# The distance decay effect and spatial reach of spillovers[1]

J.Paul Elhorst, Ioanna Tziolas, Chang Tan, Petros Milionis[2]

This paper presents a methodology to estimate the distance decay effect and spatial reach of spillover effects in the spatial Durbin (SD) model. Building on attributes of the concept of spatial autocorrelation developed by Arthur Getis, we adopt a distance-based negative exponential spatial weight matrix and parameterize it by a distance decay parameter that is estimated separately for each spatial lag. The methodology is illustrated based on the spatially augmented neoclassical growth framework, which we estimate using data for 266 NUTS-2 regions in the EU over the period 2000-2018. We find distance decay parameters ranging from 0.233 to 2.224 for the different spatial lags of the growth determinants in this model. This range highlights the restrictiveness of the SD model based on one common spatial weight matrix for all spatial lags. We also quantify and illustrate graphically the extent to which a change in each growth determinant spills over to other regions in terms of distance, slope, magnitude and significance level.

**JEL classification:** C21, C23, O47, R12

**Keywords:** Regional economic growth, growth spillovers, regional proximity, distance decay

**Paper prepared for the ERSA conference, August 28 – September 1, 2023, Alicante, Spain**

[1] This paper is inspired by the insightful work produced by Arthur Getis (1934-2022), one of the founding fathers of spatial econometrics, and is meant to serve as a tribute to his legacy in this field. Data used and programming code developed in this paper will be made available.

[2] Department of Economics, Econometrics & Finance, Faculty of Economics & Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. E-mail addresses: j.p.elhorst@rug.nl, i.tziolas@rug.nl, c.tan@rug.nl, p.milionis@rug.nl.

## 1. Introduction

As the world economy becomes increasingly integrated, there is also growing evidence that economic growth is highly correlated across space. This pattern is clearly visible in the data and has been documented in an extensive body of empirical studies (Moreno and Trehan 1997; Lopez-Bazo et al. 2004; Ertur and Koch, 2007; Ramajo et al. 2008). Although this fact is increasingly recognized by economists and policymakers, there is still no consensus in the literature on the magnitude of observed growth spillovers and the range over which they operate. This lack of consensus has been highlighted in a recent article by Rosenthal and Strange (2020) whose title raises the pressing question: "How close is close". Their answer draws on a range of research on agglomeration effects in economics and regional science, yet without providing a clear research methodology on how to estimate spillover effects.

To address this question, we propose a novel approach for the estimation of growth spillovers within the spatially augmented neoclassical growth framework that allows us to capture how fast they decay in space. Our approach draws on the work of Arthur Getis regarding the concept of spatial autocorrelation, which we "translate" into present-day spatial econometrics. Our main novelty is that we parameterize the spatial weight matrix with a parameter that captures the rate at which interactions between economies decay in terms of distance and estimate that parameter separately for each growth determinant. This way we let the data determine the slope, magnitude and significance level of economic growth spillovers as a function of distance.

We illustrate the power of this approach by using it to estimate spillover effects in GDP per capita growth across a sample of EU NUTS-2 regions over the period from 2000 to 2018. Our findings confirm not only the existence and importance of growth spillovers, but also their decrease in magnitude as distance increases, consistent with Tobler's first law of geography. Moreover, they reveal that the distance decay effect and spatial reach of each spillover effect vary significantly across different growth determinants. These findings extend prior work in the literature that has estimated spatially augmented versions of the neoclassical growth model (Lopez-Bazo et al. 2004; Ertur and Koch 2007, 2011; Elhorst et al. 2010). It also suggests that growth-related externalities are subject to potential geographic barriers in their diffusion, as some of the literature has pointed out (Döring and Schnellenbach 2006).

There is extensive work in the literature that has tried to estimate the magnitude of growth spillovers. Early work on spillovers used regional dummies (Easterly and Levine 1997) or control variables that are averaged across nearby countries (Ades and Chua 1995). Moreno and Trehan (1997) are among the first to test empirically the specific channels through which growth spillovers operate using spatial econometric techniques. They conclude that there is evidence of spillovers across countries operating both as a result of common shocks affecting economic activity and as a result of particular growth determinants. More recent work has used various approaches to measure spillovers at the sub-national level and to analyze how wide the distribution is in space (Jaffe et al. 1993; Bottazzi and Peri 2003; Funke and Niebuhr 2005; Rodriguez-Pose and Crescenzi 2008). There is also an extensive body of literature that looked at spillovers between urban regions (Glaeser et al. 1992; Henderson et al. 1995). While this literature has provided empirical evidence regarding the existence of growth spillovers, the results have been inconclusive regarding the magnitude of these spillovers (Funke and Niebuhr 2005; Ramajo et al. 2008; Benos et al. 2015; Marquez et al. 2015). One reason is that the literature has either used indirect ways to account for growth spillovers (Enflo and Hjertstrand 2009; Gardiner et al. 2011), or has attempted to estimate growth spillovers directly, using econometric specifications and spatial weight matrices which limit the decay and the spatial range of the spillovers.

Our analysis avoids the above issues by linking the magnitude of spillovers with distance in a flexible way. We let the rate of decay of growth spillovers across space be determined by the data and we permit that rate to be different across growth determinants. Utilizing the structure of the spatial Durbin (SD) model (LeSage and Pace 2009), we compute the direct and spillover effects of each growth determinant on growth rates given the estimated rates of distance decay.

The setup of this paper is as follows. In Section 2 we link our approach to attributes of the concept of spatial autocorrelation developed by Arthur Getis. In Section 3 we present the spatially augmented neoclassical model of economic growth that we use for our analysis and explain how it takes the form of the SD model. In Section 4 we introduce the parameterizations of the spatial weight matrices and show their relationship with the direct and spillover effects of the growth determinants in the SD model. In Section 5 we describe the data, report and discuss the basic results, graph the spillover effects of the economic growth determinants and examine four potential model changes. Finally, Section 6 concludes.

## 2. Arthur Getis: The concept of spatial autocorrelation

In a survey article to the Handbook of Applied Spatial Analysis (Fischer and Getis 2010), Arthur Getis summarized his contributions to the field of spatial econometrics by eleven attributes of the concept of spatial autocorrelation (Getis 2010). In this section we review these attributes and translate them into present-day spatial econometrics. According to Getis (2010, pp.257-259), '*the list should convince all of those who deal with georeferenced data that an explicit recognition of the concept is basic to any spatial analysis*':

1. *Proper specification* [to avoid misspecification] *requires that any spatial association is subsumed with the model proper.*
2. *A thorough understanding of the effects of regressor variables on a dependent variable requires that any spatial effects in both dependent and independent variables be quantified.*
3. *Spatial autocorrelation statistics are usually designed to test the null hypothesis that there is no relationship among realizations of a single variable, but the tests may be extended to consider spatial relations between variables.*
4. *Measures of spatial autocorrelation will change in certain known ways when the configuration of spatial units changes.*
5. *A focus on a single spatial unit's effect on other units and vice versa.*

6. *Measures of spatial association can identify the parameters of distance decay (for example, the parameters of a negative exponential model).*
7. *A series of measures of spatial autocorrelation over time sheds light on temporal effects.*
8. *If the goal is to avoid, as much as possible, spatial autocorrelation in the sample, then a reasonable sample design would benefit from a study of spatial autocorrelation in the region where the sample is to be selected.*
9. *Before engaging in many types of spatial analysis, it is necessary to make the assumption that spatial stationarity exists.*
10. *A means of identifying spatial clusters.*
11. *A means of identifying outliers, both spatial and non-spatial.*

Translated into present-day spatial econometrics, this is a plea for the SD model in which the spatial weight matrices take an exponential form and negatively depend on a distance decay parameter, which can differ for each spatial lag in the model.

The SD model, which received much attention in applied spatial econometric studies thanks to the work of LeSage and Pace (2009), covers the first two attributes. According to these authors, the cost of ignoring spatial lags in the regressand and the regressor variables, when relevant, is high since the coefficients of the remaining variables may then be biased. By contrast, ignoring a spatial lag in the error term, if relevant, will only result in a loss of efficiency (see also LeGallo 2014).

Regarding the third attribute, several spatial autocorrelation test statistics have been proposed and used in the applied literature to motivate the use of spatial econometric models. A common test statistic is Moran's I applied to the regressand in raw form. However, the null hypothesis that the regressand is not spatially correlated generally needs to be rejected when using this statistic. This is because Moran's I does not control for potential spatial lags in the regressor variables. Theoretically, it is possible that a standard linear regression without any spatial lags is sufficient since the regressor variables may also be spatially correlated in such a way that they fully cover the spatial correlation in the regressand. In this regard, Anselin and Rey (2014) label Moran's I as a "*non-constructive test in that the alternative is diffuse, and not a specific (focused) model*'' (p. 107).

Another commonly used approach to motivate the use of spatial econometric models is to apply the robust Lagrange multiplier tests, developed by Anselin et al. (1996). These tests analyze whether the linear regression model estimated by OLS should be extended to include a spatial lag in the regressand or the error term, known as respectively the spatial autoregressive (SAR) model and the spatial error (SE) model. However, these tests also do not control for potential spatial lags in the regressor variables. When estimating the SD model, which includes spatial lags in the regressor variables, this potential misspecification can be avoided. Furthermore, since the OLS, SAR and SE models are special cases of the SD model (LeSage and Pace 2009), it can also be tested using Wald or likelihood ratio (LR) ratio tests whether the SD model simplifies to one of these models (Elhorst 2014; Juhl 2021). The residuals of the SD model (or one of these simpler models) can then be used to test for any remaining spatial dependence. This is one of the best ways to deal with the third attribute of spatial autocorrelation. A recent test that can be used for this purpose is the cross-sectional dependence (CD) test of Pesaran (2015), because it does not require any pre-specified spatial weight matrix, which also fulfills the fourth attribute. If the CD test applied to the residuals of the SD model still points to any remaining spatial dependence, only then further adjustments may be necessary to find a proper model.

Another advantage of the SD model in empirical research over other spatial econometric models (SAR, SE and SARAR[3]) is its flexibility in modeling spillovers, and thus the fifth attribute. The main interest of many empirical researchers is not the parameter estimates of the regressor variables, but the marginal impact of changes they have on the regressand. Two marginal effects stand out: the direct effect of changing the regressor of one unit on the regressand of that unit itself, and the cumulative effect of changing the regressor of one unit on the regressand of all other units (LeSage and Pace 2009). This cumulative effect is also known as the indirect effect or, alternatively, the spillover effect, the term we employ throughout this paper. Halleck Vega and Elhorst (2015) demonstrate that only models that at least include spatial lags of the regressor variables are able to produce spillover effects that can take any empirical value relative to the direct effects. By contrast, the popular SAR, SE and SARAR models are less flexible since they impose restrictions on the magnitude of spillover effects in advance. In the SE model they are zero by construction and in the SAR and SARAR models they are the same for every regressor.

Up to now, the sixth attribute of measuring distance decay received relatively little attention in the spatial econometric literature. Most studies adopt one common spatial weight matrix for all spatial lags in the SD model. By parameterizing the distance-based negative exponential spatial weight matrix by a distance decay parameter that differs for each spatial lag, we also try to give shape to this particular attribute of Artus Getis. The present study illustrates the benefits of this approach in the context of a spatially augmented neoclassical growth framework. For this purpose, we use data of 266 NUTS-2 regions in the EU over the period 2000-2018.

Using data over a period that covers the financial crisis of 2008-2009 and the resulting Great Recession, followed by the European debt crisis of 2009-2015, we also cover the seventh attribute, as growth rates were relatively high before this recession and relatively low in the period immediately after it. Figure 1 displays the average growth rate of GDP per capita across all regions, which dropped precipitously in 2009 and then recovered gradually. Furthermore, by using data at the sub-national level, which will be characterized by a substantial level of spatial autocorrelation, we also can test whether the proposed SD model is able to cover the eighth attribute by applying the CD test on its residuals.

<< Figure 1 around here >>

[3] The abbreviation SARAR stands for the spatial autoregressive (SAR) model with spatial autoregressive (AR) error terms and thus combines the SAR and the SE models.

To test whether the ninth attribute of spatial stationarity is satisfied, we will specify in the next section which restriction on the parameters needs to be verified in the SD model. The last two attributes, the identification of spatial clusters and outlier observations, recur in our empirical analysis in Sections 5.3 and 5.4.

### 3. The spatially augmented neoclassical growth framework

The world's evolving distribution of incomes lies at the heart of the economic growth literature. Within this literature, the neoclassical growth framework is the most commonly used framework to understand the pattern of economic growth and the evolution of per capita incomes across countries and regions. The framework originates from theoretical contributions by Solow (1956) and Swan (1956) associating economic growth with the process of capital accumulation under diminishing returns. Following the standard empirical implementation of the neoclassical framework in a panel data context due to Islam (1995) leads to the following expression:

$$\Delta \ln y_{i,t} = \beta_1 \ln inv_{i,t} + \beta_2 \ln(n_{i,t} + g + \delta) + \beta_3 \ln y_{i,t-1} + \mu_i + \xi_t + \varepsilon_{i,t}, \qquad (1)$$

where $\ln y_{i,t}$ denotes the natural logarithm of GDP per capita of economy $i\ (= 1, \ldots, N)$ in period $t\ (= 1, \ldots, T)$ and $\Delta \ln y_{i,t} = \ln y_{i,t} - \ln y_{i,t-1}$ its growth rate.[4] $inv_{i,t}$ denotes the investment rate whose impact is measured by the parameter $\beta_1$. $n_{i,t}$ denotes the rate of population growth, $g$ the rate of technological progress and $\delta$ the depreciation rate.[5] The combined effect of these three variables is measured by parameter $\beta_2$. $\ln y_{i,t-1}$ is the natural logarithm of the initial level of GDP per capita at the beginning of each time period whose effect is captured by $\beta_3$. The specification also includes cross-sectional fixed effects, $\mu_i$, which reflect all time-invariant factors that lead to differences in growth rates across economies, such as geographic and institutional factors. Since growth rates are also affected by common trends, time period fixed effects, $\xi_t$, are also controlled for.[6] Finally, $\varepsilon_{i,t}$ represents the independently and identically distributed error term for all $i$ with zero mean and variance $\sigma^2$.

One important limitation of the standard neoclassical growth framework is the assumption that each economy operates in isolation of others. This assumption seems implausible especially when this framework is applied to sub-national economies between which production factors are highly mobile and technology can be easily transferred (Beugelsdijk et al. 2018). Over the past two decades, awareness of this issue has increased, leading to increased interest in the influence of an economy's spatial location on its growth rate.

[4] The presentation of this equation is based on annual data. The index *t-1* can be replaced by *t-p* if GDP per capita growth is measured over *p* years. In that case the growth rate should correspond to an average over this time period.

[5] In line with the common assumptions of the neoclassical growth framework, the rates of technological progress and depreciation, $g$ and $\delta$, are not indexed as they are assumed to be common for all economies and time periods (see also Islam 1995). In Section 5.4 we investigate what happens if we extend our specification to incorporate endogenous elements.

[6] See Section 5.4 for an alternative approach.

A prominent example is the study of Ertur and Koch (2007). In their neoclassical spatially-augmented economic-theoretical growth model, they allow for productivity spillovers between economies due to capital investment. Their model builds on and is supported by a large body of other studies highlighting the importance of technological and knowledge spillovers (e.g., Grossman and Helpman 1991; Audretsch and Feldman 2004; Autant-Bernard and LeSage 2011). Ertur and Koch (2007) also demonstrate that the empirical counterpart of their spatially augmented version of the neoclassical growth model takes the form of an SD model. This empirical model has been applied and extended in several follow-up studies, including LeSage and Fischer (2008), Elhorst et al. (2010), Yu and Lee (2012), Pfaffermayr (2012), and Lee and Yu (2016). These studies have shown that for a panel of $N$ cross-sectional observations over $T$ time periods this model in vector form reads as

$$\Delta Y_t = \rho W(\delta_0)\Delta Y_t + \tau \Delta Y_{t-1} + \eta W(\delta_1)\Delta Y_{t-1} + \lambda Y_{t-1} + \pi W(\delta_2)Y_{t-1} + \\ [X_{1t}, \dots, X_{Kt}]\beta + [W(\delta_1)X_{1t}, \dots, W(\delta_K)X_{Kt}]\theta + \mu + \xi_t \iota_N + \varepsilon_t, \quad (2)$$

where $\Delta Y_t = (\Delta \ln y_{1t}, \dots, \Delta \ln y_{Nt})^T$ denotes an $N \times 1$ vector of the regressand introduced in Equation (1). $W(\delta_0)Y_t$ represents the contemporaneous vector of the regressand $Y_t$ observed in neighboring economies and $\rho$ the spatial autoregressive response parameter of this vector. $\Delta Y_{t-1}$ and $W(\delta_1)\Delta Y_{t-1}$ denote the corresponding vectors of one-period time lags of these two variables, and $Y_{t-1}$ and $W(\delta_2)Y_{t-1}$ the initial levels of GDP per capita at the start of each time period. $[X_{1t}, \dots, X_{Kt}]$ is an $N \times K$ matrix of the regressor variables introduced in Equation (1) and $[W(\delta_1)X_{1t}, \dots, W(\delta_K)X_{Kt}]$ an $N \times K$ matrix of spatial lags of these regressor variables. The impacts of these regressor variables and their spatial lags are measured by the $K \times 1$ vectors $\beta$ and $\theta$, respectively. As explained above, $\mu = (\mu_1, \dots, \mu_N)^T$ and the set $\xi_t$ $(t = 1, \dots, T)$, where $\iota_N$ is an $N \times 1$ vector of ones, denote cross-sectional and time fixed effects respectively. The spatial weight matrix, symbolized by $W$, is an $N \times N$ matrix describing the spatial arrangement between each pair of economies $i$ and $j$, whose elements $w_{ij}$ in this paper are assumed to depend on a distance decay parameter ($\delta_k$). Its functional form is the topic of the next section.

Overall, Equation (2) shows that the GDP per capita growth rate of a given economy depends on the investment rate and the rates of population growth, technological progress and depreciation, both in the given economy and that of its neighbors, which determine the long-run equilibrium or steady state level of GDP per capita. It further depends on its lagged growth rate, as well as the contemporaneous and lagged growth rates of its neighhbors. Finally, it depends on the initial GDP per capita level in both the given and neighboring economies at the start of each time period, which reflects how far each economy is from its long-run equilibrium.

To find out under which parameter condition the spatially augmented version of the neoclassical growth framework leads to convergence or divergence, we rearrange and express Equation (2) in terms of GDP per capita levels, to get:

$$Y_t = \rho W(\delta_0)Y_t + (1+\tau+\lambda)Y_{t-1} + \big(-\rho W(\delta_0) + \eta W(\delta_1) + \pi W(\delta_2)\big)Y_{t-1} - \tau Y_{t-2}$$
$$-\eta W(\delta_1)Y_{t-2} + [X_{1t}, \dots, X_{Kt}]\beta + [W(\delta_1)X_{1t}, \dots, W(\delta_K)X_{Kt}]\theta + \mu + \xi_t\iota_N + \varepsilon_t, \quad (3)$$

Assuming row-normalized spatial weight matrices, Yu et al. (2012) show that the sum of the coefficients of the first five terms on the right-hand of this equation determines spatial stationarity, i.e., converge or divergence. This yields:

$$\rho + (1+\tau+\lambda) + (-\rho+\eta+\pi) - \tau - \eta = 1 + \lambda + \pi. \quad (4)$$

Convergence occurs if the latter sum is smaller than 1, and thus if the coefficients of the initial levels of GDP per capita in the given and neighboring economies are smaller than 0, i.e., $\lambda + \pi < 0$. In contrast divergence occurs if the sum is greater than 1, i.e., $\lambda + \pi > 0$. A special case of neither convergence or divergence occurs when $\lambda + \pi = 0$, which Yu et al. (2012) label as spatial co-integration. This corresponds to a situation in which GDP per capita growth rates in different economies fluctuate over the business cycle to a varying extent, but eventually remain on the same path during the entire sample period.

## 4. Parameterization and estimation

Following Arthur Getis' sixth attribute, a negative exponential functional form is used to specify $W(\delta_k)$. Its diagonal elements are set to zero to prevent economies from influencing themselves and its off-diagonal elements are specified by $w_{ij}(\delta_k) = \exp(-\delta_k d_{ij})$, where $d_{ij}$ denotes the geographic distance between each pair of economies $i$ and $j$. Although this functional form is commonly used, the novelty of our study is that the distance decay parameter ($\delta_k > 0$) is estimated rather than pre-specified and is allowed to be different for each spatial lag $k$ ($k = 0,1,2,3,4$). Here $k = 0$ refers to the distance decay parameter of the spatial lag in the regressand, $k = 1$ and $k = 2$ to the distance decay parameters of the time-lagged GDP per capita growth rate and the initial level of GDP per capita, and $k = 3$ and $k = 4$ to the distance decay parameters of the investment rate and the rates of population growth, technological progress and depreciation. After row-normalizing the elements $w_{ij}(\delta_k)$ of the negative exponential distance decay matrix, we obtain

$$w_{ij}(\delta_k) = \frac{\exp(-\delta_k d_{ij})}{\sum_{j=1}^{N}\exp(-\delta_k d_{ij})}. \quad (5)$$

It should be noted that previous studies generally adopt one common $W$ matrix for each spatial lag. This may be rather restrictive and lead to incorrect inferences regarding the degree of spatial interactions as it may be different for each variable. We examine whether this restriction is supported by the data in Section 5.2.

To draw conclusions regarding the marginal effects of the regressor variables and their spatial lags, the direct and spillover effects have to be invoked, as the parameters $\beta$ and $\theta$ alone provide an incomplete picture of the marginal effects in the SD model (LeSage and Pace 2009; Elhorst 2014). The direct effect ($DE_k$) measures the average impact of a change in the $k$th regressor of a given economy on its own growth rate, while the spillover effect ($SE_k$) measures the cumulative effect of changing this regressor on the growth rates of all its neighbors. The formulas for the direct and spillover effects of each standard regressor ($k = 3, 4$) read as:

$$DE_k = \frac{1}{N} tr \left\{ \left(I_N - \rho W(\delta_0)\right)^{-1} \left(\beta_k I_N + \theta_k W(\delta_k)\right) \right\}, \tag{6a}$$

$$SE_k = \frac{1}{N} \iota_N' \left\{ \left(I_N - \rho W(\delta_0)\right)^{-1} \left(\beta_k I_N + \theta_k W(\delta_k)\right) \right\} \iota_N - \frac{1}{N} tr \left\{ \left(I_N - \rho W(\delta_0)\right)^{-1} \left(\beta_k I_N + \theta_k W(\delta_k)\right) \right\}. \tag{6b}$$

Similarly, if $k = 1$ (GDP per capita growth rate), $\beta_k$ and $\theta_k$ need to be replaced by $\tau$ and $\eta$ and if $k = 2$ (the initial level of GDP per capita) by $\lambda$ and $\pi$. Halleck Vega and Elhorst (2015) demonstrate that only models that at least include spatial lags of the regressor variables ($\theta_k, \eta, \pi \neq 0$), such as the SD model, are able to produce spillover effects that can take any empirical value. By contrast, in the SE model they are zero by construction and in the SAR and SARAR models they are the same for every regressor. Parameterizing the spatial weight matrix of every regressor in the SD model enhances this flexibility by the decay parameters $\delta_k$ and $\delta_0$. This is because the slope of the distance decay effect and the distance at which they may still have effect on other units, i.e., the spatial reach, may also be different from one regressor to another. This will be further illustrated in Section 5.3.

When the distance parameters and therefore the spatial weight matrices are different, the existence of convergence, divergence or spatial cointegration can still be tested by considering $\lambda + \pi$. This is because the sum of the direct and spillover effect of the initial level of GDP per capita is equal to $\frac{1}{N} \iota_N' \left\{ \left(I_N - \rho W(\delta_0)\right)^{-1} \left(\lambda I_N + \pi W(\delta_k)\right) \right\} \iota_N$, which is smaller than, greater than or equal to zero if $\lambda + \pi$ is smaller than, greater than or equal to zero. Alternatively, one may test whether the sum of the direct and spillover effect of the initial level of GDP is smaller than, greater than or equal to zero.

We use the nonlinear quasi-maximum likelihood (QML)[7] estimator developed by Tan (2023) to estimate the parameters of Equation (2) and the corresponding variance-covariance matrix. This estimator allows us to obtain the parameter estimates of the regressor variables and the distance decay parameters simultaneously from the data, instead of relying on an arbitrarily specified $W$ matrix. The delta method (Arbia et al. 2020) or bootstrapping (LeSage and Pace 2009; Elhorst 2014) can be used to calculate the t-statistics of the direct and spillover effects.

[7] The term quasi is used since the usual assumption of normality of the error terms is not required. If the error terms are also assumed to be normally distributed, the QML estimator simplifies to an ML estimator.

## 5. Empirical analysis

### 5.1 Data

We perform the estimation of our spatial-augmented neoclassical growth framework based on a sample of 266 EU NUTS-2 regions across 27 countries over the period 2000-2018 provided by Eurostat's regional database. Conducting the analysis with EU NUTS-2 regions has the advantage of working with harmonized data on GDP and other macroeconomic aggregates, such as investment spending, which are not available at a more disaggregate level. Based on these data we can compute real GDP per capita in constant prices and adjusted for PPP (y), the investment rate (inv) and the population growth rate for each NUTS-2 regions (popg).[8] In one of the alternative model specifications, we also use data provided by Eurostat on the educational attainment of the population (educ) and the share of employment in science and technology (sci_tech). All variables are expressed in natural logarithms.

### 5.2 Basic results

Table 1 reports the estimation results of our spatially augmented neoclassical growth model for four different specifications of the spatial weight matrix or matrices. The estimates in column [1] are based on one common spatial weight matrix for all spatial lags based on the six nearest neighbor principle. This column is representative of a wide range of previous empirical studies on economic growth. Although most adopt a binary contiguity matrix based on the principle of sharing a common border, one problem is that many EU regions are islands, which would become isolated if the contiguity principle were applied to them (Anselin and Rey 2014, pp.38-40). Since the number of neighbors for non-island regions appears to be 5.98 on average, we used a six nearest neighbor matrix in column [1] so that islands can also be included in the analysis.

<< Table 1 around here >>

The estimates in column [2] are based on one common row-normalized exponential distance decay matrix using a pre-specified value of $\delta = 0.01$. This value of 0.01 has been used in several other studies based on EU regions (Pfaffermayr 2012; Ezcurra and Rios 2020) and also turns out the best choice when carrying out a Bayesian comparison test for a series of values starting with 0.01 and step size 0.01 (LeSage 2015). Using the value of 0.01 implies that all distances between regions measured in kilometers are first divided by 100. The rows of this matrix are then normalized to 1.

The estimates in column [3] are based on one common row-normalized exponential distance decay matrix whose distance decay parameter is estimated rather than pre-specified, using the non-linear estimation techniques developed by Tan (2023). The obtained estimate appears to be 1.088 (t-value 14.89), which after scaling is close to the

[8] Just as in previous studies the growth rate of the population is incremented by 0.05 to account for the rate of technological progress and the depreciation rate of capital.

pre-specified value of 0.01 in column [2]. Just as in column [2], the distances measured in kilometers are first divided by 100. If not, the estimate of the distance decay parameter would change to 0.01088. In column [4] of Table 1, the distance decay parameters of each spatial lag in the model are estimated separately.

Overall, the estimation results show a plausible model structure. To illustrate this, we focus on the estimates reported in column [4], unless otherwise stated, and explain this choice below.

The coefficient of the lagged GDP per capita growth rate in a given region is found to be positive but small and statistically insignificant, indicating that recent growth rates are not persistent. By contrast, the coefficients of both the contemporaneous and lagged growth rates in neighboring regions are found to be much larger and significant. Comparing the magnitudes of the direct and spillover effects for the lagged growth rate also reveals a staggering difference. A 1% increase in the growth rate of a region in the previous year will only lead to a 0.030% (t-value 2.08) increase in the current growth rate, whereas the resulting increase would be 0.478% (t-value 7.29) if such an increase occurred in all the neighboring regions.

Looking at the estimates for the investment rate we see a similar picture. The effect of the investment rate in the region itself and its corresponding direct effect are found to be positive, but small and insignificant. This stands in contrast with the coefficient of the investment rate in neighboring regions and its corresponding spillover effect which are much larger and statistically significant; a 1% increase in the investment rate in neighboring regions is associated with an increase in the GDP per capita growth rate of 0.027% (t-value 2.19), while such an increase in the region itself is nine times smaller.

Turning to the coefficient estimates of the population growth rate in the own and in neighboring regions, as well as its direct and spillover effects, they all appear to be negative. The difference with the previous two determinants is that only the coefficient in the region itself and its direct effect are significant. The direct effect is −0.009 (t-value −7.15), which implies that if the population grows by 1%, for example from 1 million to 1.01 million due to an influx of migrants, GDP per capita growth slows down by almost 0.1%.

Finally, looking at coefficient estimates for the initial level of GDP per capita in a given region, we see a strong and significant negative effect on GDP per capita growth (−0.090, t-value −16.25), suggesting convergent dynamics. Yet one needs to be careful since the initial level of GDP per capita in neighboring regions has a strong and significant positive effect on GDP per capita growth (0.087, t-value 6.61). The same applies to the corresponding direct (−0.090, t-value −16.43) and spillover (0.080, t-value 2.43) effects, which have opposite signs and almost sum to zero. Since we cannot reject the hypothesis that this sum is different from zero, the evidence is rather in favor of spatial cointegration, a situation that is characterized by neither convergence nor divergence over the entire sample period. This finding could be driven by the observation period and the impact of the Great Recession in 2008-2009. Some regions were hit harder than others, while after this recession some but often other regions were able to recover faster.[9]

Comparing the values of the log-likelihood function values (LogL) and two R-squared measures[10] across the four columns, it appears that as we allow for more flexibility in the spatial weight matrix, this leads to a better fit of the data. When replacing the relatively sparse six nearest neighbor matrix in column [1] with a denser exponential distance matrix in column [2],[11] all these statistics increase substantially. When the distance decay parameter in column [3] is estimated subsequently, both statistics increase further, albeit the improvement is limited since 0.01 in this particular case was already a good guess of this distance decay parameter. Finally, when conducting an LR test on the LogL value in column [4] relative to column [3], we obtain a test statistic of 21.4 (p-value 0.00), which exceeds all relevant critical values at four degrees of freedom, representing the additional number of parameters to be estimated. This finding provides empirical evidence that the distance decay parameters associated with each spatial lag should be estimated rather than pre-specified and allowed to be different. Indeed, the distance decay parameters in column [4] appear to range from 0.233 for the population growth rate to 2.224 for the initial level of GDP per capita. Apparently, the slope of the distance decay effect and the spatial reach of the spillover effects are significantly different for the different growth determinants.

Lastly, when running Pesaran's CD-test statistic on the raw data we obtain 302.3, indicating that GDP per capita growth rates are strongly spatially autocorrelated. Yet, when applied to the residuals of the four models estimated in Table 1, the test statistics drop to values between −0.908 to −0.605, which is within the confidence interval of (−1.96,+1.96).[12] This indicates that the spatial association between the regressand and regressors in these models is properly specified.

### 5.3 Graphing spillover effects

Although the coefficient estimates and the direct effects may not seem to differ much across the four columns of Table 1 at first glance, a different picture emerges when we compare the spillover effects. To illustrate this in more detail, we decompose the spillover effects in 21 distance categories and graph the obtained results in Figure 2 (the distinguished distance categories are spelled out in the note to this figure). The following explanation using the lagged GDP per capita growth rate as an example is intended to better understand these graphs and this decomposition. According to column [4] of Table 1, the spillover effect of the lagged GDP per capita growth rate is 0.478. This summary measure represents the average cumulative effect of changing this regressor in a given region on the regressand of all

[9] See also Breinlich et al. (2014), as well as Billé et al. (2023) for a similar finding in Italian regions.

[10] The corrected R-squared measures the explanatory power of the model excluding the contribution of fixed effects.

[11] A matrix is sparse if it contains many zeros and dense when it contains many nonzero elements.

[12] The CD test statistic of Pesaran (2015) converges to a standard normal distributed if $N$ and $T$ go to infinity, which implies that its critical values are $\pm 1.96$ at the 5% significance level.

other 265 regions in the sample, whether near or far, and is determined by Equation (6b). The decomposition breaks down this summary measure based on the distance to these other regions. If the distance of region $i$ to another region $j$ is $d_{ij}$ kilometers, region $j$ is assigned to distance category $[d_a, d_b]$, such that $d_a < d_{ij} < d_b$. When applied to all regions, this results in the summary measure of 0.478 reported in Table 1 being equal to the surface area under the solid line of the lagged GDP per capita growth rate in Figure 2. In addition to the lagged GDP per capita growth rate, Figure 2 also graphs the decomposed spillover effects of the other growth determinants, as well as their 95% confidence intervals, based on the estimation results reported in column [4] of Table 1.

<< Figure 2 around here >>

The first thing worth emphasizing is that the confidence intervals in Figure 2 also account for the uncertainty in the distance decay parameters. If spatial weight matrices are prespecified, as in the first two columns of Table 1, this type of uncertainty is ignored, as if the researcher does know the right specification of the spatial weight matrix. This explains why the t-values of the direct and spillover effects reported in Table 1 are lower when the distance decay parameters are estimated together with the other parameters in model, unless they take on a clearly different value than that imposed by a pre-specified spatial weight matrix. In this respect note that the t-values of the spillover effects in column [3] compared to those in column [2] go down slightly. One example is the t-value of the initial level of GDP per capita, which decreases from 2.03 in column [2] to 2.00 in column [3]. Conversely, if the distance decay parameter in column [4] is substantially different from the reported value of 1.088 in column [3], the t-values of the spillover effects increase again. One example is again the t-value of the initial level of GDP per capita, which increases from 2.00 in column [3] to 2.43 in column [4], since its distance decay parameter of 2.224 is substantially different from 1.088.

The graphs in Figure 2 show several notable patterns. For the first two distance categories up to 50 kilometers, the spillover effect of lagged GDP per capita growth in neighboring regions is greater than the direct effect in the own region. Whereas the direct effect amounts to 0.030, the spillover effect can be as high as 0.076 in these distance categories. This can be explained by the fact that some regions in our sample are located so close to each other geographically that they form a cluster. It concerns regions around the capitals of Brussels, London, Berlin, Prague and the cities of the Hague and Rotterdam (located in the same region). It is to be noted that the situation of having neighboring regions within a distance of 50 kilometers only occurs for a limited number of regions in our sample (0.13%). Normally, one would expect the spillover effect to be smaller than the direct effect, even though it is a cumulative effect measured over all other regions in the sample (see eq. 6b). However, when regions form a cluster, such as the above urbanized areas, the spillover effect may exceed the direct effect.

The spillover effect of the investment rate also decreases with distance markedly, as does GDP per capita growth, but the difference is that it only exceeds the direct effect when it comes to nearby regions up to 25 kilometers. For the spillover effect of the population growth rate, which is negative, we see that its absolute value decreases with distance, gradually reaches a value of zero, and that even for nearby regions in the smallest distance category of 25 kilometers, it is approximately five times as small as the direct effect.

Finally, the spillover effect of the initial level of GDP per capita exhibits a more complex relationship with distance. It is negative at first, then decreases in magnitude with distance, becomes positive around 100 kilometers, increases further up to 350 kilometers and finally falls back to zero over a range of 350 to 1250 kilometers. It shows that nearby regions with high levels of GDP per capita strengthen the convergence effect, whereas regions with high levels of GDP per capita located farther away and especially in the range of 150 to 500 kilometers weaken the convergence effect. Regions that do already well in terms of growth apparently benefit from richer regions within this particular spatial range, which often concern centrally-located regions in the EU though in different countries.

Another notable observation from Figure 2 is that the slope with which the spillover effects decay and their spatial reach are different for different growth determinants. The first is most obvious for the initial level of GDP per capita, which follows a completely different distance decay pattern than the other growth determinants. The second is most obvious for the population growth rate, which turns out to have a spatial reach even beyond 1500 kilometers, whereas the spatial reach of the investment rate does not tend to be greater than 700 kilometers and of both the growth rate and the initial level of GDP per capita not to be greater than 1250 kilometers. If we would have adopted one common spatial weight matrix for all spatial lags in the model, as in columns [1], [2] or [3] of Table 1 and constructed the same graphs, their slope and spatial reach would be exactly the same for every growth determinant. More specifically, the graph of the initial level of GDP per capita would change in a downward sloping graph only, while the spatial reach of the population growth rate would become the same as that of the other growth determinants.

To illustrate this, Figure 3 graphs the spillover effects of the initial level of GDP per capita and the population growth rate based on the six nearest neighbor matrix and the estimation results reported in column [1] of Table 1. Instead of the inverse U-shaped form in Figure 2 starting with negative values first, the spillover effects of the initial level of GDP per capita in Figure 3 starts with positive values and indeed is downward-sloping only. Similarly, instead of differing spatial ranges in Figure 2, the spatial range of both curves in Figure 3 indeed amounts to the same value of 500 kilometers. Further note that these differences are consistent with the estimated spillover effects reported in columns [1] and [4] of Table 1. The summary measure of the spillover effects in column [4] is 2.6 times as large for the initial level of GDP per capita and 4.2 times as large for the population growth rate compared to their counterparts in column [1].

<< Figure 3 around here >>

We conclude that the sensitivity of the spillover effects to the specification of the spatial weight matrix contrasts with the relative stability of the coefficient estimates and the direct effects seen across the different columns in Table 1. This contrast raises concerns for anyone who works with spatial econometric models in applied research. Empirical studies that want to verify whether their results are robust for the specification of the spatial weight matrix, should put more emphasis on the spillover effects rather than the parameter estimates and should consider not only different spatial weight matrices, but also different ones for each spatial lag in their model.

### 5.4 Alternative model specifications

Since the empirical literature usually works with different variants of the spatially-augmented neoclassical growth framework, in this section we draw attention to four alternative specifications. Their results are reported in Table 2.

<< Table 2 around here >>

Column [1] shows the results when the lagged growth rate is removed from the baseline model. This simpler version has been estimated in several studies, among which the original study of Ertur and Koch (2007). Due to removing this regressor, the number of observations increases from 4522 to 4788. This model run shows that our baseline version, which includes both the lagged and spatially lagged growth rates, is a better choice as it also permits the determination of the spillover effects of the GDP per capita growth rate, which according to the first graph of Figure 2 are worth to consider. When conducting an LR test on the LogL values in column [1] of Table 2 and column [4] of Table 1, adjusted for the difference in the number of observations, we obtain a test statistic of 153.5 (p-value 0.00), which indicates that this simplification is rejected by the data.

Column [2] presents the results when the time fixed effects with homogenous coefficients are replaced by cross-sectional averages of the dependent variable at time $t$ and $t-1$ with heterogeneous coefficients. This extension might be better able to distinguish global and local spillovers from each other.[13] The objection to this model specification is that it does not properly specify the spatial association between the regressand and its regressors; Pesaran's CD-test statistic applied to the residuals of this specification takes a value of 8.41, which is outside the desired confidence interval of $(-1.96,+1.96)$.

Column [3] continues with the results when the baseline model is extended to include additional explanatory variables taken from endogenous growth models (Ertur and Koch, 2011; Jung and Lopez-Bazo, 2017). It concerns the share of the population with tertiary education, as a proxy for regional differences in educational attainment, and the share of employment in science and technology, as a proxy for the share of resources used in research and development. The added-value of this extension appears to be limited

[13] A detailed explanation of this approach is provided in Elhorst (2021), which is partly inspired by previous studies of Keller (2002), Ertur and Koch (2011) and Crescenzi and Rodriguez-Pose (2013).

though. Out of all the additional parameters, six in total, only the coefficient of the share of resources used in research and development appears to be statistically significant. When conducting an LR test on the LogL values in column [3] of Table 2 and column [4] of Table 1, we obtain a test statistic of 9.4 (p-value 0.15), which indicates that this extension is rejected by the data.

Column [4] shows the results when the baseline model is not estimated based on annual observations but three-year overlapping averages, as is more common in this literature, again when removing the lagged growth rate. Due to taking averages, the number of observations decreases from 4522 to 4256. When we compare the results in columns [1] and [4], we see that in particular the significance levels of almost all coefficient estimates and implicit effects improve. This finding may be due to a reduction in the impact of observations in the year 2009, which can be seen as an outlier according to Figure 1. If outliers are an issue, this approach is recommended to narrow the graphically displayed confidence intervals of the spillover effects.

## 6. Conclusion

In his contribution to the Handbook of Applied Spatial Analysis, Arthur Getis listed eleven attributes of the concept of spatial autocorrelation. In this paper we translate these attributes into present-day spatial econometrics and present a methodology to estimate the distance decay effect and spatial reach of spillover effects in a spatial Durbin model. We apply this methodology to study spillovers in GDP per capita growth across EU regions and to illustrate these effects and their confidence intervals as a function of distance.

This approach contrasts with the standard practice in empirical studies of routinely reporting for each regressor the direct and spillover effects as two numerical summary measures. Instead, the exposition of the spillover effects based on the graphs developed in this paper is a major step forward in the literature. This is because they disentangle the spillover effects as a function of distance, which is one of the major topics in regional science, spatial economics and economic geography. Since the spillover effects of the regressors tend to be the main focus of many spatial econometric studies, these graphs may contribute to a better understanding of these effects.

Our findings confirm not only the existence and importance of growth spillovers, but also their decrease in magnitude as distance increases, consistent with Tobler's first law of geography. Normally, one would expect the spillover effects to be smaller than the direct effects, but there might be exceptions. One such exception is that the distance to neighboring regions is so small that these regions can just as well be regarded as a cluster, as a result of which the spillover effects within this cluster exceed the direct effect.

By parameterizing the spatial weight matrix of each spatial lag by a different distance decay parameter, we also found that the spatial reach of the spillover effect of each regressor is no longer the same, which from an empirical viewpoint further enhances the flexibility of these effects. This finding highlights the restrictiveness of the SD model based on one common spatial weight matrix for all spatial lags, reflecting the standard in spatial econometric research up to now. In their 2009 spatial econometric textbook, LeSage and Pace (2009, pp.72-73) partitioned the spillover effects from first to ninth-ordered neighbors numerically in an attempt to disentangle the spillover effects across space. However, hardly any study has explored this further. We hope that graphing the spillover effects for each individual regressor, as in this paper, will be followed up in more studies.

**References**


Ades A, Chua HB (1997) Thy neighbor's curse: regional instability and economic growth. Journal of Economic Growth 2: 279-304

Anselin L, Bera AK, Florax R, Yoon MJ (1996) Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics 26(1): 77-104

Anselin L, Rey SJ (2014) Modern spatial econometrics in practice. GeoDa Press LLC.

Arbia G, Bera AK, Doğan O, Taşpınar S (2020) Testing impact measures in spatial autoregressive models. International Regional Science Review 43(1-2): 40-75

Audretsch DB, Feldman MP (2004) Knowledge spillovers and the geography of innovation. In Henderson JV, Thisse JF (eds) Handbook of regional and urban economics, vol 2. Elsevier, Amsterdam, pp 2713-2739

Autant-Bernard C, LeSage JP (2011) Quantifying knowledge spillovers using spatial econometric models. Journal of Regional Science 51(3): 471-496

Benos N, Karagiannis S, Karkalakos S (2015) Proximity and growth spillovers in European regions: the role of geographical, economic and technological linkages. Journal of Macroeconomics 43: 124-139

Beugelsdijk S, Klasing MJ, Milionis P (2018) Regional economic development in Europe: the role of total factor productivity. Regional Studies 52(4): 461-476

Billé AG, Tommelleri A, Ravazzola T (2023) Forecasting regional GDPs: a comparison with spatial dynamic spatial models. Spatial Economic Analysis, forthcoming

Bottazzi L, Peri G (2003) Innovation and spillovers in regions: evidence from European patent data. European Economic Review 47(4): 687-710

Breinlich H, Ottaviano GI, Temple JR (2014) Regional growth and regional decline. In Aghion P, Durlauf SN (eds) Handbook of economic growth, vol 2. Elsevier, Amsterdam, pp 683-779

Coe DT, Helpman E (1995) International R&D spillovers. European Economic Review, 39(5): 859-887

Crescenzi R, Rodríguez-Pose A (2013) R&D, socio-economic conditions, and regional innovation in the U.S. Growth and Change 44(2): 287-320

Döring T, Schnellenbach J (2006) What do we know about geographical knowledge spillovers and regional growth?: a survey of the literature. Regional Studies 40(3), 375-395

Easterly W, Levine R (1997) Africa's growth tragedy: policies and ethnic divisions. Quarterly Journal of Economics 112(4), 1203-1250

Elhorst JP (2014) Spatial econometrics: from cross-sectional data to spatial panels. Springer, Berlin

Elhorst JP (2021) Spatial panel models and common factors. In Fischer MM, Nijkamp P (eds) Handbook of regional science. Springer, Berlin, pp. 2141-2159

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

Enflo K, Hjertstrand P (2009) Relative sources of European regional productivity convergence: a bootstrap frontier approach. Regional Studies 43(5): 643-659

Ertur C, Koch W (2007) Growth, technological interdependence and spatial externalities: theory and evidence. Journal of Applied Econometrics 22(6): 1033-1062
Ertur C, Koch W (2011) A contribution to the theory and empirics of Schumpeterian growth with worldwide interactions. Journal of Economic Growth 16(3): 215-255
Ezcurra R, Rios V (2020) Quality of government in European regions: do spatial spillovers matter? Regional Studies 54(8): 1032-1042
Fischer MM, Getis A (eds) (2010) Handbook of applied spatial analysis: software tools, methods and applications. Springer, Berlin
Funke M, Niebuhr A (2005) Regional geographic research and development spillovers and economic growth: evidence from West Germany. Regional Studies 39(1) 143-153
Gardiner B, Martin R, Tyler P (2011) Does spatial agglomeration increase national growth? Some evidence from Europe. Journal of Economic Geography 11(6): 979-1006
Getis A (2010). Spatial autocorrelation. In Fischer MM, Getis A (eds) Handbook of applied spatial analysis. Springer, Berlin, pp 255-278
Glaeser EL, Kallal HD, Scheinkman JA, Shleifer A (1992) Growth in cities. Journal of Political Economy 100(6): 1126-1152
Grossman GM, Helpman E (1991) Trade, knowledge spillovers, and growth. European Economic Review 35(2-3): 517-526
Halleck Vega S, Elhorst JP (2015) The SLX model. Journal of Regional Science 55(3): 339-363
Henderson V, Kuncoro A, Turner M (1995) Industrial development in cities. Journal of Political Economy 103(5), 1067-1090
Islam N (1995) Growth empirics: a panel data approach. The Quarterly Journal of Economics 110(4): 1127-1170
Jaffe AB, Trajtenberg M, Henderson R (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics 108(3): 577-598
Juhl S (2021) The Wald test of common factors in spatial model specification search strategies. Political Analysis 29(2): 193-211
Jung J, López-Bazo E (2017) Factor accumulation, externalities, and absorptive capacity in regional growth: evidence from Europe. Journal of Regional Science 57(2): 266-289
Keller W (2002) Geographic localization of international technology diffusion. American Economic Review 92(1): 120-142
Lee LF, Yu J (2016) Identification of spatial Durbin panel models. Journal of Applied Econometrics 31(1): 133-162
LeGallo J (2014) Cross-section spatial regression models. In Fischer MM, Nijkamp P (eds) handbook of regional science. Springer, Berlin, pp 1511-1533
LeSage JP (2015) Software for Bayesian cross section and panel spatial model comparison. Journal of Geographical Systems 17(4): 297–310
Lesage JP, Fischer MM (2008) Spatial growth regressions: model specification, estimation and interpretation. Spatial Economic Analysis 3(3): 275-304
LeSage JP, Pace RK (2009) Introduction to spatial econometrics. Chapman and Hall/CRC
López-Bazo E, Vayá E, Artis M (2004) Regional externalities and growth: evidence from European regions. Journal of Regional Science 44(1): 43-73
Márquez MA, Ramajo J, Hewings GJ (2015) Regional growth and spatial spillovers: evidence from an SpVAR for the Spanish regions. Papers in Regional Science 94: S1-S18
Moreno R, Trehan B (1997) Location and the growth of nations. Journal of Economic Growth 2(4): 399-418

Pesaran MH (2015) Testing weak cross-sectional dependence in large panels. Econometric Reviews 34(6-10): 1089-1117
Pfaffermayr M (2012) Spatial convergence of regions revisited: a spatial maximum likelihood panel approach. Journal of Regional Science 52(5): 857-873
Ramajo J, Marquez MA, Hewings GJ, Salinas MM (2008) Spatial heterogeneity and interregional spillovers in the European Union: do cohesion policies encourage convergence across regions? European Economic Review 52(3): 551-567
Rodríguez-Pose A, Crescenzi R (2008) Research and development, spillovers, innovation systems, and the genesis of regional growth in Europe. Regional Studies 42(1): 51-67
Rosenthal SS, Strange WC (2020) How close is close? The spatial reach of agglomeration economies. Journal of Economic Perspectives 34(3): 27-49
Solow RM (1956) A contribution to the theory of economic growth. Quarterly Journal of Economics 70(1): 65-94
Swan TW (1956) Economic growth and capital accumulation. Economic Record 32(2): 334-361
Tan C (2023) Modelling spatial weight matrices and lags in spatial panel models. PhD thesis, University of Groningen.
Yu J, de Jong R, Lee LF (2012) Estimation for spatial dynamic panel data with fixed effects: the case of spatial cointegration. Journal of Econometrics 167: 16–37
Yu J, Lee LF (2012) Convergence: a spatial dynamic panel data approach. Global Journal of Economics 1(01): 1-36

Table 1: Estimation results of the spatially augmented neoclassical growth model for different spatial weight matrices−

| | [1] | | [2] | | [3] | | [4] | |
|---|---|---|---|---|---|---|---|---|
| | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat |
| Δy(t-1) | 0.014 | 0.94 | 0.012 | 0.78 | 0.009 | 0.62 | 0.008 | 0.56 |
| y(t-1) | −0.094*** | −16.73 | −0.089*** | −15.88 | −0.088*** | −15.84 | −0.090*** | −16.25 |
| inv(t) | 0.002 | 0.82 | 0.002 | 0.55 | 0.001 | 0.48 | 0.002 | 0.52 |
| popg(t) | −0.008*** | −5.91 | −0.008*** | −6.03 | −0.008*** | −6.06 | −0.008*** | −6.48 |
| $W(\delta_0)$*Δy(t) | 0.534*** | 30.85 | 0.650*** | 33.64 | 0.631*** | 23.54 | 0.641*** | 22.62 |
| $W(\delta_1)$*Δy(t-1) | 0.191*** | 7.62 | 0.194*** | 6.73 | 0.198*** | 6.95 | 0.174*** | 6.36 |
| $W(\delta_2)$*y(t-1) | 0.066*** | 8.60 | 0.070*** | 8.47 | 0.068*** | 8.27 | 0.087*** | 6.61 |
| $W(\delta_3)$*inv(t) | 0.009* | 1.94 | 0.009* | 1.80 | 0.010* | 1.94 | 0.009* | 1.73 |
| $W(\delta_4)$*popg(t) | −0.002 | −1.00 | −0.005* | −1.82 | −0.005* | −1.90 | −0.016 | −1.14 |
| $\delta_0$ | | | | | 1.088*** | 14.89 | 1.047*** | 13.70 |
| $\delta_1$ | | | | | | | 2.224*** | 3.33 |
| $\delta_2$ | | | | | | | 0.633*** | 4.22 |
| $\delta_3$ | | | | | | | 1.483 | 0.95 |
| $\delta_4$ | | | | | | | 0.233 | 1.17 |
| DE_Δy(t-1) | 0.035** | 2.48 | 0.030** | 2.13 | 0.029** | 1.96 | 0.030** | 2.08 |
| DE_y(t-1) | −0.092*** | −17.58 | −0.088*** | −16.49 | −0.087*** | −16.08 | −0.090*** | −16.43 |
| DE_inv(t) | 0.004 | 1.30 | 0.003 | 0.83 | 0.002 | 0.83 | 0.003 | 0.89 |
| DE_popg(t) | −0.009*** | −6.62 | −0.009*** | −7.19 | −0.009*** | −6.92 | −0.009*** | −7.15 |
| SE_Δy(t-1) | 0.404*** | 9.29 | 0.560*** | 8.08 | 0.532*** | 7.85 | 0.478*** | 7.29 |
| SE_y(t-1) | 0.031** | 2.57 | 0.035** | 2.03 | 0.032** | 2.00 | 0.080** | 2.43 |
| SE_inv(t) | 0.021*** | 2.69 | 0.028** | 2.32 | 0.027** | 2.48 | 0.027** | 2.19 |
| SE_popg(t) | −0.014*** | −3.29 | −0.028*** | −4.26 | −0.026*** | −4.19 | −0.059 | −1.48 |
| LogL | 10919.1 | | 10974.5 | | 10975.5 | | 10986.2 | |
| $R^2$ | 0.624 | | 0.633 | | 0.633 | | 0.635 | |
| $R^2$ Excluding fixed effects | 0.426 | | 0.429 | | 0.429 | | 0.431 | |
| CD test residuals | −0.605 | | −0.789 | | −0.802 | | −0.908 | |
| # Observations | 4522 | | 4522 | | 4522 | | 4522 | |

Notes: All variables are measured in natural logarithms. Regional and time fixed effects are controlled for in all columns. *,**,***=significant at respectively the 10%, 5% and 1% significance level. [1] = Estimates with 6 nearest neighbors matrix, [2] Estimates with negative exponential matrix but pre-specified distance decay parameter of 0.01, [3] Estimates with parameterized negative exponential matrix, distance decay parameter estimated, [4] Estimates with parameterized negative exponential matrix but different decay parameters for each spatially lagged variable.

Table 2: Estimation results of alternative specifications

| | [1] | | [2] | | [3] | | [4] | |
|---|---|---|---|---|---|---|---|---|
| | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat |
| Δy(t-1) | | | 0.022 | −1.53 | 0.007 | 0.46 | | |
| y(t-1) | −0.081*** | −15.55 | −0.083*** | −15.73 | −0.091*** | −16.29 | −0.099*** | −29.80 |
| inv(t) | 0.002 | 0.77 | −0.002 | −0.08 | 0.002 | 0.61 | 0.009*** | 4.46 |
| popg(t) | −0.007*** | −5.62 | −0.009*** | −7.27 | −0.008*** | −6.43 | 0.003*** | 2.58 |
| educ(t) | | | | | −0.004 | −0.91 | | |
| sci_tech(t) | | | | | 0.011** | 2.07 | | |
| $W(\delta_0)$*Δy(t) | | | 0.112*** | 5.57 | 0.176*** | 6.38 | | |
| $W(\delta_1)$*Δy(t-1) | 0.015*** | 2.86 | 0.098*** | 12.25 | 0.078*** | 5.33 | 0.091*** | 12.84 |
| $W(\delta_2)$*y(t-1) | −0.011 | −1.36 | 0.005* | 1.05 | 0.015** | 2.26 | 0.015*** | 3.84 |
| $W(\delta_3)$*inv(t) | 0.698*** | 26.86 | −0.007 | −1.00 | −0.019 | −1.16 | −0.050*** | −4.35 |
| $W(\delta_4)$*popg(t) | | | | | 0.070 | 0.78 | | |
| $W(\delta_5)$*educ(t) | | | | | −0.007 | −0.79 | | |
| $W(\delta_6)$*sci_tech(t) | | | 0.730*** | 27.01 | 0.637*** | 22.35 | | |
| $\delta_0$ | 0.965*** | 15.01 | 0.800*** | 14.46 | 1.054*** | 13.60 | 0.760*** | 39.88 |
| $\delta_1$ | | | 2.702*** | 2.30 | 2.213*** | 3.35 | | |
| $\delta_2$ | 0.676*** | 3.99 | 0.544*** | 6.67 | 0.661*** | 3.53 | 0.687*** | 7.80 |
| $\delta_3$ | 1.316 | 1.58 | 1.516 | 0.54 | 1.296 | 1.39 | 1.284** | 2.24 |
| $\delta_4$ | 0.352 | 1.29 | 0.282 | 0.92 | 0.218 | 1.21 | 0.366*** | 4.29 |
| $\delta_5$ | | | | | 0.163 | 0.75 | | |
| $\delta_6$ | | | | | 2.946 | 0.36 | | |
| DE_Δy(t-1) | | | −0.009** | −0.66 | 0.029** | 1.97 | | |
| DE_y(t-1) | −0.090*** | −16.43 | −0.081*** | −15.77 | −0.090*** | −16.51 | −0.100*** | −30.56 |
| DE_inv(t) | 0.003 | 0.89 | 0.000 | 0.09 | 0.003 | 1.16 | 0.012*** | 6.12 |
| DE_popg(t) | −0.009*** | −7.15 | −0.010*** | −8.01 | −0.009*** | −7.08 | 0.000 | 0.18 |
| DE_educ(t) | | | | | −0.003 | −0.65 | | |
| DE_sci_tech(t) | | | | | 0.011** | 2.07 | | |
| SE_Δy(t-1) | | | 0.344*** | 6.60 | 0.473*** | 7.27 | | |
| SE_y(t-1) | 0.080** | 2.43 | 0.138** | 6.16 | 0.057 | 1.53 | 0.067*** | 2.64 |
| SE_inv(t) | 0.027** | 2.19 | 0.016** | 1.25 | 0.042*** | 2.64 | 0.086*** | 6.70 |
| SE_popg(t) | −0.059 | −1.48 | −0.049 | −1.91 | −0.066 | −1.46 | −0.195*** | −4.16 |
| SE_educ(t) | | | | | 0.185 | 0.75 | | |
| SE_lnsci_tech(t) | | | | | 0.002 | 0.09 | | |
| LogL | 11551.8 | | 11391.9 | | 10990.9 | | 12908.5 | |
| $R^2$ | 0.617 | | 0.694 | | 0.633 | | 0.802 | |
| $R^2$ Excluding fixed effects | 0.373 | | 0.492 | | 0.431 | | 0.489 | |

| | | | | |
|---|---|---|---|---|
| # Observations | 4788 | 4522 | 4522 | 4256 |

Notes: All variables are measured in natural logarithms. Estimates are based on parameterized negative exponential matrices with different decay parameters for each spatially lagged variable. Regional and time fixed effects are controlled for in all columns. *,**,***=significant at respectively the 10%, 5% and 1% significance level. [1] = Model without lagged growth rates, [2] = Time fixed effects are replaced by cross-sectional averages of the regressand at time *t* and *t*-1 and heterogeneous coefficients, [3] = Model extended to include endogenous growth variables, [4] = Model estimated based on three-year averages rather than annual observations and without lagged growth rates.

Figure 1: The average GDP per capita growth rate across all regions over time

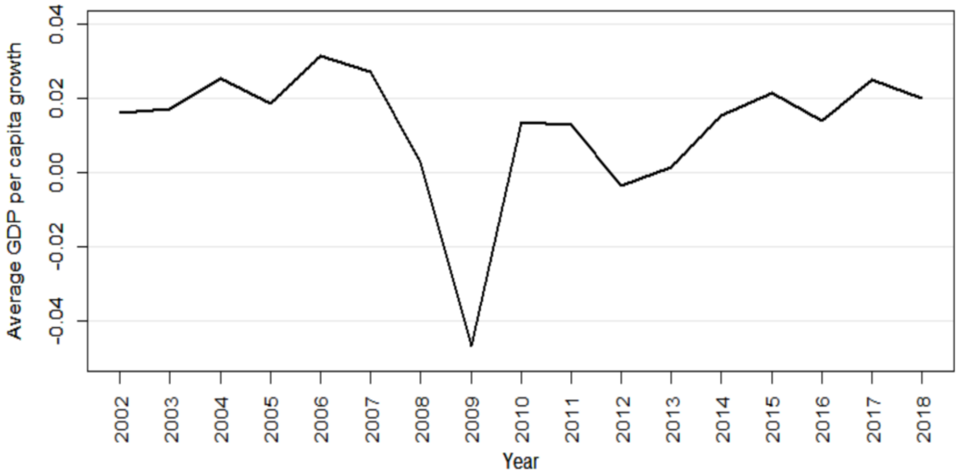
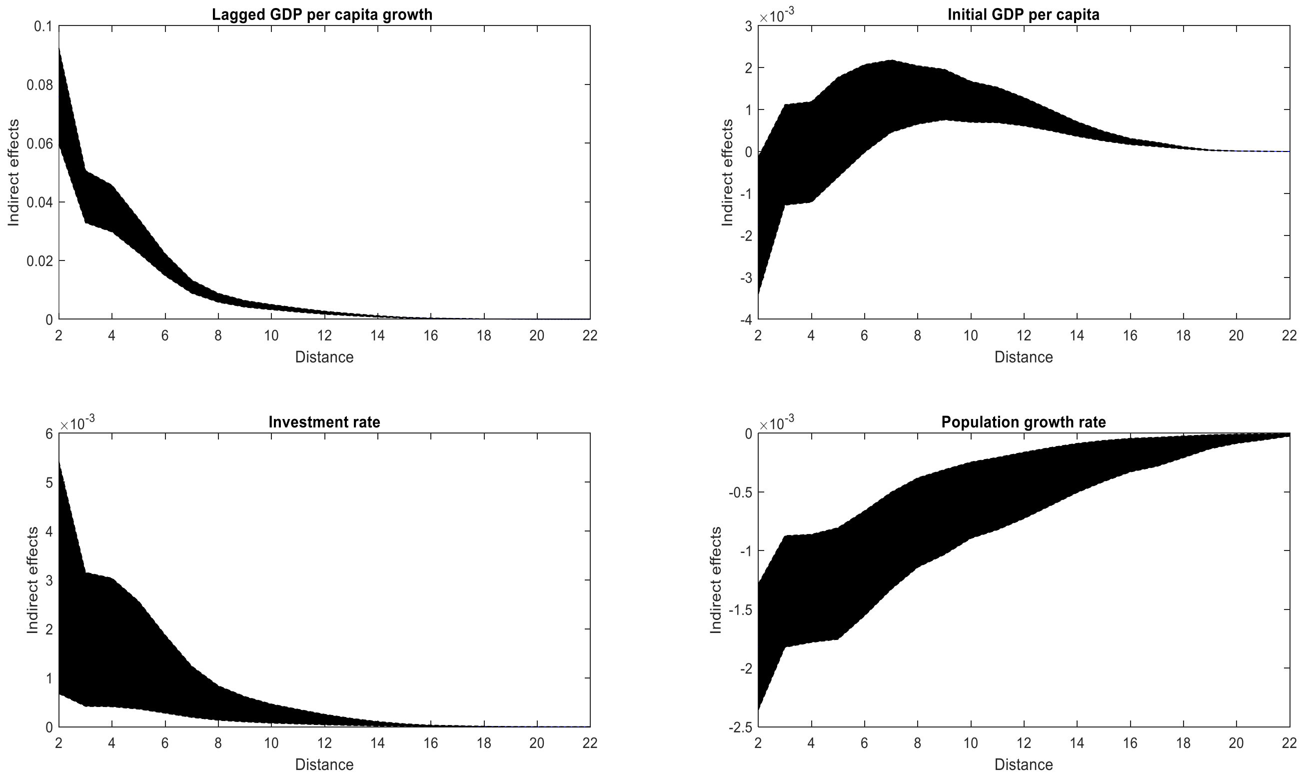
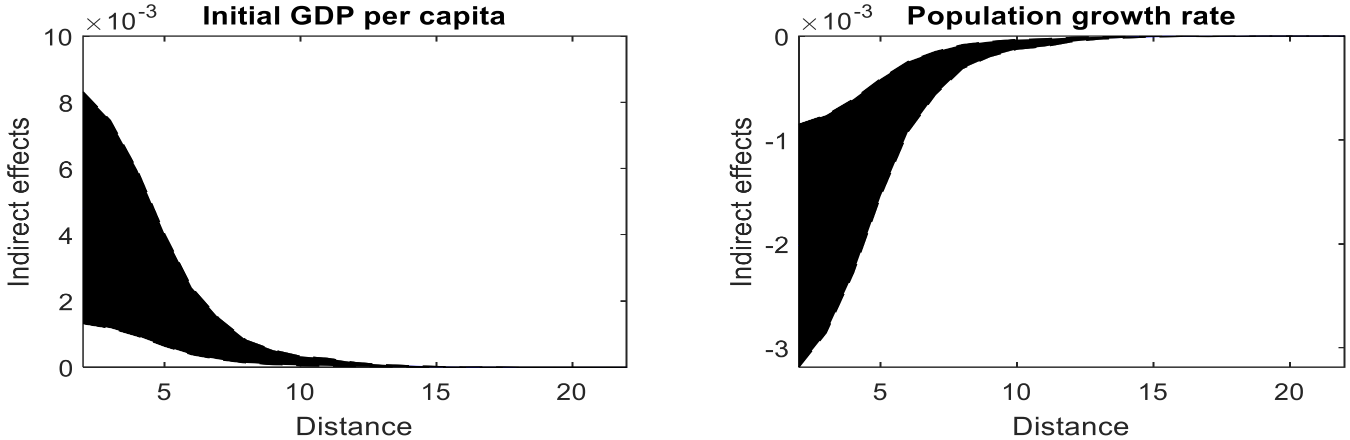


Figure 2: Spatial spillover effects of the four explanatory variables of GDP per capita growth as a function of distance




Notes: The solid lines denote spillover effects and the dotted lines the 95% confidence intervals. Spillover effects are synonymous with indirect effects. Distance is split up in 21 categories: 1 (0-25], 2 (25-50], 3 (50,100], 4 (100,150], 5 (150,200], 6 (200,250], 7 (250,300], 8 (300,350], 9 (350,400], 10 (400,450], 11 (450,500], 12 (500,600], 13 (600,700], 14 (700,800], 15 (800,900], 16 (900,1000], 17 (1000,1250], 18 (1250,1500], 19 (1500,1750], 20 (1750,2000], and 21 >2000 kilometers. The graphs are based on the estimation results reported in column [4] of Table 1.

Figure 3: Spatial spillover effects of two explanatory variables of GDP per capita growth as a function of distance based on one common six nearest neighbors spatial weight matrix




Notes: The solid lines denote spillover effects and the dotted lines the 95% confidence intervals. Spillover effects are synonymous with indirect effects. Distance is split up in 21 categories: 1 (0-25], 2 (25-50], 3 (50,100], 4 (100,150], 5 (150,200], 6 (200,250], 7 (250,300], 8 (300,350], 9 (350,400], 10 (400,450], 11 (450,500], 12 (500,600], 13 (600,700], 14 (700,800], 15 (800,900], 16 (900,1000], 17 (1000,1250], 18 (1250,1500], 19 (1500,1750], 20 (1750,2000], and 21 >2000 kilometers. The graphs are based on the estimation results reported in column [1] of Table 1.