Growth, heterogeneous technological interdependence, and spatial externalities: Theory and Evidence*

K. Miranda, M. Manjón-Antolín, and O. Martínez-Ibañez

QURE-CREIP Department of Economics, Rovira i Virgili University

Abstract

We present a growth model with interdependencies in the heterogeneous technological progress, physical capital and stock of knowledge that 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 specification against simpler alternatives that impose an homogeneous technology and limit the sources of spatial externalities. Also, our results indicate that the level of GDP per capita of the European regions is mainly determined by its unobserved heterogeneity and that of its neighbours, its past GDP per capita, and the current and past GDP per capita of their neighbours.

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

JEL Classification:

*This research was funded by grants ECO2014-55553-P, ECO2016-78652 and ECO2017-88888-P (Ministerio de Economía y Competitividad) and (Ministerio de Economía y Competitividad) and 2014FI_B00301 (Agaur, Generalitat de Catalunya). 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” (De Long and Summers, 1991; 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 models 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 employed to incorporate interactions between economies into the Solow-Swan neoclassical growth model (López-Bazo et al., 2004; Egger and Pfaffermayr, 2006; Pfaffermayr, 2009, 2012). It is interesting to note, however, that these two streams of the literature have developed rather 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, using a simplified version of the technology assumed by Ertur and Koch (2007), 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), López-Bazo et al. (2004) and Ertur and Koch (2007).

To be precise, we present a growth model with interdependencies in the (heterogeneous) technological progress, physical capital and stock of knowledge.[1] The basic framework is similar to that of Ertur and Koch (2007), but we consider additional sources of externalities across economies. While they assume that the technological progress depends on the stock of physical capital and the stock of knowledge of the other economies, we also consider the role of both the physical capital (López-Bazo et al., 2004; Egger and Pfaffermayr, 2006) and the (unobserved) initial level of technology (De Long and Summers, 1991; LeSage and Fischer, 2012) of the other economies. Moreover, we do not assume a common exogenous technological progress but account for heterogeneity in the initial level of technology, which here is interpreted as a proxy for total factor productivity (Islam, 1995).

Having presented our model, we then derive the steady-state equation and a growth-initial equation that can be taken to data. This is where the generality of our model comes at a cost, since not all the parameters of interest are identified (a limitation that also arises in the benchmark model of Ertur and Koch 2007). In essence, we cannot separate the (direct) effect that, as an input of the production function, the stock of physical capital has on the output from the (indirect) effect that it has as a driver of the technology (or we cannot separate the effect that the own stock of physical capital has on the output –via the technology– from that of the neighbouring economies). This means that, although the models of Islam (1995), López-Bazo et al. (2004) and Ertur and Koch (2007) are nested in ours, we can only statistically reject the validity of that of Islam (1995) and López-Bazo et al. (2004). Still, we argue that simple changes in the model specification (e.g., introducing the stock of physical capital lagged one period in the technological progress, rather than using its current value) and/or appropriate restrictions on the set of parameters (as e.g. Ertur and Koch 2007 do) may address this limitation. We illustrate our argument by constraining some of the parameters to be consistent with either the model of López-Bazo et al. (2004) or that of Ertur and Koch (2007).

The econometric specification of the resulting growth-initial equation corresponds to the spatial Durbin dynamic panel model (see also Elhorst et al., 2010; Yu and Lee, 2012; Ho et al., 2013), but with spatially weighted individual-specific effects. Thus, given the obvious interest

[1] It is worth noting that the model can easily be extended to incorporate the role of human capital (López-Bazo et al., 2004; Fingleton and López-Bazo, 2006). We leave this issue for future research.

in distinguishing the 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 taken on board 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 "observed" technological interdependences in the output per capita of the EU regions, that is, a positive and significant impact of the level of technology of the neighbouring regions. However, there is also evidence of "unobserved" technological interdependences in the EU regions (i.e., spatial contagion in the "unobserved productivity" accounted for the individual effects of the model). In particular, estimates of the individual effects and their spatial spillovers indicate that the richest (poorest) EU regions are likely to stand rich (poor) because of their higher (lower) "unobserved productivity" and/or higher (lower) spillovers. Lastly, our simple identification strategy produces estimates of the implied parameters that support our model specification against that of Islam (1995) and López-Bazo et al. (2004). However, our results are unclear about what technology, the one assumed by López-Bazo et al. (2004) or that assumed by Ertur and Koch (2007), fits the data better.

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

2 The Model

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 extended later 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 in time 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 variales 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. This rules out the existence of knowledge spillovers arising from technological interdependences between the regional economies. However, accounting for technological interdependences and knowledge spillovers is critical when analysing how "the relative location of an economy affects economic growth" (Elhorst et al., 2010). In the literature, depending on whether knowledge spillovers turn out to be "local" or "global" (Anselin, 2003), we find two main approaches to the introduction of spatial externalities in the Solow growth model.

On the one hand, López-Bazo et al. (2004) and Egger and Pfaffermayr (2006) consider growth models where the knowledge spillovers are local in nature, in the sense that they are limited to the neighbouring regions (at least initially).[2] To be precise, in López-Bazo et al. (2004) technology is assumed to depend on both the physical and human capital of the neighbouring regions, whereas in Egger and Pfaffermayr (2006) is assumed to grow exogenously and at the same rate in all regions (as in Mankiw et al. 1992 and Islam 1995), so that the externalities arise from the assumption that total factor productivity depends on the capital-labour ratio of the region and the spatially weighted capital-labour of the other regions. Ertur and Koch (2007), on the other hand, 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 is a geometrically weighted average of the technology of the other economies,

[2] See also Fingleton and López-Bazo (2006), Pfaffermayr (2009) and Pfaffermayr (2012).

thus making knowledge spillovers to spread over all the regions (and hence become "global"). However, it is still assumed that "some proportion of technological progress is exogenous and identical in all countries" [p. 1036].

In this paper, we extend the model of Ertur and Koch (2007) by introducing spatial dependence in the stock of capital, as well as heterogeneity and spatial dependence in the exogenous technological progress (while holding the assumption that the technological progress of an economy depends on the stock of knowledge of the other economies). In this vein, our assumed technology combines the alternative sources of spatial externalities considered in models of relative location with the unobserved heterogeneity that characterises the models of absolute location (Abreu et al., 2005). In particular, our model shares with that of Ertur and Koch (2007) the main source of parameter heterogeneity. Namely, the speed of convergence to the steady state, as discussed below. Yet we eventually estimate a constrained version in which the speed of convergence is identical for all economies (Elhorst et al., 2010; Yu and Lee, 2012). In particular, the econometric specification corresponds to a variant of the spatial Durbin dynamic panel model recently considered by Lee and Yu (2016) that includes not only individual-specific effects but also their spatial spillovers (Miranda et al., 2017a).[3]

Next we derive our empirical specification, which adopts the form of a growth-initial equation. To a large extent, our approach follows the steps of Ertur and Koch (2007). In particular, we first discuss and motivate the assumed technology, then we obtain the output per worker equation at the steady state, and finally the growth-initial equation.

2.2 Technology

Let us denote by Ω_{it} the exogenous technological progress and by $k_{it} = \dfrac{K_{it}}{L_{it}}$ the level of physical capital per worker (of region i in period t). Ertur and Koch (2007, p. 1036) assume that the

[3] As Basile (2008, p. 532-533) points out, "the local Spatial Durbin Model (...) proposed by Ertur and Koch (2007) is a general and flexible specification, since it allows identification of both spatial-interaction effects and parameter heterogeneity (...). In essence, this is the model considered here. The global Spatial Durbin Model (...) represents a less general specification, because it imposes the restriction of parameter homogeneity". In essence, this is the model we estimate. Lastly, "[t]he model proposed by López-Bazo et al. (2004) (..) imposes a further restriction on the parameters since the spatial lags of the structural characteristics of the regions are not included" (this also applies to the model proposed by Egger and Pfaffermayr 2006).

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 "the parameter ϕ describes the strength of home externalities generated by physical capital accumulation" ($0 \leq \phi < 1$) and "the degree of [regional] technological interdependence generated by the level of spatial externalities is described by γ" ($0 \leq \gamma < 1$). 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 $\sum_{j \neq i} w_{ij} = 1$ for $i = 1, \ldots, N$. Lastly, Ertur and Koch (2007) assume that $\Omega_{it} = \Omega_t = \Omega_{t=0} \exp(\mu t)$, where μ is the constant rate of growth of the exogenous technological progress. Therefore, the technology eventually assumed is $A_{it} = \Omega_{t=0} \exp(\mu t) k_{it}^{\phi} \prod_{j \neq i}^{N} A_{jt}^{\gamma w_{ij}}$.

However, as previously pointed out, there are alternative approaches to the inclusion of knowledge spillovers in the Solow model. In a series of papers, López-Bazo et al. (2004, p. 46), Egger and Pfaffermayr (2006), Fingleton and López-Bazo (2006) and Pfaffermayr (2009, 2012) argue that the physical (and human) capital may be an alternative source of externalities, "[t]he reasoning behind such spillovers [being] basically the diffusion of technology from other regions caused by investments in physical (...) capital". In mathematical terms, such a technology may adopt the following functional form:

$$A_{it} = \Omega_{t=0} \exp(\mu t) \prod_{j \neq i}^{N} k_{jt}^{\gamma w_{ij}}, \tag{2.3}$$

where, for the sake of comparability, we have used the same notation as in 2.2. However, the interpretation of the parameter γ ("assumed to be positive") is different here, for it now "measures the [strength of the] externality across economies" originated from variations in physical capital (López-Bazo et al., 2004; Fingleton and López-Bazo, 2006, p. 46). It is also important to stress that these papers maintain the assumption of an homogeneous exogenous technological progress growing at a constant rate, i.e., $\Omega_{it} = \Omega_{t=0} \exp(\mu t)$.

Our assumed technology features those displayed in 2.2 and 2.3. However, we depart from these studies in the assumptions they made with respect to the exogenous technological progress. First, they assume that it is homogeneous across regions. However, as Mankiw et al. (1992, p. 6) point out, "the $\Omega_{t=0}$ term reflects not just technology but resource endowments, climate, institutions, and so on; it may therefore differ across countries". In line with this argument, we introduce regions' heterogeneity into the definition of the exogenous technological progress by assuming that $\Omega_{it} = \Omega_{i0} \exp(\mu t)$.[4] (Lee and Yu, 2016; Miranda et al., 2017a).

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 interpretation is behind our second departure from the models of López-Bazo et al. (2004), Ertur and Koch (2007) and others, since opens the door to consider productivity spillovers as an additional source of spatial externalities (LeSage and Fischer, 2012; Miranda et al., 2017b). As De Long and Summers (1991, p. 487) point out, "it is difficult to believe that Belgian and Dutch or US and Canadian economic growth would ever significantly diverge, or that substantial productivity gaps would appear within Scandinavia".

All in all, a production technology that may account for these alternative sources of spatial dependence is the following:

$$A_{it} = \Omega_{it} \prod_{j \neq i}^{N} \Omega_{jt}^{\gamma_1 w_{ij}} k_{it}^{\phi} \prod_{j \neq i}^{N} k_{jt}^{\gamma_2 w_{ij}} \prod_{j \neq i}^{N} A_{jt}^{\gamma_3 w_{ij}} \tag{2.4}$$

with $\Omega_{it} = \Omega_{i0} \exp(\mu t)$ and Ω_{i0} non-observable (which is why Ω_{it} does not have a coefficient in

[4]Alternative ways of modelling the exogenous technological progress are $\Omega_{it} = \Omega_{t=0} \exp(\mu_i t)$ 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

2.4). Notice that $-1 \leq \gamma_3 \leq 1$ and $\gamma_2 > 0$ play the same role as γ in 2.2 and 2.3, respectively, whereas γ_1, can be interpreted as the degree of technological interdependence generated from the (unobserved) productivity spillovers. In particular, notice that $\gamma_1 = \phi = \gamma_2 = \gamma_3 = 0$ would lead us to the model proposed by Islam (1995), $\gamma_1 = \phi = \gamma_3 = 0$ (and possibly $\gamma_2 \neq 0$) to the model proposed by López-Bazo et al. (2004), and $\gamma_1 = \gamma_2 = 0$ (and possibly $\phi \neq 0$) to the model proposed by Ertur and Koch (2007). However, as shown by Miranda et al. (2017a), the fact that $\gamma_1 \neq 0$ cannot be used to discriminate between these models because they are actually observationally equivalent. Moreover, as discussed below, the parameters ϕ and γ_2 may not be identified. Therefore, whereas a test of $\gamma_3 = 0$ suffices to support our model specification (if rejected) against that of Islam (1995) and López-Bazo et al. (2004), discriminating empirically between our model and that of Ertur and Koch (2007) requires of additional identification conditions.[5]

2.3 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 2.4 expressed in logs and matrix form:

$$\begin{aligned} A &= \Omega + \gamma_1 W\Omega + \phi k + \gamma_2 W k + \gamma_3 W A \\ &= (I - \gamma_3 W)^{-1}\Omega + \gamma_1 (I - \gamma_3 W)^{-1} W\Omega + \phi (I - \gamma_3 W)^{-1} k + \gamma_2 (I - \gamma_3 W)^{-1} W k \qquad (2.5) \end{aligned}$$

where the parameters γ_1, γ_2 and γ_3 have been previously described, A is the $N \times 1$ vector of logarithms of the technology, $\Omega = \Omega_0 + \iota_N \mu t$ 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, k is the $N \times 1$ vector of of logarithms of the capital per worker, and W is the $N \times N$ spatial weight matrix that describes the spatial arrangement of the regions.

[5] We return to this issue in section 3.1, where we discuss simple identification strategies for the growth-initial equation we eventually estimate.

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

$$\begin{aligned}
\ln A_{it} &= \sum_{j=1}^{N}\sum_{r=0}^{\infty}\gamma_3^r w_{ij}^{(r)}\ln\Omega_{jt} + \gamma_1\sum_{j=1}^{N}\sum_{r=0}^{\infty}\gamma_3^r w_{ij}^{(r+1)}\ln\Omega_{jt} + \phi\sum_{j=1}^{N}\sum_{r=0}^{\infty}\gamma_3^r w_{ij}^{(r)}\ln k_{jt} \\
&+ \gamma_2\sum_{j=1}^{N}\sum_{r=0}^{\infty}\gamma_3^r w_{ij}^{(r+1)}\ln k_{jt} \\
&= \sum_{j=1}^{N}\ln\Omega_{jt}^{\sum\limits_{r=0}^{\infty}\gamma_3^r w_{ij}^{(r)}} + \sum_{j=1}^{N}\ln\Omega_{jt}^{\frac{\gamma_1}{\gamma_3}\sum_{r=1}^{\infty}\gamma_3^r w_{ij}^{(r)}} + \sum_{j=1}^{N}\ln k_{jt}^{\phi\sum\limits_{r=0}^{\infty}\gamma_3^r w_{ij}^{(r)}} + \sum_{j=1}^{N}\ln k_{jt}^{\frac{\gamma_2}{\gamma_3}\sum\limits_{r=1}^{\infty}\gamma_3^r w_{ij}^{(r)}},
\end{aligned}$$

we may rewrite 2.5 as

$$\begin{aligned}
A_{it} &= \prod_{j=1}^{N}\Omega_{jt}^{\sum\limits_{r=0}^{\infty}\gamma_3^r w_{ij}^{(r)}}\prod_{j=1}^{N}\Omega_{jt}^{\frac{\gamma_1}{\gamma_3}\sum\limits_{r=1}^{\infty}\gamma_3^r w_{ij}^{(r)}}\prod_{j=1}^{N}k_{jt}^{\phi\sum\limits_{r=0}^{\infty}\gamma_3^r w_{ij}^{(r)}}\prod_{j=1}^{N}k_{jt}^{\frac{\gamma_2}{\gamma_3}\sum\limits_{r=1}^{\infty}\gamma_3^r w_{ij}^{(r)}} \\
&= \Omega_{it}^{1+\left(\frac{\gamma_3+\gamma_1}{\gamma_3}\right)\sum_{r=1}^{\infty}\gamma_3^r w_{ii}^{(r)}}\prod_{j\neq i}^{N}\Omega_{jt}^{\left(\frac{\gamma_3+\gamma_1}{\gamma_3}\right)\sum_{r=1}^{\infty}\gamma_3^r w_{ij}^{(r)}}k_{it}^{\phi+\left(\frac{\phi\gamma_3+\gamma_2}{\gamma_3}\right)\sum_{r=1}^{\infty}\gamma_3^r w_{ii}^{(r)}}\prod_{j\neq i}^{N}k_{jt}^{\left(\frac{\phi\gamma_3+\gamma_2}{\gamma_3}\right)\sum_{r=1}^{\infty}\gamma_3^r w_{ij}^{(r)}}
\end{aligned}$$

by using $\prod_{j=1}^{N}\Omega_{jt}^{w_{ij}^{(0)}} = \Omega_{it}$ and $\prod_{j=1}^{N}k_{jt}^{\phi w_{ij}^{(0)}} = k_{it}^{\phi}$.

Also, let us now define $u_{ii} = \alpha+\phi+\left(\frac{\phi\gamma_3+\gamma_2}{\gamma_3}\right)\sum_{r=1}^{\infty}\gamma_3^r w_{ii}^{(r)}$ and $u_{ij} = \left(\frac{\phi\gamma_3+\gamma_2}{\gamma_3}\right)\sum_{r=1}^{\infty}\gamma_3^r w_{ij}^{(r)}$, with $u_{ii}+\sum_{j\neq i}^{N}u_{ij} = \sum_{j=1}^{N}u_{ij} = \alpha+\phi+\frac{\phi\gamma_3+\gamma_2}{1-\gamma_3} = \alpha+\frac{\phi+\gamma_2}{1-\gamma_3}$. Then, given that $y_{it} = A_{it}k_{it}^{\alpha}$,

$$y_{it} = \Omega_{it}^{1+\left(\frac{(\gamma_3+\gamma_1)(u_{ii}-\alpha-\phi)}{(\phi\gamma_3+\gamma_2)}\right)}\prod_{j\neq i}^{N}\Omega_{jt}^{\frac{(\gamma_3+\gamma_1)u_{ij}}{\phi\gamma_3+\gamma_2}}k_{it}^{u_{ii}}\prod_{j\neq i}^{N}k_{jt}^{u_{ij}} \qquad (2.6)$$

Notice that "this model implies spatial heterogeneity in the parameters of the production function", a feature shared with that of Ertur and Koch (2007, p. 1037). Also, in contrast to the local contagion models of López-Bazo et al. (2004) and Egger and Pfaffermayr (2006), both ours and that of Ertur and Koch (2007) are models of global contagion (Anselin, 2003). We differ, however, in that whereas in their case there are no (global) spatial externalities unless $\gamma_3 \neq 0$, there still are here if either $\gamma_1 \neq 0$ or $\gamma_2 \neq 0$ (albeit of a local nature). This is because our model features both global and local contagion. We also differ 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 $\gamma_1 = \gamma_2 = \gamma_3 = 0$.

2.4 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 matrix form, $y = A + \alpha k$, and substitute the technology by its expression in 2.5. We then pre-multiply both sides of the resulting equation by $I - \gamma_3 W$ to obtain (provided that $\gamma_3 \neq 0$ and $1/\gamma_3$ is not an eigenvalue of W)

$$y = \Omega + \gamma_1 W \Omega + (\alpha + \phi)k + (\gamma_2 - \alpha\gamma_3)Wk + \gamma_3 W y \tag{2.7}$$

Lastly, we replace in this equation 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:

$$\dot{k}_{it} = s_i y_{it} - (n_i + \delta)k_{it} \tag{2.8}$$

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 2.8 implies that the capital-output ratio is constant and converges to a balanced growth rate g defined by $\dfrac{\dot{k}_{it}}{k_{it}} = \dot{ln\, y_{it}} = \dot{ln\, k_{it}} = g = \dfrac{\mu\,(1+\gamma_1)}{(1-\gamma_3)(1-\alpha) - \phi - \gamma_2}$ (see appendix *A*). Also, it can be shown that, given a balanced growth rate g and 2.8 (see e.g. Barro and Sala-i-Martin, 2003), $\frac{k_{it}^*}{y_{it}^*} = \frac{s_i}{n_i + \delta + g}$ and $\ln k_{it}^* = \ln y_{it}^* + \ln \left(\frac{s_i}{n_i + \delta + g} \right)$.[6]

What is thus left is to introduce in 2.7 (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:

$$\begin{aligned} \ln y_{it}^* = {} & \frac{\ln \Omega_{it}}{1-\alpha-\phi} + \frac{\gamma_1}{1-\alpha-\phi} \sum_{j=1}^{N} w_{ij} \ln \Omega_{jt} + \frac{\alpha+\phi}{1-\alpha-\phi} \ln \left(\frac{s_i}{n_i+\delta+g} \right) \\ & + \frac{\gamma_2 - \alpha\gamma_3}{1-\alpha-\phi} \sum_{j=1}^{N} w_{ij} \left(\frac{s_i}{n_i+\delta+g} \right) + \frac{(1-\alpha)\gamma_3 + \gamma_2}{1-\alpha-\phi} \sum_{j=1}^{N} w_{ij} \ln y_{jt}^* \end{aligned} \tag{2.9}$$

Notice that this equation differs from that obtained by Ertur and Koch (2007) in two main features (beyond the appearance of γ_2), reflecting ultimately differences in the assumed technology. First, the heterogeneous exogenous technological progress, since Ω_{it} is assumed to be Ω_t in Ertur and Koch (2007). Second, the term $\frac{\gamma_1}{(1-\alpha-\phi)} \sum_{j=1}^{N} w_{ij} \ln \Omega_{jt}$, which is missing in their steady state equation because they assume that there are no spatial externalities in the exogenous technological progress. More generally, these features of our model are also absent in the above mentioned growth studies (López-Bazo et al., 2004; Egger and Pfaffermayr, 2006; Fingleton and López-Bazo, 2006; Pfaffermayr, 2009, 2012).

---

[6] It is also interesting to note that, if we compute the marginal productivity of capital, $\frac{\dot{k}_{it}}{k_{it}} = s_i \frac{y_{it}}{k_{it}} - (n_i + \delta)$, using the expression defining y_{it} in 2.6, we obtain $\frac{\dot{k}_{it}}{k_{it}} = s_i \Omega_{it}^{1 + \left(\frac{(\gamma_3+\gamma_1)(u_{ii}-\alpha-\phi)}{(\phi\gamma_3+\gamma_2)} \right)} \prod_{j \neq i}^{N} \Omega_{jt}^{\frac{(\gamma_3+\gamma_1)u_{ij}}{\phi\gamma_3+\gamma_2}} k_{it}^{u_{ii}-1} \prod_{j \neq i}^{N} k_{jt}^{u_{ij}} - (n_i + g)$. Therefore, provided that $\alpha + \frac{\phi+\gamma_2}{1-\gamma_3} < 1$, there are diminishing returns to the capital, as in the model of Ertur and Koch (2007, p. 1039). We differ, however, in that in our case it is not only the variations in capital what make "the rate of growth [to vary and converge] to its own steady state", but also the variations in the exogenous technological progress.

2.5 The growth-initial equation

In the standard, non-spatial growth models (see e.g. Barro and Sala-i-Martin, 2003), the analog of equation 2.9 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 $\dfrac{(1-\alpha)\gamma_3+\gamma_2}{1-\alpha-\phi}\sum_{j=1}^{N} w_{ij}\ln y^*_{jt}$ is missing). Thus, a log-linear approximation to the dynamics around the steady state using a Taylor expansion produces a growth-initial regression equation that can be estimated using the appropriate method. In our case, however, this approach would produce a rather complex system of first-order differential linear equations whose solution is not directly estimable due to the presence of variables at the steady state (Egger and Pfaffermayr, 2006, for example, approximate them using a set of exogenous variables). In particular, a log linearisation of the marginal productivity of capital, $\dfrac{\dot{k}_{it}}{k_{it}}$, around the steady state yields (see appendix *B*)

$$\frac{\dot{k}_{it}}{k_{it}} = g + (u_{ii}-1)(n_i+\delta+g)\left(\ln k_i(t) - \ln k^*_{it}\right) + \sum_{j\neq i}^{N} u_{ij}(n_i+\delta+g)\left(\ln k_{jt} - \ln k^*_{jt}\right) \tag{2.10}$$

Notice that this result coincides with the one obtained by Ertur and Koch (2007).

To tackle this issue, Ertur and Koch (2007) hypothesise that the differences between the observed and the steady state values of the capital and output per worker across regions correspond to the following expressions:

$$\begin{aligned} \ln y_{it} - \ln y^*_{it} &= \Theta_j\left(\ln y_{jt} - \ln y^*_{jt}\right) \\ \ln k_{it} - \ln k^*_{it} &= \Phi_j\left(\ln k_{jt} - \ln k^*_{jt}\right) \end{aligned} \tag{2.11}$$

This yields the following speed of convergence (see appendix *C*):

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

with

$$\lambda_i = \frac{\sum_{j=1}^{N} u_{ij} \frac{1}{\Phi_j}(n_j + g + \delta)}{\sum_{j=1}^{N} u_{ij} \frac{1}{\Phi_j}} - \sum_{j=1}^{N} u_{ij}(n_j + \delta + g)\frac{1}{\Theta_j} \tag{2.13}$$

Solving the differential equation in 2.12 for $\ln y_{it}$ (see appendix *D*), evaluating the solution at $t = t_2$:

$$\ln y_{it_2} = g\left(t_2 - t_1 e^{-\lambda_i \tau}\right) - e^{-\lambda_i \tau} \ln y_{it_1} + (1 - e^{-\lambda_i \tau}) \ln y_{i0}^* \tag{2.14}$$

with $\tau = t_2 - t_1$. In particular, under the assumption that the speed of convergence is homogeneous across regions ($\lambda_i = \lambda$ for $i = 1, \cdots, N$):

$$\ln y_{it_2} = g\left(t_2 - t_1 e^{-\lambda \tau}\right) - e^{-\lambda \tau} \ln y_{it_1} + (1 - e^{-\lambda \tau}) \ln y_{i0}^* \tag{2.15}$$

At this point it is convenient to write the previous expression in matrix form:

$$y(t_2) = g\left(t_2 - t_1 e^{-\lambda\tau}\right)\iota_N - \left(1 - e^{-\lambda\tau}\right)y(t_1) + \left(1 - e^{-\lambda\tau}\right)y^*(0) \tag{2.16}$$

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 facilitates replacing $y^*(0)$ by 2.9 at $t = 0$, which, in matrix form, is:

$$y^*(0) = (I - \rho W)^{-1}\left[\frac{1}{1-\alpha-\phi}\Omega(0) + \frac{\gamma_1}{1-\alpha-\phi}W\Omega(0) + \frac{\alpha+\phi}{1-\alpha-\phi}S + \frac{\gamma_2 - \alpha\gamma_3}{1-\alpha-\phi}WS\right] \tag{2.17}$$

where $\rho = \dfrac{(1-\alpha)\gamma_3 + \gamma_2}{1-\alpha-\phi}$ and $S = \left\{\ln\left(\dfrac{s_i}{n_i + \delta + g}\right)\right\}_{i=1,\ldots,N}$.

Thus, we introduce 2.17 in 2.16 and pre-multiply both sides of the resulting equation by

$I - \rho W$ to obtain:

$$
\begin{aligned}
y(t_2) &= g(1-\rho)\left(t_2 - t_1 e^{-\lambda\tau}\right)\iota_N + e^{-\lambda\tau}\left(I - \rho W\right)y(t_1) + \rho W y(t_2) \\
&\quad + \left(1 - e^{-\lambda\tau}\right)\left[\frac{1}{1-\alpha-\phi}\Omega(0) + \frac{\gamma_1}{1-\alpha-\phi}W\Omega(0) + \frac{\alpha+\phi}{1-\alpha-\phi}S + \frac{\gamma_2-\alpha\gamma_3}{1-\alpha-\phi}WS\right]
\end{aligned}
\tag{2.18}
$$

Alternatively, we can rewrite this equation for country i as

$$
\begin{aligned}
\ln y_{it_2} &= e^{-\lambda\tau}\ln y_{it_1} - \rho e^{-\lambda\tau}\sum_{j=1}^{N} w_{ij}\ln y_{jt_1} + \rho\sum_{j=1}^{N} w_{ij}\ln y_{jt_2} \\
&\quad + \frac{\left(1-e^{-\lambda\tau}\right)(\alpha+\phi)}{1-\alpha-\phi}\ln s_i - \frac{\left(1-e^{-\lambda\tau}\right)(\alpha+\phi)}{1-\alpha-\phi}\ln(n_i+\delta+g) \\
&\quad + \frac{\left(1-e^{-\lambda\tau}\right)(\gamma_2-\alpha\gamma_3)}{1-\alpha-\phi}\sum_{j=1}^{N} w_{ij}\ln s_j - \frac{\left(1-e^{-\lambda\tau}\right)(\gamma_2-\alpha\gamma_3)}{1-\alpha-\phi}\sum_{j=1}^{N} w_{ij}\ln(n_j+\delta+g) \\
&\quad + \left(\frac{\left(1-e^{-\lambda\tau}\right)}{1-\alpha-\phi}\ln\Omega_{i0}\right) + \left(\frac{\left(1-e^{-\lambda\tau}\right)\gamma_1}{1-\alpha-\phi}\sum_{j=1}^{N} w_{ij}\ln\Omega_{j0}\right) \\
&\quad + g(1-\rho)\left(t_2 - t_1 e^{-\lambda\tau}\right)
\end{aligned}
\tag{2.19}
$$

3 Empirical results

3.1 Model specification and identification strategies

To derive our econometric specification, notice that equation 2.19 (plus an i.i.d. shock ε and under the assumption that the speed of convergence is identical for all the economies, $\lambda_i = \lambda$) corresponds to the spatial Durbin dynamic panel model with individual-specific effects and their spatial spillovers:

$$
\begin{aligned}
y_{it} = \overline{\gamma}_1 y_{i,t-1} + \overline{\gamma}_2 \sum_{j=1}^{N} w_{ij} y_{j,t-1} + \rho \sum_{j=1}^{N} w_{ij} y_{jt} + x_{it1}\beta_1 + x_{it2}\beta_2 \\
+ \sum_{j=1}^{N} w_{ij} x_{jt1}\theta_1 + \sum_{j=1}^{N} w_{ij} x_{jt2}\theta_2 + \overline{\mu}_i + \sum_{j=1}^{N} w_{ij}\overline{\alpha}_j + f_t + \varepsilon_{it}
\end{aligned}
\tag{3.1}
$$

where $y_{it} = \ln y_{t_2}$, $y_{i,t-1} = \ln y_{t_1}$, $x_{it1} = \ln s_i$, $x_{it2} = \ln(n_i + \delta + g)$, $\overline{\gamma}_1 = e^{-\lambda\tau}$, $\overline{\gamma}_2 = -\rho e^{-\lambda\tau}$, $\beta_1 = \dfrac{\left(1-e^{-\lambda\tau}\right)(\alpha+\phi)}{1-\alpha-\phi}$, $\beta_2 = -\dfrac{\left(1-e^{-\lambda\tau}\right)(\alpha+\phi)}{1-\alpha-\phi}$, $\theta_1 = \dfrac{\left(1-e^{-\lambda\tau}\right)(\gamma_2-\alpha\gamma_3)}{1-\alpha-\phi}$, $\theta_2 = -\dfrac{\left(1-e^{-\lambda\tau}\right)(\gamma_2-\alpha\gamma_3)}{1-\alpha-\phi}$, $\overline{\mu}_i = \dfrac{\left(1-e^{-\lambda\tau}\right)}{1-\alpha-\phi}\ln\Omega_{i0}$, $\overline{\alpha}_i = \dfrac{\left(1-e^{-\lambda\tau}\right)\gamma_1}{1-\alpha-\phi}\ln\Omega_{i0}$ 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 distinguishes the spatial counterparts of the individual effects. In other words, their individual effects correspond to $\overline{\mu}_i + \sum_{j=1}^{N} w_{ij}\overline{\alpha}_j$ in 3.1. In fact, in our model the individual effects and their spatial counterparts are proportional (by a rate γ_1). 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 ($\overline{\mu}_i$) and their spatial spillovers ($\overline{\alpha}_i$). This means making use of the following correlation functions (Mundlak, 1978; Chamberlain, 1982):

$$
\begin{aligned}
\overline{\mu}_i &= c_i + \pi_{\mu 1}\left(\frac{1}{T}\sum_{t=1}^{T} x_{it1}\right) + \pi_{\mu 2}\left(\frac{1}{T}\sum_{t=1}^{T} x_{it1}\right) + \upsilon_{\mu i} \\
\overline{\alpha}_i &= \pi_{\alpha_1}\left(\frac{1}{T}\sum_{t=1}^{T} x_{it1}\right) + \pi_{\alpha_2}\left(\frac{1}{T}\sum_{t=1}^{T} x_{it2}\right) + \upsilon_{\alpha i},
\end{aligned}
\tag{3.2}
$$

where c_i is the constant term to be estimated, $\pi_{\mu 1}$, $\pi_{\mu 2}$, $\pi_{\alpha 1}$ and $\pi_{\alpha 2}$ 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 $Cov(\upsilon_{\mu i}, \upsilon_{\alpha i}) = \sigma_{\mu\alpha i}$.

The last thing to notice about our econometric specification is that not all the implied parameters (ρ, λ, α, ϕ, γ_1, γ_2, γ_3, and $\ln \Omega_{i0}$) are identified. To be precise, we do obtain a direct estimate of ρ from 3.1 and can easily obtain an estimate of λ from $\overline{\gamma_1}$. Further, we can obtain an estimate of γ_1 by regressing $\overline{\alpha}_i$ on $\overline{\mu}_i$. These parameters are thus identified. In contrast, we cannot obtain a single estimate of γ_3 (since this requires $\hat{\rho}$, $\overline{\gamma_1}$, β_1 or β_2, and θ_1 or θ_2) and $\ln \Omega_{i0}$ (since this requires $\overline{\gamma_1}$, $\overline{\mu}_i$ and β_1 or β_2) because these parameters are overidentified. Also, we cannot identify α, ϕ and γ_2 (only $\alpha + \phi$ and $\gamma_2 - \alpha\gamma_3$). The under-identification problem here, as previously pointed out, is that we cannot separate the effect that, as an input of the production function, the stock of physical capital has on the output (i.e., α) from the effect that it has as a driver of the technology (i.e., ϕ). Neither can we separate the effect that the own stock of physical capital has on the technology and, subsequently, the output (i.e., ϕ), from that of the neighbouring economies (i.e., γ_2). Still, there are ways to circumvent this identification problem.

One way is to modify the specification of the model. There are no identification problem, for example, if we are willing to assume that the stock of physical capital enters the technological progress lagged one period. That is, if we are willing to assume that $A_{it} = \Omega_{it} \prod_{j \neq i}^{N} \Omega_{jt}^{\gamma_1 w_{ij}} k_{it-1}^{\phi} \prod_{j \neq i}^{N} k_{jt-1}^{\gamma_2 w_{ij}} \prod_{j \neq i}^{N} A_{jt}^{\gamma_3 w_{ij}}$ (see, in contrast, equation 2.4). Neither there are if we argue that different arguments of the technology require different weight matrices. In mathematical terms, this means assuming that $A_{it} = \Omega_{it} \prod_{j \neq i}^{N} \Omega_{jt}^{\gamma_1 w_{ij}^{\Omega}} k_{it}^{\phi} \prod_{j \neq i}^{N} k_{jt1}^{\gamma_2 w_{ij}^{k}} \prod_{j \neq i}^{N} A_{jt}^{\gamma_3 w_{ij}^{A}}$, where w_{ij}^{Ω}, w_{ij}^{k} and w^{A} denote different weight matrices (see e.g. Lee and Yu, 2016).

These approaches, however, involve the derivation of a new model (the steady state equation and the speed of convergence, for example, would surely be altered) and/or require additional data to construct the weight matrices (in our empirical application, we may for example need data on bilateral trade flows and geographical distances between the EU regions). We thus leave these approaches for future research and concentrate here on a less demanding approach to identification. Namely, the use of appropriate constraints on the set of parameters (Ertur and Koch, 2007).

Equation 3.1 yields two main constraints between the parameters: $\beta_1 = -\beta_2$ and $\theta_1 = -\theta_2$ (see also Ertur and Koch, 2007). Thus, by imposing these constraints in 3.1, we obtain a "restricted" version of our model specification that, provided that we can find an additional constraint, is identified. This means that, if the data supports that $\beta_1 = -\beta_2$ and $\theta_1 = -\theta_2$, then γ_3 and $\ln \Omega_{i0}$ are identified and we only require an additional constraint to achieve the identification of the rest of parameters. One may be willing to assume, for example, that the impact of the own physical stock and that of the other economies in the level of technology is the same (i.e., $\phi = \gamma_2$). If that were the case, we may obtain an estimate of α and $\phi = \gamma_2$ from $\hat{\gamma}_3$, β_1 and θ_1. However, since our assumed technology encompasses that of López-Bazo et al. (2004) and Ertur and Koch (2007), we find that it is of greater interest to constrain one of the unidentified implied parameters to be consistent with either the model of López-Bazo et al. (2004) or that of Ertur and Koch (2007), and then obtain an estimate of the rest of implied parameters. Thus, under the assumption that the technology assumed by López-Bazo et al. (2004) is the appropriate (i.e., $\phi = 0$, we can obtain an estimate of α and γ_2, whereas under the assumption that the technology assumed by Ertur and Koch (2007) is the appropriate (i.e., $\gamma_2 = 0$), we can obtain an estimate of α and ϕ. Finding reasonable and statistically significant values for these estimates may help to assess the validity of our model against these alternatives.

3.2 Estimates from EU-NUTS2 regions

We use EU NUTS2 regional data from Cambridge Econometrics to estimate the model given by 3.1 and 3.2. The original data set covers 263 regions across 15 countries (Austria, Belgium, Germany, Denmark, Greece, Finland, France, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden and the United Kingdom) over the period 1980 to 2015. However, for our final sample we basically considered continental regions (Spanish, French and Italian islands, for example, were not included) and dropped the eastern Lander and Berlin as well as the Dutch region of Flevoland (because of missing data for the 1980s). We end up dealing with 193 regions. As for the time dimension, we use time intervals of 5 years —see, among others, Elhorst et al. (2010); Ho et al. (2013); Lee and Yu (2016). Our starting point is 1980 and the final year is 2015, thus resulting in 7 time periods and a balanced panel dataset.

The dependent variable is the real GDP per capita (real GDP at 2005 constant prices over total population, in thousands of people), s is the ratio between investment expenditures and gross value-added (at 2005 constant prices and as a percentage), and n is the growth rate of the working population over time (computed as in Islam 1995). As it is common in the literature (see e.g. Mankiw et al., 1992; Islam, 1995; Ertur and Koch, 2007), we assume that $\delta + g = 0.05$. Notice also that time dummies were included to account for f_t, but their coefficient estimates are not reported to save space.

[Insert Table 1 about here]

Table 1 provide descriptive statistics of these variables for the 6 periods effectively used in estimation (due to the inclusion of the lagged dependent variable in the model). If we compare the reported values with those reported by Ho et al. (2013) for 26 OECD countries over the period 1970 to 2005, we can see that most regions have a real outcome below the average of the OECD countries (the observed differences in means are way beyond the obvious differences that may arise due to the differences in time periods and/or measurement units). It is also interesting to note that the savings rate seems to be larger and the growth rate of the working population smaller in the European regions than in the OECD countries.

We estimate the model (both the unrestricted and restricted versions) using the approach and model specifications of Lee and Yu (2016) and Miranda et al. (2017a). We use the first 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 second to obtain the whole set of estimates (i.e., the basic ones plus those appearing in the correlation functions: c_i, $\pi_{\mu 1}$, $\pi_{\mu 2}$, $\pi_{\alpha 1}$ and $\pi_{\alpha 2}$); test the restricted version of the model (i.e., testing the constraints $\beta_1 = -\beta_2$ and $\theta_1 = -\theta_2$); and estimate some of the unidentified implied parameters (using the restricted version of the model and, when required, imposing an additional constraint). In essence, this is also how we have organised the discussion of the results. We will start with an analysis of the estimates of the basic and correlation functions parameters (plus ρ) in the unrestricted and restricted versions of the model, then will go on with the estimates of the implied parameters, and will conclude with a description of the geographical distribution of $\ln \hat{\Omega}_{i0}$ (the estimated "unobserved productivity" of the EU regions) and its estimated spatial spillover ($\hat{\gamma}_1 \sum_{j=1}^{N} w_{ij} \ln \hat{\Omega}_{j0}$).

[Insert Table 2 about here]

We report the estimates of the unrestricted version of the model in the last two columns of Table 2. The first column was obtained using the approach and model specification of Lee and Yu (2016), whereas the second was obtained using that Miranda et al. (2017a). Interestingly, both sets of estimates provide essentially the same picture. First, the spatial and time lagged dependent variables ($\sum_{j=1}^{N} w_{ij} y_{j,t}$ and $y_{i,t-1}$, respectively) have a high and positive coefficient, whereas the spatially weighted lagged dependent variable ($\sum_{j=1}^{N} w_{ij} y_{j,t-1}$) has a negative and smaller coefficient (see also Ho et al., 2013; Lee and Yu, 2016). The statistical significance of these parameters indicates that 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. In other words, the richest areas are likely to stand rich whereas the poorest ones can only (partially) catch up if they are geographically close to rich areas. Second, the saving rate is not statistically significant and the growth rate of labour is only statistically significant when using the approach of Miranda et al. (2017a). Thus, these factors may not directly contribute to the growth of the European regions. However, the saving rate of the neighbours does show the expected positive and statistically significant effect. Third, the (mean) growth rate of labour and its spatial counterpart are statistically significant variables in the correlations functions. In addition, all the variance components are statistically significant. This supports our correlated random effects model specification. In particular, there is evidence of correlation between the individual effects and the covariates (since π_{μ_2} is statistically significant) and there is evidence of spatial contagion in the individual effects (since π_{α_2} is statistically significant).

Our estimates of the basic parameters are largely consistent with those reported by Pfaffermayr (2009), using an analogous sample of regions (plus Switzerland's regions) and an earlier period of analysis, and Andreano et al. (2017), using an analogous sample and period of analysis. In contrast, we find some differences with those reported by Basile (2008), who consider an earlier period of analysis (and do not consider Norway's regions). To be precise, the signs of β_1, β_2 and θ_1 concur, but the statistical significance differs. Yet we find a similar statistically significant value for ρ.[7]

[Insert Table 3 about here]

As for the estimates of the restricted version of the model, we report them in the last two columns of Table 2. Again, we report estimates based on Lee and Yu (2016) in the first column and estimates based on Miranda et al. (2017a) in the second. It should be noted, however, that we only find weak evidence supporting the restricted model. Although the sum of θ_1 and θ_2 is not statistically different from zero at standard levels (the Wald test statistic is 1.95, with a p-value of 0.16), the sum of β_1 and β_2 (the Wald test statistic is 3.33, with a p-value of 0.07) and the joint test (the Wald test statistic is 13.62, with a p-value of 0.00) clearly reject the null. This may explain why the estimates of $\beta = \beta_1 = -\beta_2$ and $\theta = \theta_1 = -\theta_2$ yield opposite signs to what is usually found in the literature (Ertur and Koch, 2007; Elhorst et al., 2010).

[Insert Table 4 about here]

With this in mind, next we consider the estimated implied parameters, which are reported in Table 4. In particular, the first block of Table 4 reports results from the unrestricted model on those parameters that are directly identified from equation 3.1, λ and γ_1 (results on ρ have been previously discussed), whereas the second block reports these parameters but now estimated from the restricted model, as well as one of the parameters that is only identified in the restricted model, γ_3 ($\ln \hat{\Omega}_{i0}$ is analysed later). The last block of Table 4 reports results from our identification strategy. This means that the (first) reported estimates of α and γ_2 were obtained from the restricted model (i.e. imposing the constraints $\beta_1 = -\beta_2$ and $\theta_1 = -\theta_2$) under the assumption that the technology considered by López-Bazo et al. (2004) is the appropriate (i.e., under the additional assumption that $\phi = 0$), whereas the (second) reported estimates of

[7] Our results also differ from those reported in panel data studies analysing countries rather than regions (see e.g. Ho et al., 2013; Lee and Yu, 2016). Basically, the estimates and statistical significance of $\overline{\gamma}_1$, $\overline{\gamma}_2$, ρ and θ_2 are similar, but the signs of β_1 and β_2 are the opposite.

α and ϕ were obtained from the restricted model under the assumption that the technology considered by Ertur and Koch (2007) is the appropriate (i.e., under the additional assumption that $\gamma_2 = 0$).

First, the estimated speed of convergence, 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). Second, the impact of the (unobserved) productivity spillovers, measured by γ_1, is negative, statistically significant, and more than proportional (which is implausible...NO????). Third, these two findings remain largely the same when the estimates are obtained from the restricted model, except perhaps that the estimated degree of unobserved technological interdependence is closer to minus one (the theoretical lower bound???). Fourth, the statistical significance of the degree of observed technological interdependence, γ_3, contradicts the models of Islam (1995) and López-Bazo et al. (2004). As a caveat, however, its value is substantially smaller (about a half) than the one found by Ertur and Koch (2007) and Elhorst et al. (2010). Fifth, the estimated parameters obtained by imposing an additional constraint as an identification strategy are generally not statistically significant (and the one that is, α yields values outside the theoretical bounds). It is therefore not possible to determine which technology, that of López-Bazo et al. (2004) or that of Ertur and Koch (2007) fits the data better.

All in all, these results point to the the existence of spatial spillovers in the unobserved productivity and the level of technology. In contrast, there is no sign of the capital externalities found by either López-Bazo et al. (2004) or Ertur and Koch (2007). Also, our estimates support our model specification against that of Islam (1995) and López-Bazo et al. (2004). On the other hand, estimates obtained from the simple identification strategy we devised do not allow us to discriminate between the technology assumed by López-Bazo et al. (2004) and that assumed by Ertur and Koch (2007).

[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 restricted model) in Figure 1. In particular, Figure 1 presents a map of the European regions considered and the values of these statistics grouped by quantiles: Figure 1a reports $\ln \hat{\Omega}_{i0}$ (the "unobserved productivity") whereas Figure 1b reports $\hat{\gamma}_1 \sum_{j=1}^{N} w_{ij} \ln \hat{\Omega}_{j0}$ (the spatial spillover of the the "unobserved productivity", that is, the impact on the GDP per capita of unit i of all the units neighbouring i having their "unobserved productivity").

Results indicate that the regions with the lowest estimated "unobserved productivity" are mostly located in Scandinavia (Finland and Sweden, but also the North of Norway), Denmark, Scotland, Northern Ireland, Central-East of France, Austria, and the South-West (Portugal and Spain) and South-East (South of Italy and Greece) of Europe. Figure 1a also shows that the geographical distribution of the higher estimated "unobserved productivity" largely follows the so-called "blue banana" (from the South West of the UK to the South-West of Germany, thus covering the North of France and the Benelux), plus the Mediterranean regions of the South-West of France and the North of Italy. It is also worth noting the high values found in the Southern areas of Ireland and Norway.

Most of the regions in the high productivity group can be qualified as "rich", meaning here that their average GDP per capita over the period is in the upper quantile of the distribution. On the other hand, the same criterion would lead us to qualify most of the regions with low estimated productivities as "poor". Thus, it seems that richer/poorer regions tend to have higher/lower (unobserved) productivities. Notice, however, that the marginal effect of $\ln \Omega_{i0}$ on the GDP per capita also depends on the values of $(I - \rho W)^{-1}$, $\overline{\gamma}_1$ and $\overline{\gamma}_2 W$ (Debarsy, 2012).

As for the spillovers associated with the "unobserved productivity", Figure 1b reveals that the pattern tends to reverse mirror the one found for the estimated "unobserved productivity". This is expected given the negative and statistically significant estimate found for $\hat{\gamma}_1$. Largest values are found in the Northern regions (i.e., Ireland, the UK Midlands, Scandinavia and Denmark), but also in the East (i.e., Austria) and South (Portugal, Center and West of Spain, South of Italy and Greece). This means that these are (mostly rich) regions whose output per capita is more impacted by the "unobserved productivity" of its neighbours. South of England and Norway, East Germany, the South-East of France, North-East of Spain and the North of Italy, on the other hand, stand as the areas with the lowest spillovers. This means that these are regions whose ouput per capita is barely affected by the "unobserved productivity" of its neighbours. Since these are generally regions with low levels of GDP per capita and "unobserved productivity", our results indicate that poor regions are unlikely to increase their wealth via spillovers effects (unless these originate from the saving rates).

4 Conclusions

We present a growth model that extends previous knowledge-spillovers models in several directions. First, we do not assume a common exogenous technological progress but account for heterogeneity in the initial level of technology. Second, we assume that the technological progress depends not only on the stock of physical capital and the stock of knowledge of the other economies, but also on the physical capital and the (unobserved) initial level of technology of the other economies. Thus, our assumed technology combines the alternative sources of spatial externalities considered in previous models of relative location with the unobserved heterogeneity that characterises previous models of absolute location.

We 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. As a downside, some of the implied parameters of the model are not identified. However, we discuss alternative ways to circumvent this limitation.

We estimate the model by QML using a correlated random effects specification for the individual effects and their spatial spillovers. Results support our model specification. In particular, we find evidence of the existence of 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 its past GDP per capita and the current and past GDP per capita of their neighbours. Further, richest areas (e.g., the "blue banana") are likely to stand rich because of their higher "unobserved productivity" and/or higher spillovers in the "unobserved productivity". Thus, poor regions can only (partially) catch up if they are geographically close to rich areas whose "unobserved productivity" spills over the neighbouring regions.

Table 1: **Descriptive statistics.**

Variable	Mean	SD	Min	P25	Median	P75	Max
GDP	24,205	9,594	5,846	18,159	22,890	28,058	97,112
s	23.80	4.82	8.96	20.94	23.53	26.10	46.30
$n + \delta + g$	0.06	0.01	0.02	0.05	0.06	0.06	0.10

Note: Number of observations $= 193 \times 6 = 1,158$. 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.**

Variable	Parameter	Estimates I	Estimates II
$y_{i,t-1}$	$\overline{\gamma}_1$	0.7402^{***}	0.8850^{***}
		(0.0210)	(0.0178)
$\sum_{j=1}^{N} w_{ij} y_{j,t-1}$	$\overline{\gamma}_2$	-0.3697^{***}	-0.4568^{***}
		(0.0400)	(0.0291)
$\sum_{j=1}^{N} w_{ij} y_{j,t}$	ρ	0.5147^{***}	0.5644^{***}
		(0.0349)	(0.0290)
$\ln s_{it}$	β_1	0.0104	-0.0084
		(0.0142)	(0.0131)
$\ln(n_{it} + \delta + g)$	β_2	0.0181	0.0428^{***}
		(0.0132)	(0.0144)
$\sum_{j=1}^{N} w_{ij} \ln s_{jt}$	θ_1	0.0146^{***}	0.0450^{**}
		(0.0196)	(0.0177)
$\sum_{j=1}^{N} w_{ij} \ln(n_{jt} + \delta + g)$	θ_2	-0.0279	-0.0110
		(0.0178)	(0.0192)
$\overline{\ln s_{it}}$	π_{μ_1}		-0.0316
			(0.0241)
$\overline{\ln(n_{jt} + \delta + g)}$	π_{μ_2}		0.0731^{**}
			(0.0335)
$\sum_{j=1}^{N} w_{ij} \overline{\ln s_{jt}}$	π_{α_1}		-0.0011
			(0.0359)
$\sum_{j=1}^{N} w_{ij} \overline{\ln(n_{jt} + \delta + g)}$	π_{α_2}		-0.0982^{**}
			(0.0485)

Variance Components

σ_μ	σ_α	$\sigma_{\mu\alpha}$	σ^2_ε
0.0005^{***}	0.0007^{**}	-0.0004^{*}	0.0018^{***}
(0.0001)	(0.0003)	(0.0002)	(0.0001)

Note: Estimates I and II were obtained using the methods proposed by Lee and Yu (2016) and Miranda et al. (2017a), respectively. The dependent variable is $log(GDP)$. We denote the time-mean of a variable with an upper bar. Time dummies included but not reported. * indicates statistically significant at the 10% level, ** at the 5% level and *** at the 1% level.

Table 3: **QML estimates (Restricted model).**

Variable	Parameter	Estimates III	Estimates IV
$y_{i,t-1}$	$\overline{\gamma}_1$	0.7371^{***}	0.8828^{***}
		(0.0208)	(0.0183)
$\sum_{j=1}^{N} w_{ij} y_{j,t-1}$	$\overline{\gamma}_2$	-0.3650^{***}	-0.4742^{***}
		(0.0400)	(0.0285)
$\sum_{j=1}^{N} w_{ij} y_{j,t}$	ρ	0.5154^{***}	0.5894^{***}
		(0.0349)	(0.0269)
$\ln\left(\frac{s_{it}}{n_{it}+\delta+g}\right)$	β^c	-0.0049	-0.0240^{***}
		(0.0100)	(0.0100)
$\sum_{j=1}^{N} w_{ij} \ln\left(\frac{s_{it}}{n_{it}+\delta+g}\right)$	θ^c	0.0208^{***}	0.0285^{**}
		(0.0144)	(0.0139)
$\overline{\ln\left(\frac{s_{it}}{n_{it}+\delta+g}\right)}$	π_{μ_1}		-0.0405^{*}
			(0.0218)
$\sum_{j=1}^{N} w_{ij} \overline{\ln\left(\frac{s_{it}}{n_{it}+\delta+g}\right)}$	π_{α_1}		0.0421
			(0.0300)

Variance Components

σ_μ	σ_α	$\sigma_{\mu\alpha}$	σ^2_ε
0.0005^{***}	0.0008^{**}	-0.0005^{**}	0.0018^{***}
(0.0002)	(0.0003)	(0.0002)	(0.0001)

Note: We denote with the upper letter c the constrained parameters. In particular, $\beta = \beta_1 = -\beta_2$ and $\theta = \theta_1 = -\theta_2$. Estimates III and IV were obtained using the methods proposed by Lee and Yu (2016) and Miranda et al. (2017a), respectively. The dependent variable is $log(GDP)$. We denote the time-mean of a variable with an upper bar. Time dummies included but not reported. * indicates statistically significant at the 10% level, ** at the 5% level and *** at the 1% level.

Table 4: **Implied Parameters**

Unrestricted model		
λ	γ_1	
0.0244^{***}	-1.5617^{***}	
(0.0040)	(0.0030)	

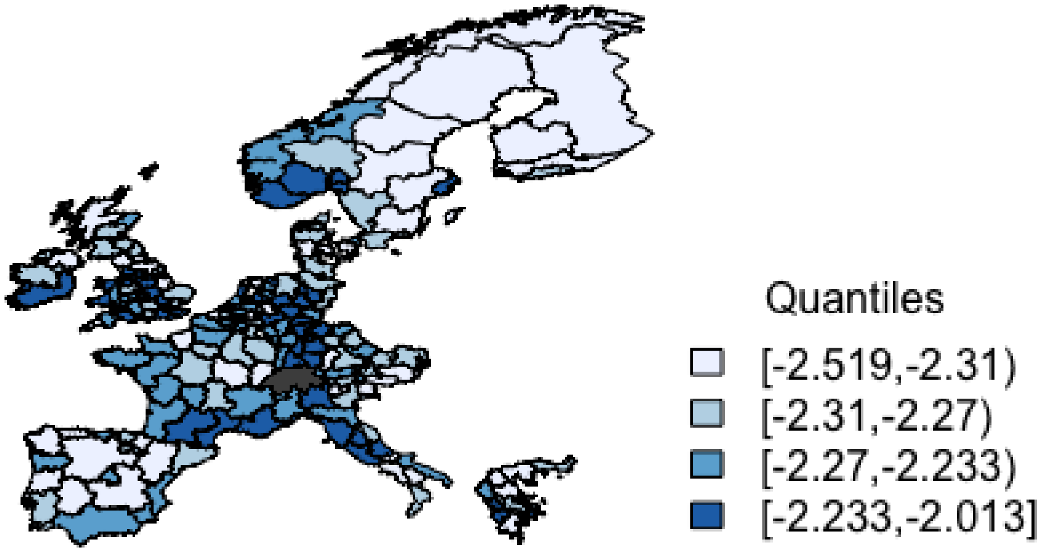
Restricted model		
λ	γ_1	γ_3
0.0244^{***}	-1.1997^{***}	0.4355^{***}
(0.0040)	(0.0004)	(0.1415)

Assumed technology	α	γ_2	ϕ
López-Bazo et al. (2004)	-0.2575^{*}	0.1937	
	(0.1445)	(0.1779)	
Ertur and Koch (2007)	-0.7024		0.4448
	(0.6223)		(0.5512)

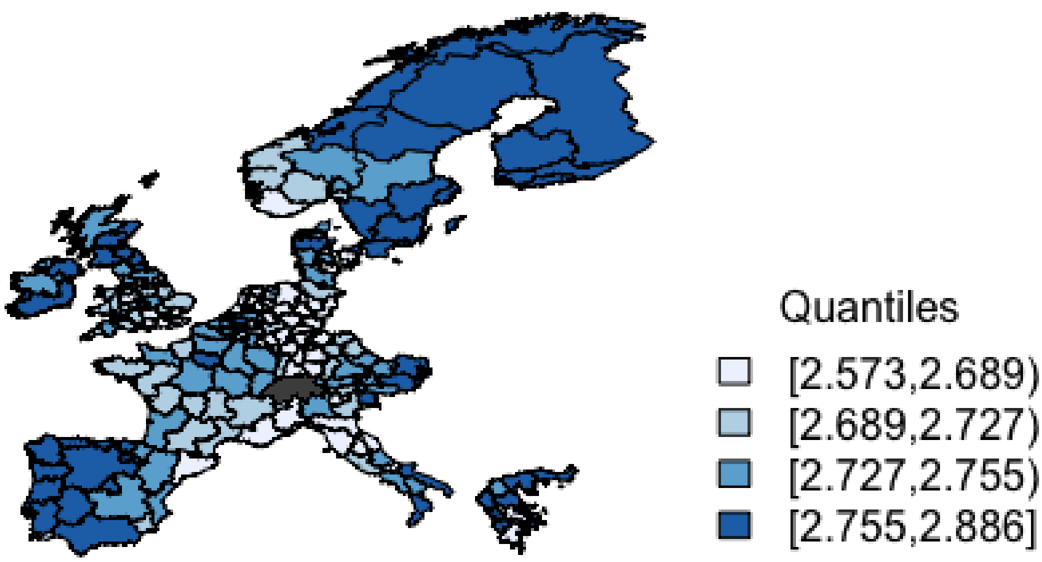
Note: Results for the "Unrestricted model" were obtained using the estimates reported in Table 2 ("Estimates II"), whereas results for the "Restricted model" were obtained using the estimates reported in Table 3 ("Estimates IV"). Except for γ_1 (OLS estimate from a linear regression without constant), standard errors were obtained using the delta method. * 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 $\hat{\gamma}_1 \sum_{j=1}^{N} w_{ij} \ln \hat{\Omega}_{j0}$.



A The balanced growth rate

From equation 2.6:

$$\ln y_{it} = \left[1 + \left(\frac{(\gamma_3+\gamma_1)(u_{ii}-\alpha-\phi)}{(\phi\gamma_3+\gamma_2)}\right)\right]\ln\Omega_{it} + \frac{(\gamma_3+\gamma_1)}{\phi\gamma_3+\gamma_2}\sum_{j\neq i}^{N} u_{ij}\ln\Omega_{jt} + u_{ii}\ln k_{it} + \sum_{j\neq i}^{N} u_{ij}\ln k_{jt}$$

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

$$\frac{d\ln y_{it}}{dt} = \left[1 + \left(\frac{(\gamma_3+\gamma_1)(u_{ii}-\alpha-\phi)}{(\phi\gamma_3+\gamma_2)}\right)\right]\mu + \frac{(\gamma_3+\gamma_1)}{\phi\gamma_3+\gamma_2}\sum_{j\neq i}^{N} u_{ij}\mu + u_{ii}g + \sum_{j\neq i}^{N} u_{ij}g$$

Also, using $u_{ii} + \sum_{j\neq i}^{N} u_{ij} = \sum_{j=1}^{N} u_{ij} = \alpha + \frac{\phi+\gamma_2}{1-\gamma_3}$,

$$\frac{d\ln y_{it}}{dt} = \left(1 - \frac{(\gamma_3+\gamma_1)(\alpha+\phi)}{(\phi\gamma_3+\gamma_2)} + \frac{(\gamma_3+\gamma_1)}{(\phi\gamma_3+\gamma_2)}\left(\frac{\alpha(1-\gamma_3)+\phi+\gamma_2}{1-\gamma_3}\right)\right)\mu + \sum_{j=1}^{N} u_{ij}g = g,$$

which after some algebra becomes:

$$\left(\frac{1+\gamma_1}{1-\gamma_3}\right)\mu + \left(\frac{\alpha(1-\gamma_3)+\phi+\gamma_2}{1-\gamma_3}\right)g = g$$

Therefore,

$$g = \frac{\mu\,(1+\gamma_1)}{(1-\gamma_3)(1-\alpha)-\phi-\gamma_2}$$

B Taylor approximation to the marginal productivity of capital

The Taylor approximation of $\frac{\dot{k}_{it}}{k_{it}}$ around the steady state $(k^*_{1t}, \cdots, k^*_{Nt})$ is

$$\begin{aligned}\frac{\dot{k}_{it}}{k_{it}} &= \frac{\dot{k}^*_{it}}{k^*_{it}} + \sum_{j=1}^{N}\left\{\left.\frac{\partial\left(\dot{k}_{it}/k_{it}\right)}{\partial \ln k_{jt}}\right|_{k^*_{jt}} (\ln k_{jt} - \ln k^*_{jt})\right\} \\ &= g + \left.\frac{\partial\left(\dot{k}_{it}/k_{it}\right)}{\partial \ln k_{it}}\right|_{k^*_{it}} (\ln k_{it} - \ln k^*_{it}) + \sum_{j\neq i}^{N}\left\{\left.\frac{\partial\left(\dot{k}_{it}/k_{it}\right)}{\partial \ln k_{jt}}\right|_{k^*_{jt}} (\ln k_{jt} - \ln k^*_{jt})\right\}\end{aligned}$$

Next we calculate the two derivatives involved. First, let us rewrite the marginal productivity of capital (see footnote 6) as

$$\frac{\dot{k}_{it}}{k_{it}} = s_i \Omega_{it}^{c_{ii}} \prod_{j\neq i}^{N} \Omega_{jt}^{c_{ij}} e^{(u_{ii}-1)\ln k_{it}} \prod_{j\neq i}^{N} e^{u_{ij}\ln k_{jt}} - (n_i + \delta)$$

with $k_{it}^{u_{ii}-1} = e^{(u_{ii}-1)\ln k_{it}}$, $c_{ii} = 1 + \left(\frac{(\gamma_3+\gamma_1)(u_{ii}-\alpha-\phi)}{(\phi\gamma_3+\gamma_2)}\right)$ and $c_{ij} = \frac{(\gamma_3+\gamma_1)u_{ij}}{\phi\gamma_3+\gamma_2}$. Thus,

$$\left.\frac{\partial\left(\dot{k}_{it}/k_{it}\right)}{\partial \ln k_{it}}\right|_{k^*_{it}} = s_i \Omega_{it}^{c_{ii}} \prod_{j\neq i}^{N} \Omega_{jt}^{c_{ij}} (u_{ii}-1) e^{(u_{ii}-1)\ln k^*_{it}} \prod_{j\neq i}^{N} e^{u_{ij}\ln k^*_{jt}}$$

Also, given that $s_i\left[\frac{y^*_{it}}{k^*_{it}}\right] - (n_i+\delta) - g = 0$, replacing y^*_{it} by 2.6 at the steady state we obtain

$$s_i \Omega_{it}^{c_{ii}} \prod_{j\neq i}^{N} \Omega_{jt}^{c_{ij}} \prod_{j\neq i}^{N} {k^*_{jt}}^{u_{ij}} = (n_i+\delta+g)\, {k^*_{it}}^{1-u_{ii}} \tag{B.1}$$

Consequently,

$$\left.\frac{\partial\left(\dot{k}_{it}/k_{it}\right)}{\partial \ln k_{it}}\right|_{k_{it}^*} = (u_{ii}-1)(n_i+\delta+g)$$

Lastly, bearing in mind that $\prod_{j\neq i}^{N} e^{u_{ij}\ln k_{jt}^*} = e^{\sum_{j\neq i}^{N} u_{ij}\ln k_{jt}^*}$,

$$\left.\frac{\partial\left(\dot{k}_{it}/k_{it}\right)}{\partial \ln k_{jt}}\right|_{k_{jt}^*} = s_i\Omega_{it}^{c_{ii}}\prod_{j\neq i}^{N}\Omega_{jt}^{c_{ij}} e^{u_{ij}\ln k_{jt}^*}u_{ij} = u_{ij}(n_i+\delta+g)$$

Therefore:

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

C Speed of convergence

Let us take the total derivative of 2.6:

$$\frac{d\ln y_{it}}{dt} = \left[1+\left(\frac{(\gamma_1+\gamma_2)(u_{ii}-\alpha-\phi)}{\phi\gamma_3+\gamma_2}\right)\right]\frac{d\ln\Omega_{it}}{dt} + \frac{(\gamma_3+\gamma_1)}{\phi\gamma_3+\gamma_2}\sum_{j\neq i}^{N} u_{ij}\frac{d\ln\Omega_{jt}}{dt}$$
$$+ u_{ii}\frac{d\ln k_{it}}{dt} + \sum_{j\neq i}^{N} u_{ij}\frac{d\ln k_{jt}}{dt}$$

Given that $\frac{d\ln\Omega_{it}}{dt} = \frac{d\ln\Omega_{jt}}{dt} = \mu$, we concentrate on the derivatives with respect to k. To this end, let us consider the final result of appendix B:

$$\begin{aligned}\frac{d\ln k_{it}}{dt} &= g + (u_{ii}-1)(n_i+\delta+g)\left(\ln k_{it} - \ln k_{it}^*\right) + \sum_{j\neq i}^{N} u_{ij}(n_i+\delta+g)\left(\ln k_{jt} - \ln k_{jt}^*\right) \\ &= g - (n_i+\delta+g)\left(\ln k_{it} - \ln k_{it}^*\right) + u_{ii}(n_i+\delta+g)\left(\ln k_{it} - \ln k_{it}^*\right) \\ &+ \sum_{j\neq i}^{N} u_{ij}(n_i+\delta+g)\left(\ln k_{jt} - \ln k_{jt}^*\right)\end{aligned}$$

Then, using equation 2.6

$$\ln y_{it} = \left[1 + \left(\frac{(\gamma_1+\gamma_2)(u_{ii}-\alpha-\phi)}{\phi\gamma_3+\gamma_2}\right)\right]\ln\Omega_{it} + \frac{(\gamma_3+\gamma_1)}{\phi\gamma_3+\gamma_2}\sum_{j\neq i}^{N} u_{ij}\ln\Omega_{jt} + u_{ii}\ln k_{it} + \sum_{j\neq i}^{N} u_{ij}\ln k_{jt}$$

and its value at the steady state

$$\ln y_{it}^* = \left[1 + \left(\frac{(\gamma_1+\gamma_2)(u_{ii}-\alpha-\phi)}{\phi\gamma_3+\gamma_2}\right)\right]\ln\Omega_{it} + \frac{(\gamma_3+\gamma_1)}{\phi\gamma_3+\gamma_2}\sum_{j\neq i}^{N} u_{ij}\ln\Omega_{jt} + u_{ii}\ln k_{it}^* + \sum_{j\neq i}^{N} u_{ij}\ln k_{jt}^*$$

we obtain

$$\ln y_{it} - \ln y_{it}^* = u_{ii}(\ln k_{it} - \ln k_{it}^*) + \sum_{j\neq i}^{N} u_{ij}(\ln k_{jt} - \ln k_{jt}^*) \tag{C.1}$$

Therefore,

$$\frac{d\ln k_{it}}{dt} = g - (n_i+\delta+g)\left(\ln k_{it} - \ln k_{it}^*\right) + (n_i+\delta+g)(\ln y_{it} - \ln y_{it}^*)$$

Plugging the previous result into the total derivative of 2.6:

$$
\begin{aligned}
\frac{d\ln y_{it}}{dt} &= \left[1+\left(\frac{(\gamma_1+\gamma_2)(u_{ii}-\alpha-\phi)}{\phi\gamma_3+\gamma_2}\right)\right]\mu+\frac{(\gamma_3+\gamma_1)}{\phi\gamma_3+\gamma_2}\sum_{j\neq i}^{N}u_{ij}\mu \\
&+u_{ii}\left(g-(n_i+\delta+g)\left(\ln k_i(t)-\ln k_i^*\right)+(n_i+\delta+g)(\ln y_i-\ln y_i^*)\right) \\
&+\sum_{j\neq i}^{N}u_{ij}\left(g-(n_j+\delta+g)\left(\ln k_j(t)-\ln k_j^*\right)+(n_j+\delta+g)(\ln y_j-\ln y_j^*)\right) \\
&=\left[1+\left(\frac{(\gamma_1+\gamma_2)(u_{ii}-\alpha-\phi)}{\phi\gamma_3+\gamma_2}\right)\right]\mu+\frac{(\gamma_3+\gamma_1)}{\phi\gamma_3+\gamma_2}\sum_{j\neq i}^{N}u_{ij}\mu+u_{ii}g+\sum_{j\neq i}^{N}u_{ij}g \\
&-\left(u_{ii}(n_i+\delta+g)\left(\ln k_{it}-\ln k_{it}^*\right)+\sum_{j\neq i}^{N}u_{ij}(n_j+\delta+g))\left(\ln k_{jt}-\ln k_{jt}^*\right)\right) \\
&+\left(u_{ii}(n_i+\delta+g)\left(\ln y_{it}-\ln y_{it}^*\right)+\sum_{j\neq i}^{N}u_{ij}(n_j+\delta+g))\left(\ln y_{jt}-\ln y_{jt}^*\right)\right)
\end{aligned}
$$

The first term in the previous expression corresponds to the balanced growth rate g (see appendix A). As for the second term, let us assume that, for each economy i, there exists Λ_i such that:

$$
\sum_{j=1}^{N}u_{ij}(n_j+g+\delta)(\ln k_{jt}-\ln k_{jt}^*)=\Lambda_i\left(u_{ii}(\ln k_{it}-\ln k_{it}^*)+\sum_{j\neq i}^{N}u_{ij}(\ln k_{jt}-\ln k_{jt}^*)\right)
$$

Thus,

$$
\begin{aligned}
\frac{d\ln y_{it}}{dt} &= g-\Lambda_i\left(u_{ii}(\ln k_{it}-\ln k_{it}^*)+\sum_{j\neq i}^{N}u_{ij}(\ln k_{jt}-\ln k_{jt}^*)\right) \\
&+u_{ii}(n_i+\delta+g)\left(\ln y_{it}-\ln y_{it}^*\right)+\sum_{j\neq i}^{N}u_{ij}(n_j+\delta+g)\left(\ln y_{jt}-\ln y_{jt}^*\right) \\
&=g-\Lambda_i(\ln y_{it}-\ln y_{it}^*)+u_{ii}(n_i+\delta+g)\left(\ln y_{it}-\ln y_{it}^*\right)+\sum_{j\neq i}^{N}u_{ij}(n_j+\delta+g)\left(\ln y_{jt}-\ln y_{jt}^*\right)
\end{aligned}
$$

where the second expression is obtained by using $C.1$.

Finally, from the first hypothesis in 2.11 we have that $(\ln y_{it}-\ln y_{it}^*)\,\Theta_j^{-1}=\ln y_{jt}-\ln y_{jt}^*$.

This allows us to obtain the speed of convergence to the steady state:

$$\begin{aligned}\frac{d\ln y_{it}}{dt} &= g - \left(\Lambda_i - u_{ii}(n_i+\delta+g) - \sum_{j\neq i}^{N} u_{ij}(n_j+\delta+g)\Theta_j^{-1}\right)(\ln y_{it} - \ln y_{it}^*) \\ &= g - \lambda_i\,(\ln y_{it} - \ln y_{it}^*)\end{aligned}$$

What is left is to derive the expressions defining Λ_i and λ_i. First, by plugging the second hypothesis in 2.11, $(\ln k_{it} - \ln k_{it}^*)\,\Phi_j^{-1} = \ln k_{jt} - \ln k_{jt}^*$, into our assumption on the existence of Λ_i:

$$\begin{aligned}\sum_{j=1}^{N} u_{ij}(n_j+g+\delta)(\ln k_{jt} - \ln k_{jt}^*) &= \Lambda_i\left(u_{ii}(\ln k_{it} - \ln k_{it}^*) + \sum_{j\neq i}^{N} u_{ij}(\ln k_{jt} - \ln k_{jt}^*)\right) \\ \sum_{j=1}^{N} u_{ij}(n_j+g+\delta)(\ln k_j - \ln k_j^*) &= \Lambda_i \sum_{j=1}^{N} u_{ij}(\ln k_j(t) - \ln k_j^*) \\ \sum_{j=1}^{N} u_{ij}(n_j+g+\delta)\,(\ln k_i(t) - \ln k_i^*)\,\Phi_j^{-1} &= \Lambda_i \sum_{j=1}^{N} u_{ij}\,(\ln k_i(t) - \ln k_i^*)\,\Phi_j^{-1} \\ \Lambda_i &= \frac{\sum_{j=1}^{N} u_{ij}\frac{1}{\Phi_j}(n_j+g+\delta)}{\sum_{j=1}^{N} u_{ij}\frac{1}{\Phi_j}}\end{aligned} \tag{C.2}$$

Second, plugging the previous result into $\lambda_i = \Lambda_i - u_{ii}(n_i+\delta+g) - \sum_{j\neq i}^{N} u_{ij}(n_j+\delta+g)\Theta_j^{-1}$ and assuming that $\Theta_i^{-1} = 1$:

$$\begin{aligned}\lambda_i &= \Lambda_i - u_{ii}(n_i+\delta+g)\Theta_i^{-1} - \sum_{j\neq i}^{N} u_{ij}(n_j+\delta+g)\Theta_j^{-1} \\ \lambda_i &= \Lambda_i - \sum_{j=1}^{N} u_{ij}(n_j+\delta+g)\Theta_j^{-1} \\ \lambda_i &= \frac{\sum_{j=1}^{N} u_{ij}\frac{1}{\Phi_j}(n_j+g+\delta)}{\sum_{j=1}^{N} u_{ij}\frac{1}{\Phi_j}} - \sum_{j=1}^{N} u_{ij}(n_j+\delta+g)\frac{1}{\Theta_j}\end{aligned}$$

D Differential equation solution

We start by noticing that the steady state in 2.9 can be written as

$$\ln y^*_{it} = \frac{1}{1-\alpha-\phi}\sum_{j=1}^{N}\sum_{r=0}^{\infty}\rho^r w^{(r)}_{ij}\ln\Omega_{jt} + \frac{\gamma_1}{1-\alpha-\phi}\sum_{j=1}^{N}\sum_{r=0}^{\infty}\rho^r w^{(r+1)}_{ij}\ln\Omega_{jt}$$
$$+\left(\frac{\alpha+\phi}{1-\alpha-\phi}\right)\sum_{j=1}^{N}\sum_{r=0}^{\infty}\rho^r w^{(r)}_{ij}\ln\left(\frac{s_j}{n_j+\delta+g}\right) + \frac{\gamma_2-\alpha\gamma_3}{1-\alpha-\phi}\sum_{j=1}^{N}\sum_{r=0}^{\infty}\rho^r w^{(r+1)}_{ij}\ln\left(\frac{s_j}{n_j+\delta+g}\right)$$

with $\rho = \frac{\gamma_2-\alpha\gamma_3}{1-\alpha-\phi}$. Using this, we can see that $\frac{d\ln y^*_{it}}{dt} = \frac{(1+\gamma_1)\mu}{1-\alpha-\phi}\left(\frac{1}{1-\rho}\right)$, since $\frac{1}{1-\rho} = \frac{1-\alpha-\phi}{(1-\alpha)(1-\gamma_3)-\phi-\gamma_2}$, then

$$\frac{d\ln y^*_{it}}{dt} = \left(\frac{1}{1-\alpha-\phi}\right)\left(\frac{1}{1-\rho}\right)\mu + \left(\frac{\gamma_1}{1-\alpha-\phi}\right)\left(\frac{1}{1-\rho}\right)\mu = g \tag{D.1}$$

Notice that $D.1$ can be seen as another differential equation, which have a particular solution on $\ln y^*_{i0}$:

$$\ln y^*_{it} = gt + \ln y^*_{i0} \tag{D.2}$$

Plugging equation $D.2$ and 2.12 we obtain:

$$\frac{d\ln y_{it}}{dt} = g - \lambda_i\left(\ln y_{it} - gt - \ln y^*_{i0}\right) \tag{D.3}$$

We use the integrating factor method to solve the differential equation in $D.3$. We first reorder terms and then multiply the equation by the integrating factor $e^{\int \lambda_i dt} = e^{\lambda_i t}$ to obtain

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

By integrating on both sides, we obtain the general solution:

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

The particular solution for $t = t_1$ implies that $C = \left(\ln y_{it_1} - gt_1 - \ln y_{i0}^*\right) e^{\lambda_i t_1}$. Thus, for any t we have:

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

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.

Andreano, M. S., Benedetti, R., and Postiglione, P. (2017). Spatial regimes in regional european growth: an iterated spatially weighted regression approach. *Quality and Quantity: International Journal of Methodology*, 51(6):2665–2684.

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.

De Long, J. B. and Summers, L. (1991). Equipment investment and economic growth. *The Quarterly Journal of Economics*, 106(2):445–502.

Debarsy, N. (2012). The mundlak approach in the spatial Durbin panel data model. *Spatial Economic Analysis*, 7(1):109–131.

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 spacetime 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.

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.

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

Pfaffermayr, M. (2012). Spatial Convergence Of Regions Revisited: A Spatial Maximum Likelihood Panel Approach. *Journal of Regional Science*, 52(5):857–873.

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