in shaping technological progress.

The rest of the paper is organised as follows. In section 2 we present the model and derive the empirical specification. In section 3 we discuss the data and the estimation results. Section 4 concludes.

2 The Model

In this section we present our growth model and derive the growth-initial equation to be estimated. To a large extent, our approach follows the steps of Ertur and Koch (2007). Thus, we first discuss and motivate the assumed technology, and then plug it into the (Cobb-Douglas) production function. We subsequently obtain the output per worker equation at the steady state, and finally the growth-initial equation. We depart from Ertur and Koch (2007) in that we derive the growth initial equation using weaker assumptions and show that our assumptions are fulfilled under certain conditions on the variables (or the parameters).

2.1 Technological interdependencies in growth

Our starting point is the Solow growth model originally proposed by Mankiw et al. (1992) using cross-section data and subsequently extended by Islam (1995) to panel data (see also Barro and Sala-i-Martin, 2003). Let us then consider a Cobb-Douglas production function for region $i = 1, \ldots, N$ in time $t = 1, \ldots, T$:

$$Y_{it} = A_{it} K_{it}^{\alpha} L_{it}^{1-\alpha}, \tag{2.1}$$

where Y_{it} denotes output, K_{it} physical capital (α is thus the capital share or output elasticity parameter), L_{it} labour, and A_{it} technology. All the variables are in levels and there are constant returns to scale in production. Also, while output, capital and labour are typically assumed to be observable, technology is assumed to be (partially) unobservable. Mankiw et al. (1992), for example, assume that $\ln A = a + \varepsilon$, where a is a constant term and ε is the standard i.i.d error.

For the purposes of this paper, a major feature of this model is that technology is assumed to grow exogenously and at the same rate in all regions. In mathematical terms, $A_{it} = \Omega_0 \exp(\mu t)$, where Ω_0 is the exogenous technological progress and μ is its constant rate of growth. This rules out the existence of knowledge spillovers arising from technological interdependence between the regional economies. However, accounting for technological interdependence and knowledge spillovers is critical when analysing how "the relative location of an economy affects economic growth" (Anselin, 2003; Fingleton and López-Bazo, 2006; Elhorst et al., 2010, p. 338).

To address this issue, Ertur and Koch (2007) assume that the technological progress of an economy depends on the stock of physical capital per worker in that economy as well as the stock of knowledge of the other economies. More specifically, they assume that the technology of an economy depends on a geometrically weighted average of the technology of other economies, thus making knowledge spillovers spread over all the regions (and hence become "global").[2] However, it is still assumed that "some proportion of technological progress is exogenous and identical in all countries" [p. 1036].

Thus, if Ω_{it} denotes the exogenous technological progress and $k_{it} = \frac{K_{it}}{L_{it}}$ the level of physical capital per worker (of region i in period t), Ertur and Koch (2007, p. 1036) assume that the technology of region i in period t is given by

$$A_{it} = \Omega_{it} k_{it}^{\phi} \prod_{j \neq i}^{N} A_{jt}^{\gamma w_{ij}}, \tag{2.2}$$

where "[t]he parameter ϕ describes the strength of home externalities generated by physical capital accumulation" and "the degree of [regional] technological interdependence generated by the level of spatial externalities is described by γ". Notice that the spatial relation between region i and its neighbouring regions is represented by a set of spatial weights or "exogenous friction terms" w_{ij}, with $j = 1, \ldots, N$, that are assumed to satisfy the following properties: $w_{ij} = 0$ if $i = j$, $0 \leq w_{ij} \leq 1$, and $\sum_j w_{ij} = 1$. Lastly, Ertur and Koch (2007) assume that $\Omega_{it} = \Omega_t = \Omega_0 \exp(\mu t)$, where μ is the constant rate of growth of the exogenous technological

[2] López-Bazo et al. (2004) and Egger and Pfaffermayr (2006), for example, consider growth models in which the knowledge spillovers are local in nature, in the sense that they are limited to the neighbouring regions.

progress. Therefore, the technology eventually assumed is $A_{it} = \Omega_0 \exp(\mu t) k_{it}^{\phi} \prod_{j \neq i}^{N} A_{jt}^{\gamma w_{ij}}$.

In this paper, we extend Ertur and Koch (2007)'s model by introducing heterogeneity and spatial dependence in the exogenous technological progress (while still assuming that the technological progress of an economy depends on the stock of knowledge of the other economies). First, as Mankiw et al. (1992, p. 6) point out, the Ω_0 "term reflects not just technology but resource endowments, climate, institutions, and so on; it may therefore differ across countries". In this respect, Prescott (1998) stresses the importance of considering differences in the adoption of technologies across economies in order to account for the observed differences in income levels. Thus, we introduce region heterogeneity into the definition of the exogenous technological progress by assuming that $\Omega_{it} = \Omega_{i0} \exp(\mu t)$.[3] Second, as Islam (1995, p. 1149) points out, Ω_{i0} "is an important source of parametric difference in the aggregate production function across [regions]". Econometrically, it can be interpreted as an individual-specific effect (possibly correlated with some of the covariates in the initial-growth specification eventually derived). Economically, it is "a measure of efficiency with which the [regions] are transforming their capital and labor resources into output and hence is very close to the conventional concept of total factor productivity" [p. 1155-1156]. This opens the door to considering productivity spillovers as an additional source of spatial externalities, since empirical evidence indicates that estimates of Ω_{i0} are spatially correlated (LeSage and Fischer, 2012; Miranda et al., 2017b). This correlation may arise, for example, from R&D spillovers (Klenow and Rodriguez-Clare, 2005) and/or analogous technology policies, practices and institutions across nearby regions (Parente and Prescott, 1994).[4]

All in all, a production technology that accounts for these alternative sources of spatial

[3]Alternative ways of modelling the exogenous technological progress are $\Omega_{it} = \Omega_0 \exp(\mu_i t)$ and $\Omega_{it} = \Omega_{i0} \exp(\mu_i t)$. However, these proposals would considerably increase the number of parameters of the model (by more than N, since it can be shown that the balanced growth rate becomes heterogeneous too) and make identification difficult, if not impossible (Lee and Yu, 2016).

[4]There may be alternative explanations for the spatial correlation of productivity. In any case, the question of how to introduce these alternative explanations in the model and test for their validity is clearly beyond the scope of this paper.

dependence is the following:

$$A_{it} = \Omega_{it} \prod_{j \neq i}^{N} \Omega_{jt}^{\varphi w_{ij}} k_{it}^{\phi} \prod_{j \neq i}^{N} A_{jt}^{\gamma w_{ij}} \tag{2.3}$$

where $\Omega_{it} = \Omega_{i0} \exp(\mu t)$, Ω_{i0} are non-observable (which is why Ω_{it} does not have a coefficient in 2.3), and φ can be interpreted as the degree of technological interdependence generated from the (unobserved) productivity spillovers. In particular, if we leave aside the heterogeneous technological progress, $\varphi = \phi = \gamma = 0$ would lead us to the model without spillovers proposed by Islam (1995) and $\varphi = 0$ to the model proposed by Ertur and Koch (2007). Notice also that, in contrast to the local contagion models of López-Bazo et al. (2004) and Egger and Pfaffermayr (2006), both ours and Ertur and Koch (2007)'s are models of global contagion (Anselin, 2003). We differ, however, in that whereas in their case there are no (global) spatial externalities in the stock of knowledge unless $\gamma \neq 0$, there are here if $\varphi \neq 0$ (albeit of a local nature). This is because our model accounts for both global and local contagion.

2.2 The production function

In order to obtain the explicit form of the Cobb-Douglas production function in 2.1 given our assumed technology, let us consider 2.3 expressed in logs and matrix form:

$$A = \Omega + \varphi W \Omega + \phi k + \gamma W A = (I - \gamma W)^{-1} \Omega + \varphi (I - \gamma W)^{-1} W \Omega + \phi (I - \gamma W)^{-1} k \tag{2.4}$$

where the parameters φ and γ have been previously described (in particular, it is now assumed that $1/\gamma$ is not an eigenvalue of W, the $N \times N$ spatial weight matrix, when $\gamma \neq 0$), A is the $N \times 1$ vector of logarithms of the technology, k is the $N \times 1$ vector of logarithms of the capital per worker, I is the $N \times N$ identity matrix, and $\Omega = \Omega_0 + \iota_N \mu t$ is the $N \times 1$ vector of logarithms of the exogenous technological progress with $\Omega_0 = (\ln \Omega_{10}, \ldots, \ln \Omega_{N0})'$ and ι_N being a $N \times 1$ vector of ones. Notice that, in this vein, the technology depends on the exogenous technological progress and the capital per worker (but not on the technology of the other economies).

Let us now denote by $w_{ij}^{(r)}$ the row i and column j element of matrix W^r. Notice that, since

W is assumed to be row-normalized, its eigenvalues are equal or smaller than one in absolute value. Thus, if $|\gamma|< 1$, $(I-\gamma W)^{-1}=\sum_{r=0}^{\infty}\gamma^r W^r$. This means that we may rewrite 2.4 as

$$\begin{aligned}A_{it} &= \prod_{j=1}^{N}\Omega_{jt}^{\sum_{r=0}^{\infty}\gamma^r w_{ij}^{(r)}}\prod_{j=1}^{N}\Omega_{jt}^{\frac{\varphi}{\gamma}\sum_{r=1}^{\infty}\gamma^r w_{ij}^{(r)}}\prod_{j=1}^{N}k_{jt}^{\phi\sum_{r=0}^{\infty}\gamma^r w_{ij}^{(r)}}\\ &= \Omega_{it}^{1+\left(\frac{\gamma+\varphi}{\gamma}\right)\sum_{r=1}^{\infty}\gamma^r w_{ii}^{(r)}}\prod_{j\neq i}^{N}\Omega_{jt}^{\left(\frac{\gamma+\varphi}{\gamma}\right)\sum_{r=1}^{\infty}\gamma^r w_{ij}^{(r)}}k_{it}^{\phi+\phi\sum_{r=1}^{\infty}\gamma^r w_{ii}^{(r)}}\prod_{j\neq i}^{N}k_{jt}^{\phi\sum_{r=1}^{\infty}\gamma^r w_{ij}^{(r)}}\end{aligned}$$

Then, given that $y_{it}=A_{it}k_{it}^{\alpha}$,

$$y_{it}=\Omega_{it}^{1+\left(\frac{(\gamma+\varphi)(u_{ii}-\alpha-\phi)}{\phi\gamma}\right)}\prod_{j\neq i}^{N}\Omega_{jt}^{\frac{(\gamma+\varphi)u_{ij}}{\phi\gamma}}k_{it}^{u_{ii}}\prod_{j\neq i}^{N}k_{jt}^{u_{ij}} \tag{2.5}$$

where $u_{ii}=\alpha+\phi+\phi\sum_{r=1}^{\infty}\gamma^r w_{ii}^{(r)}$ and $u_{ij}=\phi\sum_{r=1}^{\infty}\gamma^r w_{ij}^{(r)}$, with $u_{ii}+\sum_{j\neq i}^{N}u_{ij}=\sum_{j=1}^{N}u_{ij}=\alpha+\phi+\frac{\phi\gamma}{1-\gamma}=\alpha+\frac{\phi}{1-\gamma}$.

Notice that "this model implies spatial heterogeneity in the parameters of the production function", a feature shared with Ertur and Koch's model (2007, p. 1037). We differ, however, in that it is no longer the case that "if there are no physical capital externalities, i.e., $\phi=0$, we have $u_{ii}=\alpha$ and $u_{ij}=0$, (and) then the production function is written in the usual form" (as in e.g. Mankiw et al. 1992 and Islam 1995). As previously pointed out, here we further require that $\varphi=\gamma=0$ for the production function to be written in the usual form.

2.3 The Steady State equation

To derive the equation describing the output per worker of region i at the steady state, we proceed in the following way. First we rewrite the production function in logs and matrix form, $y=A+\alpha k$, and substitute the technology by its expression in 2.4. We then pre-multiply both sides of the resulting equation by $I - \gamma W$ to obtain

$$y = \Omega + \varphi W \Omega + (\alpha + \phi)k + \gamma W y \tag{2.6}$$

Lastly, we replace in 2.6 the log of the capital per worker in region i by its log value at the steady state, $\ln k^*_{it}$. To this end, we start by noting that the evolution of capital is governed by the following dynamic equation:

$$\dot{k}_{it} = s_i y_{it} - (n_i + \delta)k_{it} \tag{2.7}$$

where the dot over a variable denotes its derivative with respect to time, s_i is the fraction of output saved, n_i is the growth rate of labour, and δ is the annual rate of depreciation of capital (common to all regions). Given that production shows decreasing returns to scale, equation 2.7 implies that the capital-output ratio is constant and converges to a balanced growth rate g defined by $\dfrac{\dot{k}_{it}}{k_{it}} = \ln \dot{k}_{it} = \ln \dot{y}_{it} = g = \dfrac{\mu(1+\varphi)}{(1-\gamma)(1-\alpha)-\phi}$ (see appendix A). It can also be shown that, given a balanced growth rate g and 2.7 (see e.g. Barro and Sala-i-Martin, 2003), $\dfrac{k^*_{it}}{y^*_{it}} = \dfrac{s_i}{n_i+\delta+g}$ and $\ln k^*_{it} = \ln y^*_{it} + \ln\left(\dfrac{s_i}{n_i+\delta+g}\right)$. Lastly, note that if we compute the marginal productivity of capital (using the expression defining y_{it} in 2.5) we obtain $\dfrac{\dot{k}_{it}}{k_{it}} = s_i \Omega_{it}^{1+\left(\frac{(\gamma+\varphi)(u_{ii}-\alpha-\phi)}{\phi\gamma}\right)} \prod_{j\neq i}^{N} \Omega_{jt}^{\frac{(\gamma+\varphi)u_{ij}}{\phi\gamma}} k_{it}^{u_{ii}-1} \prod_{j\neq i}^{N} k_{jt}^{u_{ij}} - (n_i + g)$. Therefore, provided that $\alpha + \dfrac{\phi}{1-\gamma} < 1$, there are diminishing returns to the capital, as in Ertur and Koch (2007)'s model.

What remains to be done is to introduce in 2.6 (rewritten for economy i rather than in matrix form) the expression obtained for the log of the capital per worker in region i at the steady state. In doing so, we obtain the equation describing the output per worker of region i

at the steady state:

$$\ln y_{it}^* = \frac{\ln \Omega_{it}}{1-\alpha-\phi} + \frac{\varphi}{1-\alpha-\phi}\sum_{j=1}^{N} w_{ij}\ln \Omega_{jt} + \frac{\alpha+\phi}{1-\alpha-\phi}\ln\left(\frac{s_i}{n_i+\delta+g}\right) - \frac{\alpha\gamma}{1-\alpha-\phi}\sum_{j=1}^{N} w_{ij}\left(\frac{s_j}{n_j+\delta+g}\right) + \frac{(1-\alpha)\gamma}{1-\alpha-\phi}\sum_{j=1}^{N} w_{ij}\ln y_{jt}^* \tag{2.8}$$

Notice that this steady state equation differs from that obtained by Ertur and Koch (2007) in that the term $\frac{\varphi}{(1-\alpha-\phi)}\sum_{j=1}^{N} w_{ij}\ln \Omega_{jt}$ is missing in theirs. This arises from our assumption of heterogeneous exogenous technological progress, since Ω_{it} is assumed to be Ω_t in Ertur and Koch (2007) and, consequently, no exogenous technological interdependences are considered.

2.4 The growth-initial equation

In the standard, non-spatial growth models (see e.g. Barro and Sala-i-Martin, 2003), the analog of equation 2.8 gives an expression for the output per worker in the steady state that does not depend on the output per worker in the steady state of the other economies (i.e., the term $\frac{(1-\alpha)\gamma}{1-\alpha-\phi}\sum_{j=1}^{N} w_{ij}\ln y_{jt}^*$ is missing). Thus, a log-linear approximation to the dynamics around the steady state using a Taylor expansion produces a differential equation (see e.g. Mankiw et al., 1992) whose solution leads to a growth-initial equation that takes the form of a linear (dynamic panel data) model (Islam, 1995). In our case, however, this approach produces a complex system of first-order differential linear equations, thus raising two important questions: the stability of the solution and the derivation of an estimable growth-initial equation from such an stable solution. To address the stability issue, we simply refer to results provided by Ertur and Koch (2007). Next we concentrate on the derivation of the growth-initial equation.

Let us then consider a log linearisation of the marginal productivity of capital, $\frac{\dot{k}_{it}}{k_{it}}$, around the steady state:

$$\frac{\dot{k}_{it}}{k_{it}} = g + (u_{ii} - 1)(n_i + \delta + g)\left(\ln k_{it} - \ln k^*_{it}\right) + \sum_{j \neq i}^{N} u_{ij}(n_i + \delta + g)\left(\ln k_{jt} - \ln k^*_{jt}\right) \tag{2.9}$$

Ertur and Koch (2007) provide sufficient conditions for the (local) D-stability of the system of first-order differential equations in 2.9. Also, given 2.9, it can be shown (see appendix *B*) that the "converge equation" is

$$\dot{y}(t) - \dot{y}(t)^* = BJB^{-1}\left[y(t) - y(t)^*\right] \tag{2.10}$$

with $J = Diag(n_i + \delta + g)\left[-(1-\alpha)I + \phi(I - \gamma W)^{-1}\right]$ and $B = \alpha I + \phi(I - \gamma W)^{-1}$.

Notice that, in standard non-spatial growth models, matrix BJB^{-1} is diagonal and the elements of the diagonal correspond to the speed of convergence of each economy (Mankiw et al., 1992; Barro and Sala-i-Martin, 2003).[5] This is not the case here, which means that using a solution to 2.10 (see appendix *B* for details) to derive the growth-initial equation yields an econometric model too complicated to be estimated. Egger and Pfaffermayr (2006), for example, obtain a particular solution to the system of linear differential equations that defines their convergence equation. However, because of the intricate expression they obtain, their empirical implementation uses a reduced form version of the growth-initial equation that corresponds to a spatial autoregressive model with spatially correlated errors. In our case, it can be shown that this approach would result in a non-linear spatial model specification with heterogeneous coefficients. We leave this avenue for future research.

Ertur and Koch (2007) address this problem by imposing conditions on the converge equation, namely restricting the spatial dependence of the gap between the observed and the steady state values of the capital $(k - k^*)$ and the output per worker $(y - y^*)$. In particular, it is assumed that the gap between the observed and the steady state values of the capital and output

[5]To be precise, since it is generally assumed that $n_i \equiv n$, the speed of convergence is given by a scalar (i.e., all the economies have the same speed of convergence) and the convergence equation is just a linear differential equation.

per worker of one economy are proportional to the gap between the observed and the steady state values of the capital and output per worker of the other economies, respectively. Under these assumptions, the convergence equation in 2.10 becomes analogous to that in standard, non-spatial growth models and the growth-initial equation takes the form of a spatial Durbin model. This is the approach we follow here.

Still, a major limitation of Ertur and Koch (2007)'s approach is that they do not provide conditions for the assumed relations in the output-per-worker and output gaps to hold. We provide such a conditions in appendix C. To be precise, we provide the proof using the expression obtained for the capital in 2.9 and then show that the results also hold for the output (rather than attacking the problem directly using the complicated expression of the output in 2.10). In particular, if we define $\Phi_{ij} = \dfrac{\ln k_{it_0} - \ln k^*_{it_0}}{\ln k_{jt_0} - \ln k^*_{jt_0}}$ for all $i, j = 1, ..., N$ and some t_0, we show that the relation

$$\ln k_{it} - \ln k^*_{it} = \Phi_{ij} \left(\ln k_{jt} - \ln k^*_{jt}\right) \tag{2.11}$$

holds for all $t \geq t_0$ as long as the growth rate of labour satisfies that $n_i = \kappa \left[\sum_{j=1}^{N} u_{ij}\Phi_{ji} - 1\right]^{-1} - (g + \delta)$ for $i = 1, \ldots, N-1$ and $\kappa = (n_N + g + \delta)\left[\sum_{j=1}^{N} u_{Nj}\Phi_{jN} - 1\right]$.[6]

We also show that, under these conditions, the speed of convergence, λ_i, is given by the following differential equation (see appendix C):

$$\frac{d \ln y_{it}}{dt} = g - \lambda_i \left(\ln y_{it} - \ln y^*_{it}\right), \tag{2.12}$$

where $\lambda_i = \sum_{j=1}^{N} G_{ij}\Theta^{-1}_{ij}$, G_{ij} is the row i and column j element of matrix BJB^{-1} and

[6]Our results can also be interpreted as either a condition on the observed and steady-state capitals in t_0 or as a set of constraints on the parameters of the model (α, ϕ, γ and κ). Notice, however, that these approaches do not yield an explicit solution for the variables and parameters involved, respectively, which means that the existence of such a solution is generally not guaranteed (obviously, in the case of the parameters, a necessary condition for the existence of a solution is that $N \leq 4$).

$\Theta_{is} = \sum_{j=1}^{N} u_{ij}\Phi_{ij}/\Phi_{is}\sum_{j=1}^{N} u_{sj}\Phi_{sj}$ are the proportionality parameters of the output-per-worker gap, i.e., $\ln \frac{y_{it}}{y_{it}^*} = \Theta_{is} \ln \frac{y_{st}}{y_{st}^*}$.

Lastly, under the simplifying assumption that the speed of convergence is homogeneous across regions, i.e., $\lambda_i = \lambda$ for $i = 1, \ldots, N$, the growth initial equation can be written as (see appendix C for details):

$$
\begin{aligned}
\ln y_{it_2} = {} & e^{-\lambda\tau} \ln y_{it_1} - \rho e^{-\lambda\tau} \sum_{j=1}^{N} w_{ij} \ln y_{jt_1} + \rho \sum_{j=1}^{N} w_{ij} \ln y_{jt_2} \\
& + \frac{\left(1 - e^{-\lambda\tau}\right)(\alpha + \phi)}{1 - \alpha - \phi} \ln s_i - \frac{\left(1 - e^{-\lambda\tau}\right)(\alpha + \phi)}{1 - \alpha - \phi} \ln(n_i + \delta + g) \\
& - \frac{\left(1 - e^{-\lambda\tau}\right)\alpha\gamma}{1 - \alpha - \phi} \sum_{j=1}^{N} w_{ij} \ln s_j + \frac{\left(1 - e^{-\lambda\tau}\right)\alpha\gamma}{1 - \alpha - \phi} \sum_{j=1}^{N} w_{ij} \ln(n_j + \delta + g) \\
& + \frac{\left(1 - e^{-\lambda\tau}\right)}{1 - \alpha - \phi} \ln \Omega_{i0} + \frac{\left(1 - e^{-\lambda\tau}\right)\varphi}{1 - \alpha - \phi} \sum_{j=1}^{N} w_{ij} \ln \Omega_{j0} \\
& + g(1 - \rho)\left(t_2 - t_1 e^{-\lambda\tau}\right)
\end{aligned}
\tag{2.13}
$$

where $\tau = t_2 - t_1$, $t_2 > t_1 \geq t_0$, and $\rho = \frac{(1-\alpha)\gamma}{1 - \alpha - \phi}$.

3 Empirical results

3.1 Model specification and identification

To derive our econometric specification, notice that equation 2.13 (plus an i.i.d. shock ε) corresponds to the spatial Durbin dynamic panel model with individual-specific effects and their spatial spillovers:

$$z_{it} = \overline{\gamma}_1 z_{i,t-1} + \overline{\gamma}_2 \sum_{j=1}^{N} w_{ij} z_{j,t-1} + \rho \sum_{j=1}^{N} w_{ij} z_{jt} + \beta_1 x_{1_{it}} + \beta_2 x_{2_{it}} + \theta_1 \sum_{j=1}^{N} w_{ij} x_{1_{jt}} + \theta_2 \sum_{j=1}^{N} w_{ij} x_{2_{jt}}$$
$$+ \mu_i + \sum_{j=1}^{N} w_{ij} \alpha_j + f_t + \varepsilon_{it} \tag{3.1}$$

where $z_{it} = \ln y_{it_2}$, $z_{i,t-1} = \ln y_{it_1}$, $x_{1_{it}} = \ln s_{it}$, $x_{2_{it}} = \ln(n_{it} + \delta + g)$, $\overline{\gamma}_1 = e^{-\lambda\tau}$, $\overline{\gamma}_2 = -\rho e^{-\lambda\tau}$, $\beta_1 = \frac{\left(1 - e^{-\lambda\tau}\right)(\alpha + \phi)}{1 - \alpha - \phi}$, $\beta_2 = -\frac{\left(1 - e^{-\lambda\tau}\right)(\alpha + \phi)}{1 - \alpha - \phi}$, $\theta_1 = \frac{\left(1 - e^{-\lambda\tau}\right)\alpha\gamma}{1 - \alpha - \phi}$, $\theta_2 = -\frac{\left(1 - e^{-\lambda\tau}\right)\alpha\gamma}{1 - \alpha - \phi}$, $\mu_i = \frac{\left(1 - e^{-\lambda\tau}\right)}{1 - \alpha - \phi} \ln \Omega_{i0}$, $\alpha_i = \frac{\left(1 - e^{-\lambda\tau}\right)\varphi}{1 - \alpha - \phi} \ln \Omega_{i0}$ and $f_t = g(1 - \rho)\left(t_2 - t_1 e^{-\lambda\tau}\right)$.

This means that equation 3.1 corresponds to the model specification discussed by Lee and Yu (2016), except that their model does not distinguish the spatial counterparts of the individual effects ($\sum_{j=1}^{N} w_{ij}\alpha_j$). In other words, their individual effects correspond to $\mu_i + \sum_{j=1}^{N} w_{ij}\alpha_j$ in 3.1. In fact, in our model, the individual effects and their spatial counterparts are proportional (by a rate φ). This is, therefore, a particular case of the more general specification proposed by Miranda et al. (2017a).

To distinguish the individual effects from their spatial spillovers, we assume a correlated random effects specification for the individual effects (μ_i) and their spatial spillovers (α_i). This means making use of the following correlation functions (Mundlak, 1978; Chamberlain, 1982):

$$\begin{aligned} \mu_i &= \pi_{\mu_1}\left(\frac{1}{T}\sum_{t=1}^{T} x_{1_{it}}\right) + \pi_{\mu_2}\left(\frac{1}{T}\sum_{t=1}^{T} x_{2_{it}}\right) + \upsilon_{\mu i} \\ \alpha_i &= \pi_{\alpha_1}\left(\frac{1}{T}\sum_{t=1}^{T} x_{1_{it}}\right) + \pi_{\alpha_2}\left(\frac{1}{T}\sum_{t=1}^{T} x_{2_{it}}\right) + \upsilon_{\alpha i}, \end{aligned} \tag{3.2}$$

where π_{μ_1}, π_{μ_2}, π_{α_1} and π_{α_2} are the parameters associated with the period-means of the regressors, and $\upsilon_{\mu i}$ and $\upsilon_{\alpha i}$ are random error terms with $E(\upsilon_{\mu i}) = 0 = E(\upsilon_{\alpha i})$, $Var(\upsilon_{\mu i}) = \sigma_\mu^2$, $Var(\upsilon_{\alpha i}) = \sigma_\alpha^2$ and $Cov(\upsilon_{\mu i}, \upsilon_{\alpha i}) = \sigma_{\mu\alpha}$. Notice that there are alternative specifications that are nested in this error term structure. Notably, the standard "random effects" without spatial contagion (which is derived from our model by imposing the constraints $\pi_{\mu_1} = \pi_{\mu_2} = \pi_{\alpha_1} = \pi_{\alpha_2} = 0$, $\sigma^2_\alpha = 0$ and $\sigma^2_\mu \neq 0$) and a "fixed effects" error term (obtained when imposing $\pi_{\mu_1} \neq 0$, $\pi_{\mu_2} \neq 0$, $\pi_{\alpha_1} = \pi_{\alpha_2} = \sigma^2_\alpha = 0$, and $\sigma^2_\mu \neq 0$) analogous to that discussed by Mundlak (1978) and Chamberlain (1982).

The last thing to notice about our econometric specification is that the implied parameters (α, ϕ, φ, γ, λ and $\ln \Omega_{i0}$) are not identified. In particular, we cannot obtain a single estimate of α (since this can be obtained from λ, ϕ and either β_1 or β_2, but also from λ, ϕ, γ and either θ_1 or θ_2), ϕ (since this can be obtained from λ, α and either β_1 or β_2, but also from λ, α, γ and either θ_1 or θ_2), φ (since this can be obtained from each (α_i, μ_i) pair, but also from either π_{μ_1} and π_{α_1} or π_{μ_2} and π_{α_2}), γ (since this requires ρ, $\overline{\gamma}_1$, either β_1 or β_2, and either θ_1 or θ_2, respectively), λ (since this can be obtained from $\overline{\gamma}_1$, but also from $\overline{\gamma}_2$ and ρ) and $\ln \Omega_{i0}$ (since this requires either μ_i, $\overline{\gamma}_1$ and either β_1 or β_2, or α_i, $\overline{\gamma}_1$, φ and either β_1 or β_2) because in principle these parameters are overidentified. However, it is easy to see that equations 3.1 and 3.2 contain three sets of constraints on the parameters: *i)* $\beta_1 = -\beta_2$ and $\theta_1 = -\theta_2$ (arising from the assumption that the production function is homogeneous of degree one, thus making the output per capita depend only on the stock of physical capital); *ii)* $\overline{\gamma}_2 = -\rho\overline{\gamma}_1$ (arising from the assumed spatial-time dynamics of the technology); and *iii)* $\alpha_i = \varphi\mu_i$ (i.e., $\pi_\alpha = \varphi\pi_\mu$, $\sigma^2_\alpha = \varphi^2\sigma^2_\mu$ and $\sigma_{\mu,\alpha} = \varphi\sigma^2_\mu$, which arise from the assumed spatial contagion in the heterogeneous exogenous technology and unobserved productivity). By imposing these six constraints on 3.1 and 3.2 (i.e., the "unconstrained model"), we obtain a constrained version of our model in which all the implied parameters are identified.[7]

To this end, we start by replacing 3.2 in 3.1, which in matrix form yields:

$$Z_t = \overline{\gamma}_1 Z_{t-1} + \overline{\gamma}_2 W Z_{t-1} + \rho W Z_t + X_t\beta + W X_t\theta + \overline{X}\Pi_\mu + W\overline{X}\Pi_\alpha + f_t + \eta_t \tag{3.3}$$

where $X_t = \left(x_{1_t} \,\vdots\, x_{2_t} \right)$, $\overline{X}$ denote period-means of X_t, $\beta = (\beta_1, \beta_2)'$, $\theta = (\theta_1, \theta_2)'$,

[7] While *i)* also arises in Ertur and Koch (2007)'s model, *ii)* and *iii)* are specific to our model specification. In this respect, notice that Elhorst et al. (2010, p. 343) also consider the constraint $\overline{\gamma}_2 = -\rho\overline{\gamma}_1$. However, while in our case this arises directly from the derivation of our model specification, they argue that this "constraint is unnecessarily restrictive because no theoretical or empirical reason exists to impose it".

$\Pi_\mu = (\pi_{\mu_1}, \pi_{\mu_2})'$, $\Pi_\alpha = (\pi_{\alpha_1}, \pi_{\alpha_2})'$, and the error term is $\eta_t = \upsilon_\mu + W\upsilon_\alpha + \varepsilon_t$, with variance-covariance matrix given by $J_T \otimes (\sigma^2_\mu I + \sigma_{\mu\alpha}(W + W') + \sigma^2_\alpha WW') + \sigma^2_\varepsilon I_{NT}$, J_T being a $T \times T$ matrix of ones and I_{NT} being the $NT \times NT$ identity matrix. This is the unconstrained version of our econometric model.

Let us now define $S_1 = I - \rho W$, $S_2 = I + \varphi W$ and $X^*_{it} = \ln\left(\frac{s_{it}}{n_{it} + \delta + g}\right) = \ln(S_{it})$. Then, the constrained model is given by

$$S_1 Z_t = \overline{\gamma}^c_1 S_1 Z_{t-1} + \beta^c X^* + \theta^c W X^* + S_2 \overline{X} \Pi^c_\mu + f_t + \eta^c_t \tag{3.4}$$

with $\overline{\gamma}_2 = -\rho\overline{\gamma}^c_1$, $\beta_1 = -\beta_2 = \beta^c$, $\theta_2 = -\theta_1 = \theta^c$ and $\Pi_\alpha = \varphi\Pi^c_\mu$ and $\eta^c_t = \varepsilon_t + S_2\upsilon_\mu$, with variance-covariance matrix given by $J_T \otimes (\sigma^2_\mu S_2 S'_2) + \sigma^2_\varepsilon I_{NT}$. Notice that, in contrast to 3.3, the estimation of the constrained version of our econometric model in 3.4 (see e.g. Lee and Yu, 2016; Miranda et al., 2017a) provides an estimate of: *i)* the capital share, α (from $\overline{\gamma}^c_1$, β^c and θ^c); *ii)* the externalities of the stock of physical capital, ϕ (from $\overline{\gamma}^c_1$, β^c and θ^c); *iii)* the degree of technological interdependence between the unobserved productivity, φ (directly from S_2); *iv)* the degree of technological interdependence between the economies, γ (from $\overline{\gamma}^c_1$, β^c and θ^c); *v)* the speed of convergence, λ (from $\overline{\gamma}^c_1$); and *vi)* the unobserved productivity, $\ln\Omega_{i0}$ (from μ_i, β^c and $\overline{\gamma}^c_1$). In particular, obtaining a statistically significant estimate of φ should be interpreted as supportive evidence for our model (against that of Ertur and Koch, 2007). Also, obtaining a statistically significant estimate of γ would lead us to reject the model of Islam (1995).

3.2 Estimates from EU-NUTS2 regions

We estimate the model given by 3.3 using the approach and model specifications of Lee and Yu (2016) and Miranda et al. (2017a). We use the former as a benchmark for our basic parameters ($\overline{\gamma}_1$, $\overline{\gamma}_2$, ρ, β_1, β_2, θ_1 and θ_2, which, since all the variables are in logs, can be interpreted as elasticities) and the latter to obtain the whole set of estimates (i.e., the basic ones plus those appearing in the correlation functions: π_{μ_1}, π_{μ_2}, π_{α_1} and π_{α_2}), test the validity of the constrained version of the model (using a Likelihood Ratio test), and estimate the implied parameters (using the constrained version of the model). We also follow this scheme in the discussion of the results. This means that we will start with an analysis of the estimates of the basic and correlation functions parameters in the unconstrained and constrained models, then we will move on to the estimates of the implied parameters, and we will conclude with a description of the geographical distribution of the estimated "unobserved productivity" of the EU regions ($\ln \hat{\Omega}_{i0}$) and their estimated spatial spillover ($\hat{\varphi} \sum_{j=1}^{N} w_{ij} \ln \hat{\Omega}_{j0}$).

First, however, a word about the data. We use EU NUTS2 regional data from Cambridge Econometrics to estimate our model. In particular, our initial sample is analogous to the one analysed by Elhorst et al. (2010), so we can use their results as a benchmark to which ours will be compared. Thus, we initially consider 189 regions across 14 EU countries (Austria, Belgium, Germany, Denmark, Greece, Finland, France, Ireland, Italy, the Netherlands, Portugal, Spain, Sweden and the United Kingdom) using time intervals of five years (see also Ho et al., 2013; Lee and Yu, 2016) between 1982 and 2002. This results in a balanced panel dataset with four time periods (1982-1987, 1987-1992, 1992-1997, 1997-2002).[8]

It is worth noting, however, that we have explored alternative samples to check the robustness of our results. First, we extended our initial sample to cover the years of the recent global crisis (the time intervals 2002-2007 and 2007-2012). Second, we considered different time intervals in a wider time period (1980 to 2015, with observations for 1980-1985, 1985-1990, and up to 2010-2015, which was the last available period at the moment of writing this paper). Third, we considered alternative groups of countries (e.g., including Norway, which is a non-EU country, and/or dropping Portugal, Ireland, Italy, Spain and/or Greece, which are countries that have faced (severe) problems with economic growth over the last decade). In all these cases, the estimates we obtained for the (un)constrained model remained largely unaltered. We illustrate this by reporting results from these alternative sampling schemes: the period 2002 to 2012, the period 1980 to 2015, the period 1982 to 2002 without including Portugal, Ireland, Italy, Spain and Greece (the so-called "PIIGS") and the period 1982 to 2002

[8] To be precise, the (small) differences between our sample and that of Elhorst et al. (2010) are the following. First, they have data on Luxembourg and the period 1977-1982. Second, in their sample "the islands (such as those associated with southern European countries) are assumed to be connected to the mainland, so that each region has at least one neighbour" (p. 353). Here we only consider continental regions, which means that our sample does not include the Spanish cities of Ceuta and Melilla, the French "Départements d'outre mer", and the Greek, Finnish, French, Italian and Spanish islands.

without including Greece (since in all these cases the results were not substantially different when Norway was included).

All these estimates were obtained using real GDP per capita as the dependent variable (i.e., y_{it} is real GDP at 2005 constant prices over total population, in thousands of people). As for the explanatory variables, s_{it} is the ratio between investment expenditures and gross value-added (at 2005 constant prices and as a percentage) and n_{it} is the growth rate of the population over time (computed as in Islam 1995). As is common in the literature (see e.g. Mankiw et al., 1992; Islam, 1995; Ertur and Koch, 2007), we assume that $\delta + g = 0.05$ to compute the depreciation rate. Note also that time dummies and a constant term were included in the set of explanatory variables to account for f_t. Lastly, W is a contiguity weight matrix.

[Insert Table 1 about here]

Table 1 provides descriptive statistics for the dependent and main explanatory variables (i.e., y_{it}, s_{it} and n_{it}). In particular, we report the statistics for the five samples considered and the periods effectively used in estimation in each case (notice that we lose one observation due to the inclusion of the lagged dependent variable in the model). The differences in the values of the statistics across the samples considered are of small magnitude, particularly between the original sample and the same sample without Greece. In fact, the observed differences arise in the GDP and the saving rate, whereas the distribution of the depreciation rate remains almost unaltered across samples. It is also interesting to note that the recent economic crisis seems to have increased the levels of GDP and savings, but mostly for those regions that were already at the top of the distribution (i.e., the centre of the distribution of these variables has shifted to the right and the upper tail has increased, thus making differences between the extremes larger). The effect is similar when the PIIGS are dropped from the original sample, except that now it it is the lower tail of the distribution the one that increases (i.e., we are dropping regions with levels of GDP and savings that are lower than those of the rest of the sample).

[Insert Table 2 about here]

We move now to the analysis of the estimates of the model and, as previously pointed out, start by considering the estimates of the unconstrained version of the model. These are reported in Table 2. In particular, the first reported estimates (in column two) were obtained using the approach and model specification of Lee and Yu (2016), whereas the rest (columns three to seven) were obtained using that of Miranda et al. (2017a). We report results for the initial sample (period 1982 to 2002) in columns two and three and, subsequently, for the other samples considered (periods 1982 to 2012, 1980 to 2015, 1982 to 2002 without the PIIGS, and 1982 to 2002 without Greece).

We find a remarkable regularity in both the values and the statistical significance of the coefficients across the samples and estimation approaches considered. Perhaps the only differences worth mentioning are: *i)* the slightly lower value of the coefficient associated with the time-lagged dependent variable ($\overline{\gamma}_1$) when the model is estimated using Lee and Yu (2016)'s approach; and *ii)* the lack of statistical significance of the coefficient associated with the saving rate (β_1) when the years of the recent crisis are considered (i.e., the samples covering the periods 2002 to 2012 and 1980 to 2015). This caveat aside, all sets of estimates provide essentially the same picture.

In particular, the basic parameters are all statistically significant (except for θ_2) and have the predicted signs (see Ertur and Koch, 2007).[9] Consistent with the constraint $\overline{\gamma}_2 = -\rho\overline{\gamma}_1$, the spatial and time lagged dependent variables have a high and positive coefficient, whereas the spatially weighted lagged dependent variable has a negative and smaller coefficient in absolute value (see also Ho et al., 2013; Lee and Yu, 2016). Thus, the level of GDP per capita of the European regions is largely determined by its past GDP per capita, and the current and past GDP per capita of their neighbours. Further, the saving rate of an economy contributes positively to its GDP per capita, but its depreciation rate and the saving rate of the neighbouring regions both contribute negatively. All in all, these results indicate that richest

[9]Our estimates of the basic parameters are largely consistent with those reported by Basile (2008) using an analogous sample of regions and the period 1988 to 2000. They also concur with those reported in panel data studies analysing countries rather than regions (see e.g. Ho et al., 2013; Lee and Yu, 2016). In contrast, we find some differences with those reported by Pfaffermayr (2009), who considers an analogous period of analysis but whose sample includes regions in Norway and Switzerland.

areas are likely to stay rich (more so they if are geographically close to rich areas, like e.g. in the so-called "blue banana") while poorer areas can only (partially) catch up if they increase their saving rates and/or are geographically close to rich areas.

As for the correlation functions parameters, there is evidence of *i)* correlation between the individual effects and the covariates (since both the –mean– saving and depreciation rates are statistically significant) and *ii)* spatial contagion in the individual effects (since the spatially weighted –mean– saving rate is generally statistically significant). In addition, two of the variance components, σ^2_μ and σ^2_ε, are statistically significant. This supports our correlated random effects model specification. In particular, results are consistent with the constraint $\alpha_i = \varphi\mu_i$, which implies a "fixed effects" error term model with proportional spatial contagion (Miranda et al., 2017a).

[Insert Table 3 about here]

Next we consider the results for the constrained version of the model, which are reported in Table 3. Before discussing the estimates, however, we should assess the validity of equation 3.4 in the different samples considered. To this end, we used a Likelihood Ratio test. We found that the "fully" constrained version of the model (i.e., the model resulting from imposing the constraints $\beta_1 = -\beta_2$, $\theta_1 = -\theta_2$, $\overline{\gamma}_2 = -\rho\overline{\gamma}_1$ and $\alpha_i = \varphi\mu_i$) was statistically supported only in the last two samples (i.e., the period 1982 to 2002 without the PIIGS and without Greece).[10] Estimates from this fully constrained version of the model are reported in Table 3*b*. Still, after testing the validity of each constraint individually, we found that a "partially" constrained version of the model in which only the constraint $\alpha_i = \varphi\mu_i$ was imposed was not rejected in the first three samples (periods 1982 to 2002, 1982 to 2012, and 1980 to 2015). Estimates from this partially constrained version of the model (including φ, which is identified) are reported in Table 3*a*.

At first sight, there is very little to comment on the results reported in Table 3*a* since,

[10] In particular, the Likelihood Ratio test statistics we obtained in the first three samples were 18.42 (period 1982 to 2002), 42.06 (period 1982 to 2012) and 27.26 (period 1980 to 2015), all statistically significant at standard levels. The Likelihood Ratio test statistics of the other samples (the period 1982 to 2002 without the PIIGS and without Greece) are reported in the last row of Table 3.

as expected, they are very similar to the ones obtained from the unconstrained model (see Table 2). Yet two things are worth mentioning. First, the correlation functions parameters and the variance components parameters are all statistically significant. This again supports our correlated random effects specification. Second, the coefficient reflecting the degree of technological interdependence generated from the productivity spillovers, φ, shows a negative and (at least in two of the samples considered) statistically significant value. Also, the estimates we obtain for φ are similar across the samples considered. Given the imposed constraint, $\alpha_i = \varphi\mu_i$, this indicates that there exists a negatively proportional relation between the individual effects of the EU regions and their spatial spillovers. We will return to this point when we analyse the geographical distribution of $\ln\hat{\Omega}_{i0}$ and $\hat{\varphi}\sum_{j=1}^{N} w_{ij}\ln\hat{\Omega}_{j0}$.

As for the estimates of the "fully" constrained version of the model, the first thing to notice is that they are similar in the two samples considered (except for the lack of statistical significance of θ^c in the sample without the PIIGS). In particular, the basic parameters are all statistically significant and have the predicted signs (see Ertur and Koch, 2007). Also, if we compare our results with those obtained by Elhorst et al. (2010), our estimates of the difference in the logs of the saving and depreciation rates, and of its spatial counterpart, are both larger (and statistically significant, whereas only the former is in their case). The estimated coefficient of the spatially lagged dependent variable, on the other hand, is analogous to the one reported by Elhorst et al. (2010). Lastly, the other parameters have estimated values in line with those obtained for the "partially" constrained version of the model.

[Insert Table 4 about here]

We then use these "fully constrained" estimates to obtain the implied parameters of the theoretical model. These are reported in Table 4. Our main findings are the following. First, the statistical significance of the degree of technological interdependence generated from the (unobserved) productivity spillovers, φ, supports our assumed technology (against the related alternatives of Islam 1995 and Ertur and Koch 2007). Second, the estimated speed of convergence, as measured by λ, is around 2% and statistically significant, which is a standard result in the literature (Barro and Sala-i-Martin, 2003; López-Bazo et al., 2004; Ertur and

Koch, 2007; Lee and Yu, 2016). Third, the statistical significance of the degree of technological interdependence, as measured by γ, supports Ertur and Koch (2007)'s model and contradicts the model of Islam (1995). Moreover, its value is similar to that found by Ertur and Koch (2007) and Elhorst et al. (2010), somewhere in between them. Fourth, the estimates of the capital share, as measured by α are in line with those obtained in the literature (Barro and Sala-i-Martin, 2003; Ertur and Koch, 2007; Elhorst et al., 2010). Fifth, the parameter capturing capital externalities at the global level (ϕ, through γ) is not statistically significant. In other words, there is no sign of the capital externalities in technology found by Ertur and Koch (2007).

All in all, our estimates support our model specification against that of Islam (1995) and Ertur and Koch (2007). They also point to the existence of spatial spillovers in the unobserved productivity and the level of technology. That is, we find evidence supporting the existence of both local and global spillovers in the stock of knowledge. In contrast, we find no evidence of spatial externalities in the stock of capital.

[Insert Figure 1 about here]

To conclude our empirical analysis, we report the geographical distribution of the estimated "unobserved productivity" and its spatial spillover (to reiterate, obtained from the constrained model in 3.4) in Figure 1. More precisely, Figure 1 presents a map of the European regions considered and the values of these statistics grouped by quantiles: Figure 1*a* reports $\ln \hat{\Omega}_{i0}$ (the "unobserved productivity") whereas Figure 1*b* reports $\hat{\varphi} \sum_{j=1}^{N} w_{ij} \ln \hat{\Omega}_{j0}$ (the spatial spillover of the "unobserved productivity", that is, the impact on the technology of unit i of all the units neighbouring i having their "unobserved productivity"). Notice that we have opted to use the estimates from the 1982-2002 sample without Greece to construct Figure 1 because this allows us to analyse a larger number of regions. It is important to stress, however, that results were not substantially different when we used the 1982-2002 sample without the PIIGS. Notice also that, given the negative and statistically significant value found for φ, there is a negatively proportional relation between the unobserved productivity of each EU region and the spatial contagion of this unobserved productivity on its neighbouring regions.[11]

With this in mind, we start by noting the considerable heterogeneity that Figure 1*a* displays, which contradicts the standard assumption of homogeneous exogenous technological progress. In particular, the results indicate that the regions with the lowest estimated "unobserved productivity" are mostly located in Scandinavia (Finland and Sweden), Scotland and North of England, Northern Ireland, Central-South of France, South-Est of Germany, Austria, Central and North-West of Spain, and North-West and South of Italy. Figure 1*a* also shows that the geographical distribution of the higher estimated "unobserved productivity" covers the so-called "blue banana" (from the South of the UK to the South-West of Germany, thus including the North of France, Belgium and the Netherlands), plus Denmark and the Mediterranean regions of the South-West of France and Central Italy.

What is also interesting to note is that about half of the regions in the high productivity group can be qualified as "rich", meaning here that their average GDP per capita over the periods considered is in the upper quantile of the distribution. Likewise, the same criterion would lead us to qualify about half of the regions with low estimated productivities as "poor". Thus, it seems that many of the richer/poorer regions tend to have higher/lower (unobserved) productivities. In fact, the Spearman rank correlation between $\ln \hat{\Omega}_{i0}$ and the average GDP per capita is 0.36 and statistically significant.

As for the spillovers associated with the "unobserved productivity", Figure 1*b* reveals that the pattern tends to be opposite to the one found for the estimated "unobserved productivity". In particular, the largest values are found in the Northern regions (i.e., Scandinavia, East of Ireland, the UK Midlands and South of Scotland), but also in the East (i.e., Austria) and South (South-West of France, North and West of Spain, and South of Italy) of Europe. This means that these are (often poor) regions whose "unobserved productivity" is more impacted by the "unobserved productivity" of its neighbours. The South of England and Ireland, Belgium, the Netherlands, and West Germany, on the other hand, stand as the areas with the lowest

[11]These spillovers correspond to the (local) spill-in effects proposed by Miranda et al. (2017b). We do not report the spill-out effects because, given the proportional relation that imposes the constraint $\alpha_i = \varphi\mu_i$, its geographical distribution is no more informative than that of $\ln \hat{\Omega}_{i0}$ (in fact, since both $\ln \hat{\Omega}_{i0}$ and φ take negative values, the spill-out effects take positive values and are larger/smaller the smaller/larger $\ln \hat{\Omega}_{i0}$ is).

spillovers. This means that these are (mostly rich) regions whose output per capita is barely affected by the "unobserved productivity" of its neighbours.

4 Conclusions

In this paper we present a growth model that extends previous knowledge-spillover models in several directions. First, we do not assume common exogenous technological progress but account for heterogeneity in the initial level of technology. Second, we derive the speed of convergence and the growth-initial equation under weaker assumptions than in previous literature and provide conditions for these assumptions to hold. We then use EU-NUTS2 regional information from Cambridge Econometrics to test whether the data supports the main features of our growth model. In particular, our econometric specification is derived from the growth-initial equation of the model and takes the form of a spatial Durbin dynamic panel model with spatially weighted individual effects.

We estimate the model by QML using a correlated random effects specification for the individual effects and their spatial spillovers. Our results support our model specification. Also, they are largely *i)* consistent with other studies using analogous data; and *ii)* robust to the use of alternative specifications, samples and estimation approaches. In particular, we find evidence of the existence of (global) spatial spillovers arising from the level of technology, but not from the investment in capital. Also, our estimates indicate that the level of GDP per capita of the European regions is largely determined by their past GDP per capita and the current and past GDP per capita of their neighbours, their saving rate and that of their neighbours, and their depreciation rate. However, the role of unobservable characteristics is worth noting: the richest areas (e.g., the "blue banana") are rich partially because of their higher "unobserved productivity" and a number of poor regions benefit from "unobserved productivity" spillovers.

Table 1: **Descriptive statistics**

(a) *Sample I: 1982-2002*

Variable	Mean	St. Dev.	Min	P25	Median	P75	Max
GDP	23,393	9,961	6,321	18,554	22,307	26,227	133,452
s	23.39	4.50	9.98	20.65	23.08	25.77	46.08
$n+\delta+g$	0.05	0.00	0.04	0.05	0.05	0.06	0.07

(b) *Sample II: 1982-2012*

Variable	Mean	St. Dev.	Min	P25	Median	P75	Max
GDP	25,355	11,536	6,321	19,698	23,997	28,934	176,529
s	23.69	4.76	9.98	20.64	23.47	26.17	48.84
$n+\delta+g$	0.05	0.01	0.04	0.05	0.05	0.06	0.08

(c) *Sample III: 1980-2015*

Variable	Mean	St. Dev.	Min	P25	Median	P75	Max
GDP	25,322	11,842	5,798	19,567	24,020	28,829	191,016
s	23.51	4.81	9.39	20.50	23.26	25.83	46.31
$n+\delta+g$	0.05	0.01	0.04	0.05	0.05	0.06	0.07

(d) *Sample IV: 1982-2002 w/o PIIGS (Portugal, Ireland, Italy, Spain and Greece)*

Variable	Mean	St. Dev.	Min	P25	Median	P75	Max
GDP	25,317	10,247	12,208	20,464	23,397	27,307	133,452
s	23.28	4.44	10.82	20.65	23.00	25.41	46.08
$n+\delta+g$	0.05	0.00	0.04	0.05	0.05	0.06	0.07

(e) *Sample V: 1982-2002 w/o EL (Greece)*

Variable	Mean	St. Dev.	Min	P25	Median	P75	Max
GDP	23,936	9,881	6,321	19,188	22,620	26,525	133,452
s	23.34	4.41	9.98	20.65	23.08	25.70	46.08
$n+\delta+g$	0.05	0.00	0.04	0.05	0.05	0.06	0.07

Note: Number of observations: $189\times 4 = 756$ (Sample I), $189\times 6 = 1,134$ (Sample II), $189 \times 7 = 1,323$ (Sample III), $139 \times 4 = 556$ (Sample IV), and $180 \times 4 = 720$ (Sample V). GDP is real GDP (at 2005 constant prices, in Euros) per capita (using total population, in thousands of people). s is the ratio between investment expenditures and gross value-added (as a percentage and at 2005 constant prices, in Euros). n is is the working-age population growth rate (computed as in Islam 1995) and $\delta + g = 0.05$ (as in e.g. Mankiw et al., 1992; Islam, 1995; Ertur and Koch, 2007).

Table 2: **QML estimates (unconstrained model)**

	Sample I (1982-2002)	Sample I (1982-2002)	Sample II (1982-2012)	Sample III (1980-2015)	Sample I (w/o PIIGS)	Sample I (w/o EL)
$\overline{\gamma}_1$	0.6291*** (0.0304)	0.9049*** (0.0145)	0.9177*** (0.0160)	0.8520*** (0.0294)	0.8681*** (0.0221)	0.8980*** (0.0157)
$\overline{\gamma}_2$	−0.3202*** (0.0556)	−0.4317*** (0.0366)	−0.4746*** (0.0290)	−0.3934*** (0.0338)	−0.4757*** (0.0412)	−0.4706*** (0.0362)
ρ	0.5281*** (0.0432)	0.5047*** (0.0380)	0.5603*** (0.0277)	0.5463*** (0.0273)	0.5587*** (0.0513)	0.5357*** (0.0383)
β_1	0.1149*** (0.0283)	0.0774** (0.0354)	0.0124 (0.0187)	−0.0053 (0.0149)	0.0604 (0.0405)	0.1031*** (0.0349)
β_2	−0.1624*** (0.0434)	−0.1952*** (0.0542)	−0.1742*** (0.0370)	−0.1045*** (0.0320)	−0.1564*** (0.0506)	−0.1536*** (0.0529)
θ_1	−0.0944*** (0.0339)	−0.0907** (0.0419)	−0.0526** (0.0259)	0.0018 (0.0187)	−0.1090*** (0.0506)	−0.1154*** (0.0410)
θ_2	0.0553 (0.0577)	0.0528 (0.0714)	0.0337 (0.0482)	0.0317 (0.0404)	0.1085 (0.0703)	0.0446 (0.0697)
π_{μ_1}		−0.1131*** (0.0397)	−0.0526** (0.0259)	−0.0606** (0.0306)	−0.1185** (0.0482)	−0.1432*** (0.0403)
π_{μ_2}		0.3486*** (0.0728)	0.3321*** (0.0596)	0.3310*** (0.0752)	0.3358*** (0.0888)	0.3037*** (0.0737)
π_{α_1}		0.0954* (0.0502)	0.0829** (0.0337)	0.0613 (0.0393)	0.1223* (0.0637)	0.1189* (0.0508)
π_{α_2}		−0.1637 (0.1112)	−0.1244 (0.0846)	−0.1453 (0.1011)	−0.2721 (0.1360)	−0.0975 (0.1137)
σ^2_μ		0.0006*** (0.0002)	0.0004** (0.0001)	0.0010** (0.0003)	0.0013*** (0.0003)	0.0007*** (0.0002)
σ^2_α		1.7×10^{-5} (0.0004)	0.0002 (0.0003)	0.0008 (0.0005)	0.0001 (0.0008)	1.2×10^{-5} (0.0004)
$\sigma_{\mu\alpha}$		0.0001 (0.0002)	−0.0002 (0.0002)	−0.0007* (0.0004)	−0.0003 (0.0006)	−0.0001 (0.0003)
σ^2_ε		0.0035*** (0.0002)	0.0030*** (0.0002)	0.0028*** (0.0002)	0.0019*** (0.0002)	0.0028*** (0.0002)

Note: All estimates were obtained using the approach proposed by Miranda et al. (2017a), except for those in column two, which were obtained using the approach proposed by Lee and Yu (2016). Time dummies included but not reported. The symbol * indicates statistically significant at the 10% level, ** at the 5% level and *** at the 1% level.

Table 3: **QML estimates (constrained model)**

(a) Partially constrained model

	Sample I (1982-2002)	Sample II (1982-2012)	Sample III (1980-2015)
$\overline{\gamma}_1$	0.9028***	0.9217***	0.8674***
	(0.0144)	(0.0147)	(0.0271)
$\overline{\gamma}_2$	−0.4455***	−0.4853***	−0.4099***
	(0.0373)	(0.0281)	(0.0319)
ρ	0.5253***	0.5608***	0.5434***
	(0.0384)	(0.0271)	(0.0273)
β_1	0.0595	0.0019	−0.0086
	(0.0394)	(0.0196)	(0.0151)
β_2	−0.1891***	−0.1717***	−0.1044***
	(0.0562)	(0.0372)	(0.0321)
θ_1	−0.0453	−0.0116	0.0142
	(0.0377)	(0.0174)	(0.0153)
θ_2	0.0323	0.0283	0.0361
	(0.0790)	(0.0487)	(0.0391)
$\pi_{\mu_1^c}$	−0.0903***	−0.0336	−0.0467*
	(0.0446)	(0.0261)	(0.0270)
$\pi_{\mu_2^c}$	0.3423***	0.3309***	0.3320***
	(0.0781)	(0.0597)	(0.0707)
$\sigma_\mu^{2^c}$	0.0005**	0.0004***	0.0009***
	(0.0002)	(0.0001)	(0.0003)
$\sigma_\varepsilon^{2^c}$	0.0035***	0.0031***	0.0029***
	(0.0002)	(0.0002)	(0.0001)
φ	−0.3803	−0.3773*	−0.4644**
	(0.3303)	(0.2267)	(0.1903)
LR-test	4.82	5.35	3.02

(b) Fully constrained model

	Sample I (w/o PIIGS)	Sample I (w/o EL)
$\overline{\gamma}_1^c$	0.8700***	0.9026***
	(0.0195)	(0.0131)
ρ^c	0.5747***	0.5496***
	(0.0405)	(0.0361)
β^c	0.0797**	0.1083***
	(0.0323)	(0.0333)
θ^c	−0.0575	−0.0780**
	(0.0318)	(0.0383)
$\pi_{\mu_1^c}$	−0.1257***	−0.1445***
	(0.0412)	(0.0392)
$\pi_{\mu_2^c}$	0.2195***	0.2507***
	(0.07530)	(0.0617)
σ_μ^c	0.0013***	0.0006***
	(0.0003)	(0.0002)
$\sigma_\varepsilon^{2^c}$	0.0020***	0.0031***
	(0.0002)	(0.0002)
φ	−0.3854*	−0.4432*
	(0.2108)	(0.2459)
LR-test	9.88	7.65

Note: The superscript c denotes constrained parameters (see section 3.2 for details). All estimates were obtained using the approach proposed by Miranda et al. (2017a). Time dummies are included but not reported. The symbol * indicates statistically significant at the 10% level, ** at the 5% level and *** at the 1% level. LR-test is the Likelihood Ratio test statistic of the hypothesis that the constraint $\alpha_i = \varphi\mu_i$ is valid (Table 3a) and the constraints $\beta_1 = -\beta_2$, $\theta_1 = -\theta_2$, $\overline{\gamma}_2 = -\rho\overline{\gamma}_1$ and $\alpha_i = \varphi\mu_i$ are valid (Table 3b).

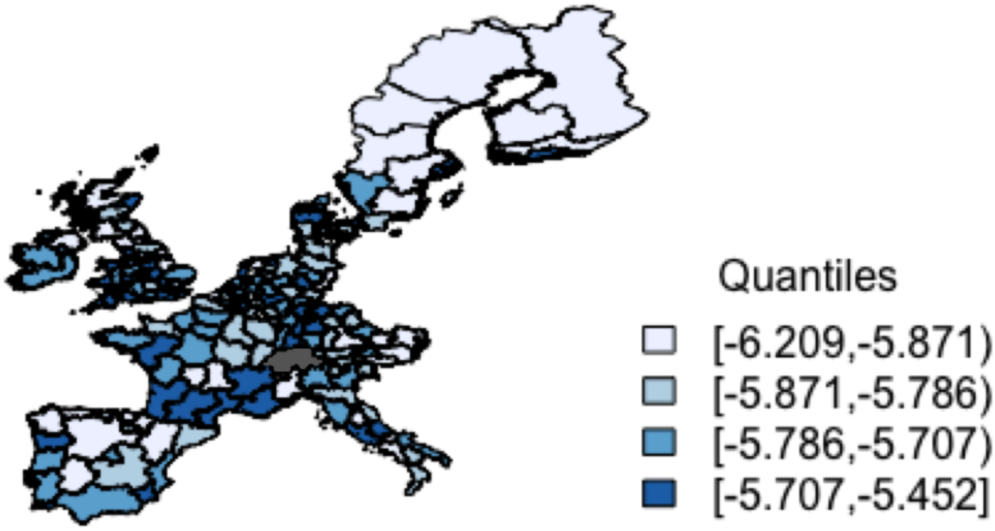
Table 4: **Implied Parameters**

	Sample I (w/o PIIGS)	Sample I (w/o EL)
α	0.5930***	0.4349***
	(0.1216)	(0.1374)
ϕ	-0.0665	-0.0548
	(0.0744)	(0.0983)
φ	-0.3854*	-0.4432*
	(0.2108)	(0.2459)
λ	0.0205***	0.0279***
	(0.0029)	(0.0045)
γ	0.6394***	0.6305***
	(0.1346)	(0.1162)

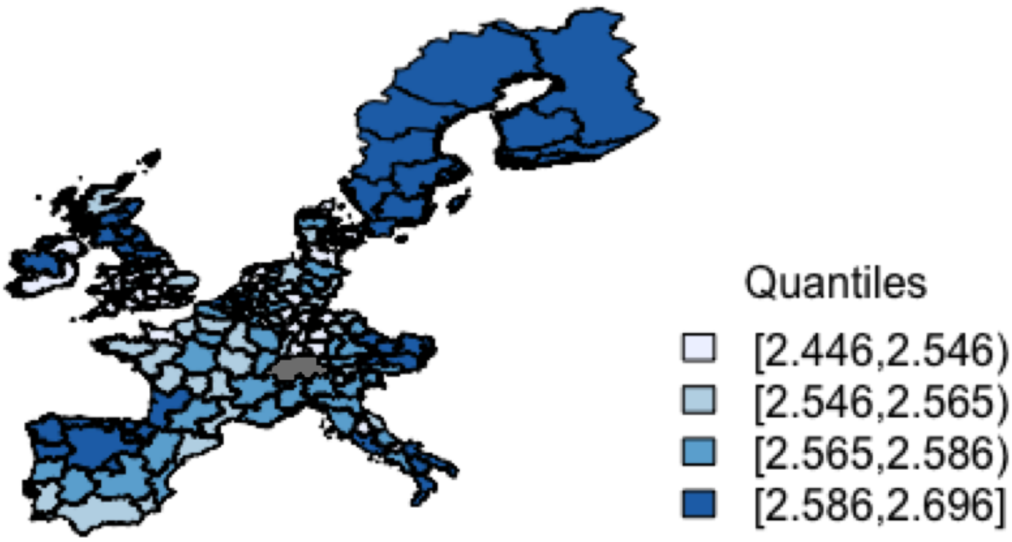
Note: * indicates statistically significant at the 10% level, ** at the 5% level and *** at the 1% level.

Figure 1: **Estimated individual effects and their spatial spillovers**

(a) Geographical distribution of $ln\hat{\Omega}_{i0}$



(b) Geographical distribution of $\hat{\varphi}\sum_{j=1}^{N} w_{ij} \ln \hat{\Omega}_{j0}$



A The balanced growth rate

From equation 2.5:

$$\ln y_{it} = \left[1 + \left(\frac{(\gamma+\varphi)(u_{ii}-\alpha-\phi)}{\phi\gamma}\right)\right]\ln\Omega_{it} + \left(\frac{\gamma+\varphi}{\phi\gamma}\right)\sum_{j\neq i}^{N} u_{ij}\ln\Omega_{jt} + u_{ii}\ln k_{it} + \sum_{j\neq i}^{N} u_{ij}\ln k_{jt}$$

Since $\ln\Omega_{it} = \ln\Omega_{i0} + \mu t$, then:

$$\frac{d\ln y_{it}}{dt} = \left[1 + \left(\frac{(\gamma+\varphi)(u_{ii}-\alpha-\phi)}{\phi\gamma}\right)\right]\mu + \left(\frac{\gamma+\varphi}{\phi\gamma}\right)\sum_{j\neq i}^{N} u_{ij}\mu + u_{ii}g + \sum_{j\neq i}^{N} u_{ij}g$$

Also, using $u_{ii} + \sum_{j\neq i}^{N} u_{ij} = \sum_{j=1}^{N} u_{ij} = \alpha + \frac{\phi}{1-\gamma}$,

$$\frac{d\ln y_{it}}{dt} = \left(1 - \frac{(\gamma+\varphi)(\alpha+\phi)}{\phi\gamma} + \left(\frac{\gamma+\varphi}{\phi\gamma}\right)\left(\frac{\alpha(1-\gamma)+\phi}{1-\gamma}\right)\right)\mu + \left(\alpha + \frac{\phi}{1-\gamma}\right)g = g,$$

which after some algebra becomes:

$$\left(\frac{1+\varphi}{1-\gamma}\right)\mu + \left(\frac{\alpha(1-\gamma)+\phi}{1-\gamma}\right)g = g$$

Therefore,

$$g = \frac{\mu\,(1+\varphi)}{(1-\gamma)(1-\alpha)-\phi}$$

B The convergence equation

Folowing Egger and Pfaffermayr (2006), let us start by calculating $\dfrac{d\ \ln k_{it}}{dt} - \dfrac{d\ \ln k^*_{it}}{dt}$:

$$
\begin{aligned}
\dot{k}_{it} - \dot{k}^*_{it} &= -(n_i + \delta + g)(\ln k_{it} - \ln k^*_{it}) + u_{ii}(n_i + \delta + g)(\ln k_{it} - \ln k^*_{it}) \\
&\quad + \sum_{j \neq i}^{N} u_{ij}(n_i + g + \delta)(\ln k_{jt} - \ln k^*_{jt}) \\
&= -\left(1 - \alpha\right)\left(n_i + \delta + g\right)(\ln k_{it} - \ln k^*_{it}) + \phi \sum_{j=1}^{N} \sum_{r=0}^{\infty} \gamma^r w^{(r)}_{ij} (n_i + \delta + g)(\ln k_{jt} - \ln k^*_{jt})
\end{aligned}
$$

where the second term is easily obtained using the definitions of u_{ii} and u_{ij}, and the fact that $\sum_{r=1}^{\infty} \gamma^r w^{(r)}_{ij} = \sum_{r=0}^{\infty} \gamma^r w^{(r)}_{ij} - w^{(0)}_{ij}$. In matrix notation:

$$
\begin{aligned}
\dot{k}(t) - \dot{k}(t)^* &= Diag(n_i + \delta + g)\left[-\left(1 - \alpha\right) I + \phi \left(I - \gamma W\right)^{-1}\right] \left(k(t) - k(t)^*\right) \\
&= Diag(n_i + \delta + g) P Diag\left(-\left(1 - \alpha\right) + \frac{\phi}{1 - \gamma \tau_i}\right) P^{-1} \left(k(t) - k(t)^*\right) \\
&= J\left(k(t) - k(t)^*\right) \qquad \text{(B.1)}
\end{aligned}
$$

where we use the term (t) after a matrix to stress its time dependence, $Diag()$ denotes a diagonal matrix whose elements correspond to the expression in brackets, P is the matrix of eigenvectors of W, and τ_i is the i-th eigenvalue of W. Notice that, since W is row-normalized, $|\tau_i| \leq 1$.

Now, from equation 2.5,

$$
\ln y_{it} = \left[1 + \frac{\varphi(u_{ii} - \alpha - \phi)}{\phi\gamma}\right] \ln \Omega_{it} + \frac{(\gamma + \varphi)}{\phi\gamma} \sum_{j \neq i}^{N} u_{ij} \ln \Omega_{jt} + u_{ii} \ln k_{it} + \sum_{j \neq i}^{N} u_{ij} \ln k_{jt}
$$

and its value at the steady state,

$$
\ln y^*_{it} = \left[1 + \frac{\varphi(u_{ii} - \alpha - \phi)}{\phi\gamma}\right] \ln \Omega_{it} + \frac{(\gamma + \varphi)}{\phi\gamma} \sum_{j \neq i}^{N} u_{ij} \ln \Omega_{jt} + u_{ii} \ln k^*_{it} + \sum_{j \neq i}^{N} u_{ij} \ln k^*_{jt},
$$

we obtain that

$$\ln y_{it} - \ln y_{it}^* = u_{ii}(\ln k_{it} - \ln k_{it}^*) + \sum_{j\neq i}^{N} u_{ij}(\ln k_{jt} - \ln k_{jt}^*) \tag{B.2}$$

Also, using the definitions of u_{ii} and u_{ij}, and the fact that $\sum_{r=1}^{\infty}\gamma^r w_{ij}^{(r)} = \sum_{r=0}^{\infty}\gamma^r w_{ij}^{(r)} - w_{ij}^{(0)}$, we may rewrite $B.2$ as

$$\ln y_{it} - \ln y_{it}^* = \alpha(\ln k_{it} - \ln k_{it}^*) + \phi\sum_{j=1}^{N}\sum_{r=0}^{\infty}\gamma^r w_{ij}^{(r)}(\ln k_{jt} - \ln k_{jt}^*)$$

In matrix notation:

$$y(t) - y^*(t) = \left[\alpha I + \phi\left(I - \gamma W\right)^{-1}\right]\left(k(t) - k(t)^*\right) = B\left(k(t) - k(t)^*\right), \tag{B.3}$$

Thus,

$$\dot{y}(t) - \dot{y}(t)^* = \left[\alpha I + \phi\left(I - \gamma W\right)^{-1}\right]\left(\dot{k}(t) - \dot{k}(t)^*\right) = B\left(\dot{k}(t) - \dot{k}(t)^*\right) \tag{B.4}$$

Lastly, plugging $B.1$ into $B.4$ and replacing $k(t) - k(t)^*$ by its expression from $B.3$, we obtain the "convergence equation" (Egger and Pfaffermayr, 2006; Ertur and Koch, 2007):

$$\dot{y}(t) - \dot{y}(t)^* = BJB^{-1}\left[y(t) - y(t)^*\right] = B\left[Diag(n_i + \delta + g)(-I + B)\right]B^{-1}\left(y(t) - y(t)^*\right) \tag{B.5}$$

We conclude by noting that if we denote by $\exp\{\}$ the exponential function, the solution to the first-order differential equation in $B.5$ expressed in terms of $y(t_1) - y^*(t_1)$ is:

$$y(t) - y^*(t) = \exp\left\{(t - t_1)BJB^{-1}\right\}\left[y(t_1) - y^*(t_1)\right]$$

C Speed of convergence and the growth-initial equation

Following Ertur and Koch (2007), let us consider the following condition

$$\ln k_{it} - \ln k_{it}^* = \Phi_{ij}\left(\ln k_{jt} - \ln k_{jt}^*\right) \tag{C.1}$$

for all $i, j = 1, ..., N$ and $t \geq t_0$. Let us also define $\Phi_{ij} = \dfrac{\ln k_{it_0} - \ln k_{it_0}^*}{\ln k_{jt_0} - \ln k_{jt_0}^*}$, a set of $N-1$ "proportionality parameters" satisfying $\Phi_{ij} = \Phi_{ji}^{-1}$, $\Phi_{sj} = \Phi_{ij}/\Phi_{is}$ and $\Phi_{ii} = 1$. Next we show that $C.1$ holds if

$$n_i = \kappa\left[\sum_{j=1}^{N} u_{ij}\Phi_{ji} - 1\right]^{-1} - (g+\delta) \tag{C.2}$$

for $i = 1, \ldots, N-1$ and $\kappa = (n_N + g + \delta)\left[\sum_{j=1}^{N} u_{Nj}\Phi_{jN} - 1\right]$.

We start by rewriting equation 2.9 as

$$\frac{\partial \ln \frac{k_{it}}{k_{it}^*}}{\partial t} = (n_i + g + \delta)\left[\sum_{j=1}^{N} u_{ij}\Phi_{ji} - 1\right]\ln\frac{k_{it}}{k_{it}^*}$$

Then, for some constant κ_i,

$$\ln\frac{k_{it}}{k_{it}^*} = \kappa_i \exp\left((n_i + g + \delta)\left[\sum_{j=1}^{N} u_{ij}\Phi_{ji} - 1\right]t\right)$$

and so

$$\ln\frac{k_{it+1}}{k_{it+1}^*} = \exp\left((n_i + g + \delta)\left[\sum_{j=1}^{N} u_{ij}\Phi_{ji} - 1\right]\right)\ln\frac{k_{it}}{k_{it}^*}.$$

Now suppose that condition $C.1$ holds in some $t \geq t_0$. Then,

$$\begin{aligned}
\ln \frac{k_{it+1}}{k^*_{it+1}} &= \exp\left((n_i + g + \delta)\left[\sum_{j=1}^{N} u_{ij}\Phi_{ji} - 1\right]\right)\Phi_{is} \ln \frac{k_{st}}{k^*_{st}} \\
&= \frac{\exp\left((n_i + g + \delta)\left[\sum_{j=1}^{N} u_{ij}\Phi_{ji} - 1\right]\right)}{\exp\left((n_s + g + \delta)\left[\sum_{j=1}^{N} u_{sj}\Phi_{js} - 1\right]\right)}\Phi_{is} \ln \frac{k_{st+1}}{k^*_{st+1}}
\end{aligned}$$

In other words, given that $C.1$ holds in t_0 (by definition of Φ_{ij}), $C.1$ holds in all $t \geq t_0$ as long as, for all $i = 1, \ldots, N$ and $s = 1, \ldots, N$,

$$\frac{\exp\left((n_i + g + \delta)\left[\sum_{j=1}^{N} u_{ij}\Phi_{ji} - 1\right]\right)}{\exp\left((n_s + g + \delta)\left[\sum_{j=1}^{N} u_{sj}\Phi_{js} - 1\right]\right)} = 1,$$

which is obviously equivalent to condition $C.2$ given that $C.2$ implies that $(n_i + g + \delta)\left[\sum_{j=1}^{N} u_{ij}\Phi_{ji} - 1\right] = \kappa$ for all $i = 1, \ldots, N$.

Now it is easy to show that, if condition $C.2$ is satisfied for $\Phi_{ij} = \dfrac{\ln k_{it_0} - \ln k^*_{it_0}}{\ln k_{jt_0} - \ln k^*_{jt_0}}$, then the following relation assumed by Ertur and Koch (2007)

$$\ln y_{it} - \ln y^*_{it} = \Theta_{ij}\left(\ln y_{jt} - \ln y^*_{jt}\right) \tag{C.3}$$

also holds, with Θ_{ij} being another set of $N-1$ "proportionality parameters" ($\Theta_{ij} = \Theta^{-1}_{ji}$, $\Theta_{sj} = \Theta_{ij}/\Theta_{is}$ and $\Theta_{ii} = 1$). To this end, let us rewrite $B.2$ as

$$\ln y_{it} - \ln y^*_{it} = \sum_{j=1}^{N} u_{ij}(\ln k_{jt} - \ln k^*_{jt}),$$

which, using $\ln k_{it} - \ln k_{it}^* = \Phi_{ij}\left(\ln k_{jt} - \ln k_{jt}^*\right)$, becomes

$$\ln \frac{y_{it}}{y_{it}^*} = \sum_{j=1}^{N} u_{ij}\Phi_{ij} \ln \frac{k_{it}}{k_{it}^*}$$

Therefore,

$$\frac{\ln \frac{y_{it}}{y_{it}^*}}{\ln \frac{y_{st}}{y_{st}^*}} = \frac{\sum_{j=1}^{N} u_{ij}\Phi_{ij}}{\sum_{j=1}^{N} u_{sj}\Phi_{sj}} \frac{\ln \frac{k_{it}}{k_{it}^*}}{\ln \frac{k_{st}}{k_{st}^*}} = \frac{\sum_{j=1}^{N} u_{ij}\Phi_{ij}}{\sum_{j=1}^{N} u_{sj}\Phi_{sj}} \frac{\ln \frac{k_{it}}{k_{it}^*}}{\Phi_{si}\ln \frac{k_{it}}{k_{it}^*}} = \frac{\sum_{j=1}^{N} u_{ij}\Phi_{ij}}{\Phi_{si}\sum_{j=1}^{N} u_{sj}\Phi_{sj}}$$

and

$$\ln \frac{y_{it}}{y_{it}^*} = \Theta_{is} \ln \frac{y_{st}}{y_{st}^*}$$

with $\Theta_{is} = \dfrac{\sum_{j=1}^{N} u_{ij}\Phi_{ij}}{\Phi_{si}\sum_{j=1}^{N} u_{sj}\Phi_{sj}}$.

Lastly, we use C.3 to derive the growth initial equation. To this end, we start by noting that the steady state in 2.8 can be written as

$$\ln y_{it}^* = \frac{1}{1-\alpha-\phi}\sum_{j=1}^{N}\sum_{r=0}^{\infty}\rho^r w_{ij}^{(r)} \ln \Omega_{jt} + \frac{\varphi}{1-\alpha-\phi}\sum_{j=1}^{N}\sum_{r=0}^{\infty}\rho^r w_{ij}^{(r+1)} \ln \Omega_{jt}$$
$$+ \left(\frac{\alpha+\phi}{1-\alpha-\phi}\right)\sum_{j=1}^{N}\sum_{r=0}^{\infty}\rho^r w_{ij}^{(r)} \ln\left(\frac{s_j}{n_j+\delta+g}\right) - \frac{\alpha\gamma}{1-\alpha-\phi}\sum_{j=1}^{N}\sum_{r=0}^{\infty}\rho^r w_{ij}^{(r+1)} \ln\left(\frac{s_j}{n_j+\delta+g}\right)$$

with $\rho = \dfrac{(1-\alpha)\gamma}{1-\alpha-\phi}$. Using this, we can see that $\dfrac{d\ln y_{it}^*}{dt} = \dfrac{(1+\varphi)\mu}{1-\alpha-\phi}\left(\dfrac{1}{1-\rho}\right)$. In fact, since $\dfrac{1}{1-\rho} = \dfrac{1-\alpha-\phi}{(1-\alpha)(1-\gamma)-\phi}$,

$$\frac{d\ln y_{it}^*}{dt} = g,$$

which can be seen as another differential equation whose particular solution expressed in terms of $\ln y_{i0}$ is

$$\ln y^*_{it} = gt + \ln y^*_{i0} \tag{C.4}$$

Notice also that plugging C.3 into B.5 we obtain that, for $t \geq t_0$,

$$\frac{d\ \ln y_{it}}{dt} = g - \lambda_i \left(\ln y_{it} - \ln y^*_{it}\right), \tag{C.5}$$

where $\lambda_i = \sum_{j=1}^{N} G_{ij}\Theta^{-1}_{ij}$ is the "speed of convergence" of each economy and G_{ij} is the row i and column j element of matrix BJB^{-1}.[12]

Then, plugging equation C.4 into C.5 we obtain:

$$\frac{d \ln y_{it}}{dt} = g - \lambda_i \left(\ln y_{it} - gt - \ln y^*_{i0}\right)$$

We use the integrating factor method to solve this differential equation. Thus, we first reorder terms and then multiply the equation by the integrating factor $e^{\int \lambda_i dt} = e^{\lambda_i t}$ to obtain

$$\frac{d}{dt}\left(e^{\lambda_i t} \ln y_{it}\right) = e^{\lambda_i t} g + \lambda_i e^{\lambda_i t}\left(gt + \ln y^*_{i0}\right),$$

which, by integrating on both sides, provides the general solution:

$$\ln y_{it} = gt + \ln y^*_{i0} + Ce^{-\lambda_i t}$$

for $t \geq t_0$. Then, specifying the constant C in terms of the function evaluated at $t = t_1$, $C = \left(\ln y_{it_1} - gt_1 - \ln y^*_{i0}\right) e^{\lambda_i t_1}$, the solution for any t is given by

$$\ln y_{it} = g\left(t - t_1 e^{-\lambda_i (t-t_1)}\right) + \ln y_{it_1} e^{-\lambda_i (t-t_1)} + \left(1 - e^{-\lambda_i (t-t_1)}\right) \ln y^*_{i0},$$

[12] Notice that Ertur and Koch (2007) derive an analogous expression for the speed of convergence but assuming that both C.1 and C.3 hold. Here, however, because of the result obtained in B.1 (following Egger and Pfaffermayr, 2006), we only require that one of the proportionality relations they consider hold (either the one on the capital, C.1, or that on the output-per-worker, C.3).

which, at $t = t_2 > t_1$, is

$$\ln y_{it_2} = g\left(t_2 - t_1 e^{-\lambda_i \tau}\right) - e^{-\lambda_i \tau} \ln y_{it_1} + (1 - e^{-\lambda_i \tau}) \ln y^*_{i0} \tag{C.6}$$

with $\tau = t_2 - t_1$.

At this point it is convenient to write the previous expression in matrix form under the simplifying assumption that the speed of convergence is homogeneous across regions:

$$y(t_2) = g\left(t_2 - t_1 e^{-\lambda \tau}\right) \iota_N + e^{-\lambda \tau} y(t_1) + \left(1 - e^{-\lambda \tau}\right) y^*(0) \tag{C.7}$$

where $y(t_2)$ is a $N \times 1$ vector containing the log of the outcome per worker at t_2, ι_N is a $N \times 1$ vector of ones, $y(t_1)$ is a $N \times 1$ vector containing the log of the outcome per worker at t_1, and $y^*(0)$ is a $N \times 1$ vector containing the log of the initial level of output per worker at the steady state. The reason for this is that it facilitates replacing $y^*(0)$ by 2.8, which, in matrix form, is:

$$y^*(0) = (I - \rho W)^{-1} \left[\frac{1}{1-\alpha-\phi}\Omega(0) + \frac{\varphi}{1-\alpha-\phi} W\Omega(0) + \frac{\alpha+\phi}{1-\alpha-\phi} S - \frac{\alpha\gamma}{1-\alpha-\phi} WS\right] \tag{C.8}$$

where it is assumed that $1/\rho$ is not an eigenvalue of W when $\rho \neq 0$ and $S = \left\{\ln\left(\frac{s_i}{n_i+\delta+g}\right)\right\}_{i=1,\ldots,N}$.

All is left is to introduce C.8 in C.7 and pre-multiply both sides of the resulting equation by $I - \rho W$ to obtain:

$$\begin{aligned} y(t_2) = {} & g(1-\rho)\left(t_2 - t_1 e^{-\lambda \tau}\right) \iota_N + e^{-\lambda \tau} (I - \rho W)\, y(t_1) + \rho W y(t_2) \\ & + \left(1 - e^{-\lambda \tau}\right) \left[\frac{1}{1-\alpha-\phi}\Omega(0) + \frac{\varphi}{1-\alpha-\phi} W\Omega(0) + \frac{\alpha+\phi}{1-\alpha-\phi} S - \frac{\alpha\gamma}{1-\alpha-\phi} WS\right] \end{aligned} \tag{C.9}$$

References

Abreu, M., de Groot, H., and Florax, R. (2005). Space and growth: A survey of empirical evidence and methods. *Région et Développement*, 21:13–44.

Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. *International Regional Science Review*, 26(2):153–166.

Barro, R. J. and Sala-i-Martin, X. (2003). *Economic Growth, 2nd Edition*. The MIT Press.

Basile, R. (2008). Regional economic growth in Europe: A semiparametric spatial dependence approach. *Papers in Regional Science*, 87(4):527–544.

Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics*, 18(1):5–46.

Egger, P. and Pfaffermayr, M. (2006). Spatial convergence. *Papers in Regional Science*, 85(2):199–215.

Elhorst, P., Piras, G., and Arbia, G. (2010). Growth and convergence in a multiregional model with space time dynamics. *Geographical Analysis*, 42(3):338–355.

Ertur, C. and Koch, W. (2007). Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics*, 22(6):1033–1062.

Fingleton, B. and López-Bazo, E. (2006). Empirical growth models with spatial effects. *Papers in Regional Science*, 85(2):177–198.

Ho, C.-Y., Wang, W., and Yu, J. (2013). Growth spillover through trade: A spatial dynamic panel data approach. *Economics Letters*, 120(3):450–453.

Islam, N. (1995). Growth empirics: A panel data approach. *The Quarterly Journal of Economics*, 110(4):1127–1170.

Klenow, P. J. and Rodriguez-Clare, A. (2005). Externalities and Growth. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1, pages 817–861. Elsevier.

Lee, L.-F. and Yu, J. (2016). Identification of spatial Durbin panel models. *Journal of Applied Econometrics*, 31(1):133–162.

LeSage, J. P. and Fischer, M. M. (2012). Estimates of the impact of static and dynamic knowledge spillovers on regional factor productivity. *International Regional Science Review*, 35(1):103–127.

López-Bazo, E., Vayá, E., and Artís, M. (2004). Regional externalities and growth: Evidence from European regions. *Journal of Regional Science*, 44(1):43–73.

Mankiw, N. G., Romer, D., and Weil, D. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, 107(2):407–437.

Miranda, K., Martínez-Ibañez, O., and Manjón-Antolín, M. (2017a). A correlated random effects spatial durbin model. Department of Economics, Rovira i Virgili Universiy.

Miranda, K., Martínez-Ibañez, O., and Manjón-Antolín, M. (2017b). Estimating individual effects and their spatial spillovers in linear panel data models: Public capital spillovers after all? *Spatial Statistics*, 22(1):1–17.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica*, 46(1):69–85.

Parente, S. and Prescott, E. (1994). Barriers to technology adoption and development. *Journal of Political Economy*, 102(2):298–321.

Pfaffermayr, M. (2009). Conditional [beta]- and [sigma]-convergence in space: A maximum likelihood approach. *Regional Science and Urban Economics*, 39(1):63–78.

Prescott, E. C. (1998). Needed: A theory of total factor productivity. *International Economic Review*, 39(3):525–551.

Yu, J. and Lee, L.-F. (2012). Convergence: A spatial dynamic panel data approach. *Global Journal of Economics*, 1(1):1–36.