

Article

Functional-Coefficient Spatial Durbin Models with Nonparametric Spatial Weights: An Application to Economic Growth

Mustafa Koroglu ^{1,2} and Yiguo Sun ^{1,*}

¹ Department of Economics and Finance, University of Guelph, Guelph, ON N1G2W1, Canada; mkoroglu@uoguelph.ca

² Department of Economics, Süleyman Şah University, Istanbul 34956, Turkey

* Correspondence: yisun@uoguelph.ca; Tel.: +1-519-824-4120

Academic Editor: Isabel Casas

Received: 6 November 2015; Accepted: 19 January 2016; Published: 3 February 2016

Abstract: This paper considers a functional-coefficient spatial Durbin model with nonparametric spatial weights. Applying the series approximation method, we estimate the unknown functional coefficients and spatial weighting functions via a nonparametric two-stage least squares (or 2SLS) estimation method. To further improve estimation accuracy, we also construct a second-step estimator of the unknown functional coefficients by a local linear regression approach. Some Monte Carlo simulation results are reported to assess the finite sample performance of our proposed estimators. We then apply the proposed model to re-examine national economic growth by augmenting the conventional Solow economic growth convergence model with unknown spatial interactive structures of the national economy, as well as country-specific Solow parameters, where the spatial weighting functions and Solow parameters are allowed to be a function of geographical distance and the countries' openness to trade, respectively.

Keywords: functional coefficients; local linear regression; nonparametric 2SLS estimator; series estimator; Solow economic growth convergence model; spatial Durbin model

JEL: C14; C21; O47

1. Introduction

Ever since the seminal work of [1], there has been a significant amount of empirical work studying variation in economic growth rates across countries. Particularly, more and more economists have paid attention to the impacts of economic interaction and spillover effects on the regional and national economy in the past two decades; see, e.g., [2] for taxation and the global allocation of capital, [3] for cross-border foreign direct investment decisions, [4,5] for economic growth models with worldwide interactions, [6] for country interactions in discretionary fiscal policy and [7] for an overview of empirical studies of strategic interaction among governments over environmental standards and public expenditures. In the meanwhile, econometric theory in parametric spatial regression models has been introduced and well developed to analyse the spatial and economic externalities; for a detailed survey on parametric spatial econometric models, see [8–11]. This paper joins the others to examine the impact of cross-country economic externalities on national growth through a Solow growth model augmented with economic externalities.

The role of spatial dependence in regional economic growth has received substantial attention in the empirical growth literature in the recent decade; see, e.g., [12] for a survey on economic growth and space. It has been recognized that a nation's per capita GDP growth rate is affected not only by its

own values of determinants, such as savings, population growth rate and initial level of income, but also by its neighbouring nations' per capita GDP growth rates and the values of these determinants. For example, Ertur and Koch [4] developed a theoretical growth model with spatial externality resulting from technological interdependence among economies and proposed a spatially-augmented Solow economic growth model yielding a conditional convergence equation with heterogeneous Solow parameters. Note that heterogeneous Solow parameters are also supported by similar studies with no spatial interactions; see, e.g., [13–15].

Fitting a spatial Durbin model (SDM) using data from 91 non-oil regions/and countries for the period from 1960 to 1995, Ertur and Koch [4] found positive and significant spatial dependence across these economies together with predicted signs for all coefficients. However, their study suffers from two potential problems. First, the parametric SDM requires researchers to pre-determine the non-stochastic spatial dependence structure among economies before estimating parameters appearing in the model, and the misspecified spatial interactive relations can incur inconsistent estimation and misleading inference. Second, the subsampling method may not be the best way of studying heterogeneous Solow parameters. In this paper, working on the same dataset used in [4], we therefore aim to re-examine the spatial spillover effects of economic growth, while estimating in a nonparametric way the true spatial dependence structure among economies and allowing the Solow parameters to vary with respect to the trade openness of an economy.

Specifically, we propose a functional-coefficient spatial Durbin model with nonparametric spatial weights and estimate the unknown spatial weights and coefficient curves via a series approximation approach by a nonparametric two-stage least squares method. Based on the first-step consistent estimator, we then construct a second-step estimator for the unknown functional coefficients, which is oracle efficient in the sense that the limit distribution of the second-step estimator is the same regardless of whether the spatial weights are known. Moreover, we give our inference on spatial dependence through average direct and indirect impact values with standard errors calculated from the bootstrap method.

The remainder of this paper is organized as follows. Section 2 introduces our proposed semiparametric spatial Durbin model. Section 3 presents our estimation methodology. Section 4 reports results from a small Monte Carlo simulation to examine the finite sample performance of our proposed estimators. Section 5 gives our empirical results. Section 6 concludes.

2. Model

In the empirical economic growth literature, DeLong and Summers [16], to the best of our knowledge, is the first study to investigate spatial correlation taking geographical distance into account. Using a sample of 61 countries for the period from 1960 to 1985, they find no significant spatial correlation in their sample. Moreno and Trehan [17], on the other hand, augments [1]'s model with a spatial interactive term and find highly significant spillover effects between geographical neighbours, and they argue that using a border dummy variable instead of a spatial lag term neglects the influence of neighbour countries that do not have a common border with the country of interest; relevant literature includes [18–20]. Moreover, Ertur, le Gallo, and Baumont [21] provide strong evidence of spatial dependence in economic convergence processes among European regional economies. Using 155 European regions over the period 1988–2000, Basile [22] also finds some evidence of spatial spillovers across countries.

The question of how to measure spatial interactive relations between any pair of spatial units is answered by defining a neighbourhood set for each spatial unit according to some selected relevant variables. For example, Cliff and Ord [8] specify spatial weights for spatial unit i as the ratio of the length of the common border between units i and j to the geographical distance between them, $g_{ij} = b_{ij}^{\beta}/d_{ij}^{\alpha}$, with some parameters $\alpha > 0$ and $\beta > 0$. A common approach in practice is to use only distance-based weights with a decay parameter α ; *i.e.*, the spatial weight from unit j on unit i is defined as $g_{ij} = g(d_{ij}, \alpha)$, where $g(\cdot)$ is a known function and α is a parameter to be

estimated. The popularly-used distance function includes an inverse power function $g_{ij} = 1/d_{ij}^\alpha$ or a negative exponential function $g_{ij} = \exp(-\alpha d_{ij})$ (e.g., [23]) for some $\alpha > 0$. Moreover, if a cut-off distance is not used, then a non-sparse spatial weight matrix is constructed. This implies that every region is a neighbour of other regions, but the spatial weights depreciate as the distance between two regions increases. For recent studies on spatial weights, see [24,25]. Moreover, the term “neighbours” can also refer to contiguities defined by economic distances (e.g., [26–28]) and social networks ([29]).

As the functional form $g(\cdot)$ is unknown in practice, to avoid misspecifying the wrong spatial weighting function, one can alternatively estimate the unknown spatial weighting function $g(\cdot)$ from the data via a nonparametric series estimation method. Compared to the parametric spatial modelling approach, the nonparametric approach enables researchers to impose less restrictive assumptions on the spatial weight function; see [27,30] for details. Alternatively, Ahrens and Bhattacharjee [31] proposed to estimate the unknown spatial weights via the LASSO estimation method when the unknown spatial weights matrix is sufficiently sparse.

Ertur and Koch [4] derive a theoretical Solow economic growth model augmented with global technological interdependence. They then approximate their theoretical model by a parametric spatial Durbin model via a linearization procedure and calculate the spatial weights from both the inverse power function and the exponential function of geographic distances with $\alpha = 2$ for the sake of robustness, as the true spatial weighting function is unknown. As the linearization and selected parametric spatial weights may result in a model misspecification problem, in this paper, in order to better approximate [4]’s theoretical model, we therefore propose a semiparametric growth model that extends [4]’s parametric model by allowing nonparametric spatial weights, as well as varying Solow coefficients. Specifically, our proposed functional-coefficient spatial Durbin model with nonparametric spatial weights is given by:

$$Y_i = \sum_{j \neq i} g(Z_{ij}) Y_j + \sum_{t=1}^p \sum_{j \neq i} m_t(Z_{ij}) X_{tj} + X_i^T \theta(D_i) + u_i, \quad i = 1, \dots, n, \quad (1)$$

where Y_i is a scalar dependent variable, $X_i = [X_{1i}, \dots, X_{pi}]^T$ is a $p \times 1$ vector, D_i is a continuous scalar random variable and Z is a non-stochastic spatial covariate with $Z_{ii} = 0$ and $Z_{ij} > 0$ for $i \neq j$. Moreover, $g(\cdot)$, $m_t(\cdot)$ for $t = 1, \dots, p$ and $\theta(\cdot)$ are all unknown measurable smooth functions with $g(0) = 0$ and $m_t(0) = 0$, $t = 1, \dots, p$. In Model (1), X_i has a functional coefficient depending on D_i , and the unknown spatial weights are a function of non-stochastic geographic distance Z_{ij} . The first two terms in the right-hand side of Equation (1), $\sum_{j \neq i} g(Z_{ij}) Y_j$ and $\sum_{j \neq i} m_t(Z_{ij}) X_{tj}$ for each $t = 1, \dots, p$, are called the spatial lag of the dependent variable and the spatially-lagged exogenous variables, respectively. In the spatial econometrics literature, this kind of model specification is referred to as the spatial Durbin model, where (X_i, D_i, u_i) are independently distributed across i while only the dependent variable Y_i is dependently distributed across spatial units. The detailed data information on X_i, D_i, Y_i and Z_{ij} is delayed to Section 5. If the model includes only the spatial lag of the dependent variable, it is called a pure spatial autoregressive (SAR) model. Basile *et al.* [32] introduced the spatial autoregressive semiparametric geoadditive models to account for spatial dependence, spatial unobserved heterogeneity and unknown functional curves of regressors simultaneously, where the spatial autoregression is represented by a pre-determined spatial lag term of the dependent variable.

Let G_n and $M_{n,t}$ be $n \times n$ unknown spatial weight matrices with its $(i, j)^{th}$ element being $g_{ij} = g(Z_{ij})$ and $m_{t,ij} = m_t(Z_{ij})$, respectively, for $t = 1, \dots, p$, $i = 1, \dots, n$, $j = 1, \dots, n$. We then obtain a reduced form of Model (1) written in matrix form:

$$Y = (I_n - G_n)^{-1} \left[\sum_{t=1}^p M_{n,t} X_t + mtk \{X, \theta(D)\} + U \right], \quad i = 1, \dots, n, \quad (2)$$

where $Y = [Y_1, \dots, Y_n]^T$, $X_t = [X_{t1}, \dots, X_{tn}]^T$ and $U = [u_1, \dots, u_n]^T$ are all $n \times 1$ vectors and $mtk\{X, \theta(D)\}$ is an $n \times 1$ vector with the i^{th} element equal to $X_i^T \theta(D_i)$. Furthermore, I_n is an $n \times n$ identity matrix.

Let $\lambda_i(A_n)$ be the i^{th} eigenvalue of an $n \times n$ matrix A_n , $\rho(A_n) = \max_{1 \leq i \leq n} |\lambda_i(A_n)|$ and $\|A_n\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ and $\|A_n\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$ be the respective row and column norm of A_n . Furthermore, $C > 0$ is a finite positive number that takes different values at different appearances. Below, we impose some regularity conditions on Model (1).

Assumption A1: (i) $\{Y_i\}$ is generated from Model (1), and $\{(X_i, D_i)\}$ is independently distributed with finite second moments; (ii) $g(\cdot)$, $m_t(\cdot)$ and $\theta(\cdot)$ are all uniformly bounded up to their respective p^{th} -order derivatives for some $p > 2$; (iii) $\{u_i\}$ is an independent sequence with zero mean, $E[u_i | X_i = x, D_i = d] = 0$ and $E[u_i^2 | X_i = x, D_i = d] = \sigma_i^2(x, d) > 0$ for all i and $(x, d) \in R^p \times R$, and $\sup_{(x,d) \in R^p \times R} \max_{1 \leq i \leq n} E[|u_i|^{2+\delta} | X_i = x, D_i = d] \leq C < \infty$ for some $\delta > 0$ and a positive constant C .

Assumption A2: (i) There exist a positive integer N and a constant $c_G \in (0, 1)$, such that for all $n > N$, $\rho(G_n) \leq c_G$; (ii) $\|G_n\|_j \leq C < \infty$, $\|M_{n,t}\|_j \leq C < \infty$ for all t and $\|(I_n - G_n)^{-1}\|_j \leq C < \infty$ for $j = 1$ and ∞ and some finite value $C > 0$.

Assumption A1 (i) states that the explanatory variables (X_i, D_i) are independent, while the dependent variable Y_i exhibits spatial dependence; and Assumption A1 (iii) allows the error term, u_i , to be independent with heteroskedasticity, and the bounded higher order moment is required for deriving the limiting normal distribution of the proposed estimator. By [33] (p. 421), Assumption A2 (i) ensures that $I_n - G_n$ is a non-singular matrix with $(I_n - G_n)^{-1} = \sum_{j=0}^{\infty} G_n^j$, which implies that $\{Y_i\}$ is spatially stationary. In addition, we have $\max_{1 \leq i, j \leq n} |g_{ij}| \leq \rho(G_n) < 1$ by Properties 4.66 and 4.67 in [33] (p. 68). It is ready to show that $n^{-1}Y^T Y = O_p(1)$, and $E(Y_i^2 | X_1 = x_1, \dots, X_n = x_n, D_1 = d_1, \dots, D_n = d_n)$ is continuously differentiable and uniformly bounded under Assumptions A1 and A2. In addition, Assumption A2 (ii) is a regularity condition (see, e.g., Assumption 1 in [34]), and it holds if the spatial weight function, $g(z)$, decreases to zero for large z and:

$$\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n \sum_{j=1}^n I(Z_{ij} \in \mathcal{Z}) < \infty \text{ for any fixed bound set } \mathcal{Z}, \tag{3}$$

where the indicator function $I(\mathcal{A}) = 1$ if event \mathcal{A} holds, and zero otherwise.

Note that a parametric SDM is given by $Y_i = \rho \sum_{j \neq i} w_{ij} Y_j + \delta \sum_{t=1}^p \sum_{j \neq i} p_{t,ij} X_{tj} + X_i^T \theta_0 + u_i$, $i = 1, \dots, n$, where W_n and $P_{n,t}$ are spatial weight matrices with their respective $(i, j)^{th}$ element equal to w_{ij} and $p_{t,ij}$. Therefore, if the parametric spatial Durbin model holds true, the spatial weight matrices G_n and $M_{n,t}$ in Model (1) are equivalent to ρW_n and $\delta P_{n,t}$ in the spatial Durbin model, respectively. From an estimation and econometric modelling viewpoint, the normalization of spatial weight matrices in the spatial Durbin model is used to identify the spatial multiplier parameters (ρ, δ) , but this is not necessary in our proposed Model (1). Therefore, allowing nonparametric spatial weights saves us from applying an *ad hoc* spatial weight matrix normalization procedure as in the parametric SDM.

3. Estimation Methodology

If the dependent variable exhibits spatial autocorrelation, it must be accounted for by incorporating the spatially-lagged dependent variable into the model. If this variable is not included in the model, there would be an omitted variable type specification error due to the fact that unobserved factors may have a direct effect on the response variable. Moreover, the presence of the spatial lag-dependent variable in the model results in a simultaneity bias problem. This can be seen explicitly from the reduced form Model (2). From Model (2), we see $E(G_n Y U^T) = G_n (I_n - G_n)^{-1} E(U U^T)$ is a non-zero matrix, so that the spatial lag of the dependent variable, $G_n Y$, is correlated with the error term, U , which, therefore, results in an endogeneity problem in Model (1). Therefore, the ordinary least squares (or OLS) estimator would be biased and inconsistent. Moreover,

the term $(I_n - G_n)^{-1}$ in Equation (2) explains that region i is affected not only by its own determinants, but also by its neighbouring regions' values. This has been called a global interaction effect in [4] (p. 1044) and [24] (p. 15). Another source of the spatial endogeneity problem is due to the endogeneity of spatial covariate in the model. Recently, Kelejian and Piras [35] and Sun [30] estimated the spatial panel data model and the SAR model with an endogenous spatial weight matrix in a nonparametric way, respectively. Moreover, Qu and Lee [36] proposed estimators for the parametric SAR model with an endogenous spatial covariate. This paper only deals with the endogeneity of the spatial lag of the dependent variable, as our spatial weight matrix is non-stochastic.

The endogeneity problem can be addressed by using the maximum likelihood estimation (or MLE) method, as well as the instrumental variable (or IV) approach. Ord [37] was the first to examine the MLE of SAR models. He proposed to use the eigenvalues of the spatial weights matrix to alleviate the computational complexity of the MLE method in large sample sizes. Lee [38] derived the large sample properties of the quasi-MLE without a normality assumption on error terms, while Bao and Ullah [39] obtained the second order bias of the maximum likelihood estimator for spatial autoregressive models. As the (quasi-) maximum likelihood estimator can be computationally difficult in moderate or large-sized samples, Kelejian and Prucha [40] proposed a two-stage least squares (or 2SLS) estimator for a SAR model with spatial autoregressive errors, while Lee [41] proposed an asymptotically-optimal 2SLS estimator. As the spatial weights in Model (1) are unknown, the 2SLS estimation methods derived in [40,41] are not feasible; we therefore use a series approximation method to recover the unknown spatial weight function and estimate all unknown functions via a nonparametric 2SLS (or NP2SLS) estimation method. For an overview of the sieve estimation method, see [42].

Specifically, we approximate the unknown weighting functions $g(\cdot)$ and $m_t(\cdot)$, $t = 1, \dots, p$, and the vector of functional coefficients $\theta(\cdot)$ by series expansions:

$$g^*(z) = \sum_{l=1}^{L_n} \alpha_l \phi_l(z), \tag{4}$$

$$m^*(z) = [m_1^*(z), \dots, m_p^*(z)]^T = \left[\sum_{l=1}^{L_n} \gamma_{1l} \phi_l(z), \sum_{l=1}^{L_n} \gamma_{2l} \phi_l(z), \dots, \sum_{l=1}^{L_n} \gamma_{pl} \phi_l(z) \right]^T, \tag{5}$$

and:

$$\theta^*(d) = [\theta_1^*(d), \dots, \theta_p^*(d)]^T = \left[\sum_{l=1}^{L_n} \beta_{1l} \phi_l(d), \sum_{l=1}^{L_n} \beta_{2l} \phi_l(d), \dots, \sum_{l=1}^{L_n} \beta_{pl} \phi_l(d) \right]^T, \tag{6}$$

respectively, where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{L_n})^T$, $\gamma_t = (\gamma_{t1}, \gamma_{t2}, \dots, \gamma_{tL_n})^T$ and $\beta_t = (\beta_{t1}, \beta_{t2}, \dots, \beta_{tL_n})^T$ for $t = 1, \dots, p$ are all $L_n \times 1$ vectors of unknown coefficients, $\{\phi_j(\cdot)\}_{j=1}^{L_n}$ is a sequence of square integrable orthonormal basis functions over the interval $[0, \infty)$ and L_n denotes the number of basis functions. The following assumption regulates the sparseness of the weight matrix and the smoothness of unknown functions.

Assumption A3: (i) There exists a positive constant sequence $\{v_n\}$, such that:

$$\max_{1 \leq l \leq L_n, 1 \leq i \leq n} \sum_{j \neq i}^n |\phi_l(Z_{ij})| \leq C v_n \text{ and } \max_{1 \leq l \leq L_n, 1 \leq j \leq n} \sum_{i=1}^n |\phi_l(Z_{ij})| \leq C v_n. \tag{7}$$

(ii) There exist $L_n \times 1$ vectors α , γ_t , and β_t for $t \in \{1, \dots, p\}$, such that:

$$\max_{1 \leq i \leq n} \sum_{j \neq i} |g(Z_{ij}) - \alpha^T \Phi_{L_n}(Z_{ij})| = O(L_n^{-\zeta}) \tag{8}$$

$$\max_{1 \leq t \leq p} \max_{1 \leq i \leq n} \sum_{j \neq i} |m_t(Z_{ij}) - \gamma_t^T \Phi_{L_n}(Z_{ij})| = O(L_n^{-\zeta}) \tag{9}$$

and:

$$\sup_{d \in R} \max_{1 \leq t \leq p} |\theta_t(d) - \beta_t^T \Phi_{L_n}(d)| = O(L_n^{-\zeta}),$$

respectively, for some $\zeta > 2$ as $L_n \rightarrow \infty$, where $\Phi_{L_n}(z) = [\phi_1(z), \phi_2(z), \dots, \phi_{L_n}(z)]^T$ is an $L_n \times 1$ vector of basis functions.

It is not necessary to know the exact order of v_n in Assumption A3 (i), and the consistency of our proposed estimator does not require $v_n \equiv 1$, as assumed in [27]’s Assumption (vi). From approximation theory in mathematics, Assumption A1 (ii) is a necessary condition for Assumption A3 (ii). However, (8) and (9) also require spatial units expanding sparsely as more spatial units are included, for example when (3) holds true; and the consistency of our proposed estimator relies on increasing domain asymptotic theory. Moreover, we use Laguerre polynomial series to approximate the unknown functions, as it is one of the common choices for series expansions when a function has a domain over $[0, \infty)$ ([42] (p. 5574)). In addition, L_n acts as a smoothing parameter that increases slowly with the sample size. In other words, it is required to have $L_n \rightarrow \infty$ and $L_n/n \rightarrow 0$ as $n \rightarrow \infty$. An introduction of series estimation methods in a nonparametric framework can be found in [43] (Chapter 15).

Now, we approximate Model (1) by:

$$Y_i \approx \sum_{l=1}^{L_n} \alpha_l \sum_{j \neq i} \phi_l(Z_{ij}) Y_j + \sum_{t=1}^p \sum_{l=1}^{L_n} \gamma_{tl} \sum_{j \neq i} \phi_l(Z_{ij}) X_{tj} + \sum_{t=1}^p \sum_{l=1}^{L_n} \beta_{tl} \phi_l(D_i) X_{ti} + u_i, i = 1, \dots, n. \quad (10)$$

To derive our first-step estimator, we rewrite Model (10) in matrix form as follows:

$$Y \approx V_n \xi + U, \quad (11)$$

where we denote the i^{th} row vector of an $n \times [(2p + 1)L_n]$ matrix V_n by:

$$V_{n,i}^T = [(\sum_{j \neq i} \Phi_{L_n}(Z_{ij}) Y_j)^T, (\sum_{j \neq i} \Phi_{L_n}(Z_{ij}) X_{1j})^T, \dots, (\sum_{j \neq i} \Phi_{L_n}(Z_{ij}) X_{pj})^T, (\Phi_{L_n}(D_i) X_{1i})^T, \dots, (\Phi_{L_n}(D_i) X_{pi})^T]$$

and a $[(2p + 1)L_n] \times 1$ vector of parameters $\xi = [\alpha^T, \gamma_1^T, \gamma_2^T, \dots, \gamma_p^T, \beta_1^T, \beta_2^T, \dots, \beta_p^T]^T$.

The specification of the instrumental variable matrix is of great importance to obtain a consistent estimator. Since the number of endogenous variables increases with the number of approximating functions, L_n , it is intuitively appealing to instrument the endogenous variables, $\sum_{j \neq i} \phi_l(Z_{ij}) Y_j$, $l = 1, \dots, L_n$, by $(\sum_{j \neq i} \Phi_{L_n}(Z_{ij}) X_{t_1j}) X_{t_2i}$ and $\sum_{j \neq i} \Phi_{L_n}(Z_{ij}) D_j$, $t_1, t_2 \in \{1, 2, \dots, p\}$ for $p > 1$ as in our empirical application; see, e.g., [44]. Since $X_{t_1j} X_{t_2i}$ and D_j are exogenous and relevant in predicting Y_j , we would expect the proposed instrumental variables to serve as valid instruments for $\sum_{j \neq i} \Phi_{L_n}(Z_{ij}) Y_j$. Therefore, we define the i^{th} row vector of an $n \times [(2p + 2)L_n]$ instrumental matrix Q_n as:

$$Q_{n,i}^T = [(\sum_{j \neq i} \Phi_{L_n}(Z_{ij}) X_{1j}) X_{1i})^T, (\sum_{j \neq i} \Phi_{L_n}(Z_{ij}) X_{1j})^T, \dots, (\sum_{j \neq i} \Phi_{L_n}(Z_{ij}) X_{pj})^T, (\Phi_{L_n}(D_i) X_{1i})^T, \dots, (\Phi_{L_n}(D_i) X_{pi})^T, (\sum_{j \neq i} \Phi_{L_n}(Z_{ij}) D_j)^T].$$

We then can estimate ξ from (11) by the 2SLS estimation method. Note that we do not pursue optimal instrument variables in this paper due to the complexity of this approach in the semiparametric setup and the fact that the oracle efficiency of the second-step estimator of $\theta(\cdot)$ does not rely on the use of optimal instruments in the first-step estimation.

To ensure the existence of our 2SLS estimator, we assume that the exogenous regressors matrix X_n , the instrumental variables matrix Q_n and $V_n^T Q_n(Q_n^T Q_n)^{-1} Q_n^T V_n$ all have full column rank. Moreover, for the relevance of the instruments, we assume that $E[Q_n^T V_n]$ has a full column rank. Otherwise, we can remove linearly-dependent terms as long as the number of instruments in Q_n is more than the number of endogenous variables L_n plus the number of exogenous regressors $2p$. Lee [45] (p. 493) argues that the 2SLS estimator would be inconsistent if (X_i, D_i) are both irrelevant in predicting $\{Y_i\}$. Therefore, throughout this paper, we assume that X and D contain relevant variables in predicting $\{Y_i\}$ and $\theta(\cdot)$ takes non-zero values over any non-empty interval, so that there is no need to use the quadratic moments as additional orthogonal relations, as suggested in [45]. Our empirical application in this paper satisfies this assumption by both economic theory and empirical findings observed from the economic growth literature.

To construct a consistent estimator for $g(\cdot)$, $m_t(\cdot)$ and $\theta_t(\cdot)$, $t = 1, \dots, p$, we consider the following nonparametric 2SLS objective function:

$$\min_{\xi} \left[Q_n^T (Y - V_n \xi) \right]^T \left[Q_n^T (Y - V_n \xi) \right]. \tag{12}$$

The nonparametric 2SLS estimator of ξ solves (12) and is given by:

$$\hat{\xi} = [V_n^T Q_n(Q_n^T Q_n)^{-1} Q_n^T V_n]^{-1} V_n^T Q_n(Q_n^T Q_n)^{-1} Q_n^T Y,$$

and hence, the corresponding nonparametric 2SLS estimators¹ of unknown functions are given by:

$$\hat{g}(z) = \sum_{l=1}^{L_n} \hat{\alpha}_l \phi_l(z), \tag{13}$$

$$\hat{m}(z) = [\hat{m}_1(z), \dots, \hat{m}_p(z)]^T = \left[\sum_{l=1}^{L_n} \hat{\gamma}_{1l} \phi_l(z), \sum_{l=1}^{L_n} \hat{\gamma}_{2l} \phi_l(z), \dots, \sum_{l=1}^{L_n} \hat{\gamma}_{pl} \phi_l(z) \right]^T, \tag{14}$$

and

$$\hat{\theta}(d) = [\hat{\theta}_1(d), \dots, \hat{\theta}_p(d)]^T = \left[\sum_{l=1}^{L_n} \hat{\beta}_{1l} \phi_l(d), \sum_{l=1}^{L_n} \hat{\beta}_{2l} \phi_l(d), \dots, \sum_{l=1}^{L_n} \hat{\beta}_{pl} \phi_l(d) \right]^T. \tag{15}$$

Next, we propose a second-step estimator for the functional coefficients, $\theta(d)$, using the local linear regression approach. We would expect the local linear estimate of $\theta(d)$, $\tilde{\theta}(d)$, to have an improvement over the first-step estimator, $\hat{\theta}(d)$. Sun [30] considered a semiparametric spatial autoregressive model that has a mathematical representation of Model (1) with $M_{n,t} = 0$ for all $t = 1, \dots, p$ and has recently shown that the local linear estimator of $\theta(\cdot)$ can be oracle efficient under some regularity conditions in the sense that its limiting distribution does not depend on whether or not the spatial weights are known. This is a general result from non-/semi-parametric additive models. As the unknown functions $(g(Z_{ij}), m_1(Z_{ij}), \dots, m_p(Z_{ij}))$ and $\theta(D_i)$ enter Model (1) additively, we expect that $\tilde{\theta}(d)$ is oracle efficient, as well. We assume that as $n \rightarrow \infty$, $h \rightarrow 0$, $nh \rightarrow \infty$ and $nh^5 \rightarrow c \in (0, \infty)$, where h is the bandwidth, which controls the size of the local neighbourhood around an interior point d . Moreover, let $K(\cdot)$ be a kernel function, which assigns more weights to the data closer to point d , satisfying: (i) $\int K(a) da = 1$; (ii) $K(a) = K(-a)$; and (iii) $\int a^2 K(a) da > 0$.

The estimation procedure for $\tilde{\theta}(d)$ is given as follows:

(i) We replace $g(z)$ and $m_t(z)$ in (1) by $\hat{g}(z)$ and $\hat{m}_t(z)$, respectively, and treat $\hat{Y}_i = Y_i - \sum_{j \neq i} \hat{g}(Z_{ij}) Y_j - \sum_{t=1}^p \sum_{j \neq i} \hat{m}_t(Z_{ij}) X_{tj}$ as the dependent variable.

¹ The asymptotic properties of the series estimator without an endogenous variable follows from [46] and with some extensions from [47]. Moreover, we expect the asymptotic properties of our proposed estimators closely follow [30].

(ii) Applying the first-order Taylor series expansion of $\theta(D)$ around d , $\theta(D) \approx \theta(d) + \theta'(d)(D - d)$, we calculate the local linear estimator from a minimization of a kernel-weighted objective function:

$$(\tilde{\theta}(d), \tilde{\theta}'(d)) = \arg \min_{\theta(d), \theta'(d)} \sum_{i=1}^n [\hat{Y}_i - X_i^T \theta(d) - X_i^T \theta'(d)(D_i - d)]^2 K((D_i - d)/h) \tag{16}$$

where $\tilde{\theta}(d)$ estimates $\theta(d)$ and $\tilde{\theta}'(d)$ estimates $\theta'(d)$, the first order derivative of $\theta(d)$.

For a complete treatment of local linear estimator, see [48]. As the mathematical proofs of the consistency of the first-step estimator of $(\hat{g}(z), \hat{m}_1(z), \dots, \hat{m}_p(z), \hat{\theta}(d)^T)^T$ and the limiting result of the second-step estimator $\tilde{\theta}(d)$ closely follow those given in [30], the proofs are omitted from the paper.

4. Monte Carlo Simulations

In this section, we present the results from a very small Monte Carlo simulation study to assess the finite-sample properties of our estimators and more simulation results can be obtained from the authors upon request. We generate the data from the following regression model:

$$Y_i = \sum_{j \neq i} g(Z_{ij}) Y_j + \sum_{j \neq i} m(Z_{ij}) X_j + X_i \exp(-(4D_i - 1)^2) + u_i, i = 1, 2, \dots, n, \tag{17}$$

where we randomly draw $u_i \sim \text{i.i.d.}N(0, 0.5)$, $D_i \sim \text{i.i.d.}U[0, 1]$ and $X_i = 0.5D_i + \eta_i$ with $\eta_i \sim \text{i.i.d.}N(0, 1)$ independent of $\{u_i\}$. For the exogenous variable, Z , we first randomly generate n observations from the $U[0, R_n]$ distribution with $R_n = 0.001n^{1.6}$, by which we control the sparseness of spatial units. Then, we calculate Z_{ij} as the absolute distance between observations i and j . The specification of spatial weight functions requires that $g(\cdot)$ and $m(\cdot)$ are both decreasing and non-negative functions. We therefore set $g(z) = m(z) = 0.01 \exp(-z/0.01)$ for $z > 0$ with $g(0) = 0$ and $m(0) = 0$. The random variables, u_i , D_i and Z_i , are all mutually independent.

We consider a sample size $n \in \{100, 200, 400\}$. The number of replications is 1000 for each n in the Monte Carlo experiments. Moreover, we set $L_n = 1, 2, 3$ for each sample size, respectively. In the second-step estimation of coefficient functions, we select the bandwidth via a cross-validation method and use the Gaussian kernel function. To measure the performance of the estimators, we compute the root mean squared errors (or RMSEs) for each simulation. In Table 1, we report the averages of the RMSEs computed over 1000 repetitions, where $\hat{g}(\cdot)$, $\hat{m}(\cdot)$ and $\hat{\theta}(\cdot)$ denote the NP2SLS estimators of $g(\cdot)$, $m(\cdot)$ and $\theta(\cdot)$, respectively; $\tilde{\theta}(\cdot)$ is the second-step estimator of $\theta(\cdot)$, and $\vec{\theta}(\cdot)$ is the local linear estimator of $\theta(\cdot)$, while $g(\cdot)$ and $m(\cdot)$ are known. Furthermore, we also estimate the average direct impact (ADI) and the average indirect impact (AII) and report their corresponding RMSEs in Table 1. Specifically, we first obtain the reduced form model from (17):

$$Y = (I_n - G_n)^{-1} [M_n X + \theta(D) \circ X + U], \tag{18}$$

where Y , X , $\theta(D)$ and U are all $n \times 1$ vectors, and " \circ " denotes the Hadamard multiplication. Then, the expected marginal effect of X is given by the following $n \times n$ matrix:

$$\frac{\partial E(Y|X, D)}{\partial X} \equiv (I_n - G_n)^{-1} [M_n + \text{diag}\{\theta(D)\}] = S(G_n, M_n, D)$$

from which we obtain $\text{ADI} = n^{-1} \text{tr}\{S(G_n, M_n, D)\}$ and $\text{AII} = n^{-1} \mathbf{i}'_n S(G_n, M_n, D) \mathbf{i}_n - \text{ADI}$; see LeSage and Pace [10], where \mathbf{i}_n is the $n \times 1$ vector of ones and $\text{diag}\{\theta(D)\}$ is a diagonal matrix. Replacing the two unknown spatial weight matrices and $\theta(D)$ by their estimates, we obtain the estimates for ADI and AII.

From Table 1, we observe that there is a decrease in the RMSEs for all three estimators as the sample size increases in each design. Moreover, the second-step estimator always performs better than the nonparametric 2SLS estimator. The relative ratios of the RMSEs of the second-step estimator

and $\vec{\theta}(\cdot)$ generally reduce as the sample size increases. Therefore, our simulation results support the consistency of our proposed estimators. As for the ADI and AII, we also see an overall decreasing pattern in the RMSEs as the sample size increases, where the AII is less accurately estimated than the ADI, as the former is calculated from $n(n - 1)$ terms and the latter is calculated from n elements only.

Table 1. Average RMSEs.

n	$\hat{g}(\cdot)$	$\hat{m}(\cdot)$	$\hat{\theta}(\cdot)$	$\tilde{\theta}(\cdot)$	$\vec{\theta}(\cdot)$	ADI	AII
100	0.2109	0.0734	0.3535	0.2159	0.1557	0.1061	0.4386
200	0.1901	0.0593	0.2579	0.1680	0.1168	0.0761	0.2913
400	0.0500	0.0160	0.2013	0.1001	0.0859	0.0432	0.2989

5. Empirical Application

Monte Carlo simulations results given in Section 4 support the consistency of our proposed estimation method. We are now in a position to re-investigate cross-country growth patterns. We want to evaluate the impact of a country’s initial income, savings rate, population growth rate and openness, as well as neighbour countries’ economic growth spillovers on a country’s economic growth rate. We follow [4] in using a sample of 91 countries listed in [1], which is the Heston-Summers data taken from Penn World Table 6.1. Consider the following conditional convergence Solow growth model ²:

$$\begin{aligned}
 Y_i = & \sum_{j \neq i} g(Z_{ij})Y_j + \sum_{j \neq i} m(Z_{ij}) \ln(y60_j) + \theta_1(\text{lopen}_i) + \theta_2(\text{lopen}_i) \ln(y60_i) \\
 & + \theta_3(\text{lopen}_i) \ln(s_i) + \theta_4(\text{lopen}_i) \ln(n_i + 0.05) + u_i,
 \end{aligned}
 \tag{19}$$

for $i = 1, 2, \dots, 91$, where Y_i is the i^{th} country’s average growth rate of real GDP per capita between 1960 and 1995, $y60_i$ is the i^{th} country’s initial real GDP per capita in 1960, s_i is the i^{th} country’s average saving rate, n_i is the i^{th} country’s average growth rate of working-age population (ages between 15 and 64), lopen_i is a scalar development index of a country defined as the logarithm of the i^{th} country’s average ratio of total imports plus exports over its real GDP over the period from 1960 to 1995, and Z_{ij} is the great-circle distance between i^{th} and j^{th} countries’ capitals ³.

We approximate $g(\cdot)$, $m(\cdot)$ and $\theta_i(\cdot)$, $i = 1, 2, 3, 4$, using the Laguerre polynomials with $L_n = 2$. Moreover, a cross-validation selected bandwidth, h_{opt} , is calculated as 0.5285. We obtained $\rho(\hat{G}_n) = 0.041$, which suggests a spatial stationarity in the data. The distribution of the estimated residuals is approximately normal as the q-q plot of estimated residuals is close to linear. The coefficient estimates, $\tilde{\theta}_1(\cdot)$, $\tilde{\theta}_2(\cdot)$, $\tilde{\theta}_3(\cdot)$ and $\tilde{\theta}_4(\cdot)$ are presented in Figure 1, where the solid lines with circles display the second-step estimates. The dashed lines represent the estimates of the spatially-augmented Solow growth model of [4] using the inverse power spatial weight function, which we include as a baseline. We interpret Figure 1 as follows. First, in Figure 1b, we see that there is a negative relation between the initial level of income and the economic growth rate, except for Mauritius, Hong Kong, Zambia, Cameroon and Singapore, which confirms a conditional β -convergence hypothesis. Moreover, we observe that $\tilde{\theta}_2(\cdot)$ is increasing in openness, which,

² Since the sample size is less than 100, we include only one spatially-lagged exogenous variable, $M_n \ln(y60)$, to have better finite sample estimation accuracy. Moreover, the reason behind this choice is that spatial lag effects from the savings rate and the population growth rate were not found significant in Ertur and Koch [4]’s Table IV.

³ We follow [4] in calculating the variable Z_{ij} :

$$Z_{ij} = radius * \arccos[\cos(|long_i - long_j|) \cos(lat_i) \cos(lat_j) + \sin(lat_i) \sin(lat_j)],$$

where $radius$ is taken as the Earth’s radius and lat_i and $long_i$ are the latitude and longitude for country i , respectively.

however, results in a gradually declining degree of convergence. In addition, we see that the nonparametric model reveals slightly weaker conditional economic growth convergence as compared to the parametric model.

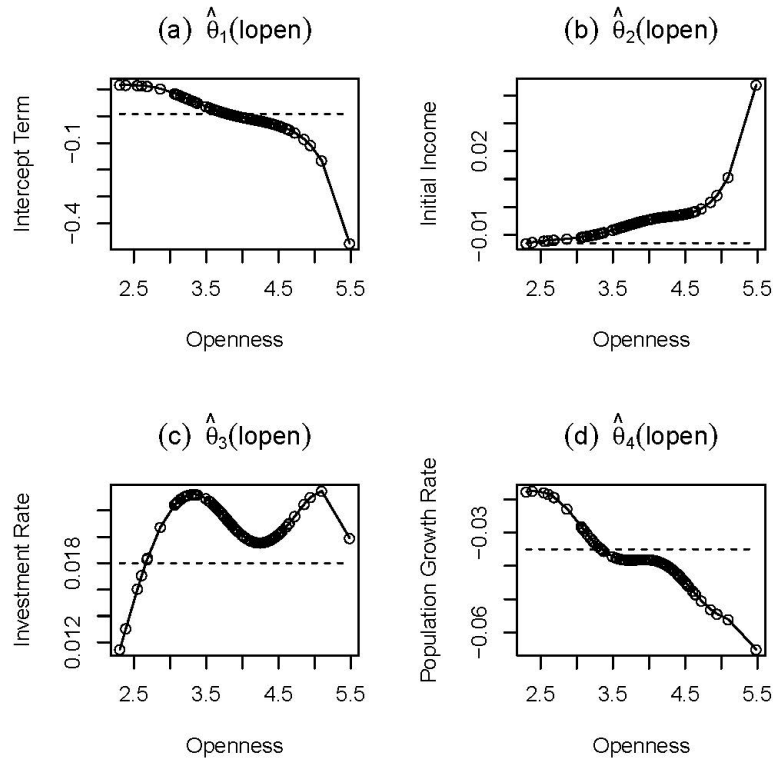


Figure 1. Estimated coefficient curves.

Second, in Figure 1c, we see that $\tilde{\theta}_3(\cdot)$ exhibits a positive, but not a monotonic, relation between the real investment rate and the real GDP per capita growth rate. Our estimate of the coefficient of the investment rate fluctuates as the trade openness of countries increases. For the economies with a trade openness higher than 15% of GDP, our result indicates that the nonparametric model sees stronger positive impact of the investment rate on the real GDP per capita growth rate than the parametric model does. Third, in Figure 1d, it is observed that the population growth rate has a negative impact on the real GDP per capita growth rate. For the countries whose trade openness ranges between 29% and 65%, our estimates for the coefficient of the population growth rate are relatively flat. Moreover, we note that the magnitude of the negative effect of the population growth rate is getting larger as the trade openness of countries increases, especially when the trade openness is over 65% of GDP. Overall, Figure 1 can be interpreted as the fact that an open economy suffers from higher negative impact of the population growth rate, but at the same time takes the advantage of high initial real GDP per capita.

Next, due to the cross-country interactions through spatial weights, the functional coefficient estimates have a different interpretation than the one obtained from the non-spatial model. In order to correctly interpret these estimates, we rewrite the estimated model in a reduced form as follows:

$$\begin{aligned}
 Y = & (I_n - \hat{G}_n)^{-1} [\hat{M}_n \ln(y60) + \tilde{\theta}_1(\text{lopen}) + \tilde{\theta}_2(\text{lopen}) \odot \ln(y60) \\
 & + \tilde{\theta}_3(\text{lopen}) \odot \ln(s) + \tilde{\theta}_4(\text{lopen}) \odot \ln(n + 0.05) + \tilde{U}], \tag{20}
 \end{aligned}$$

where \tilde{U} is the $n \times 1$ vector of residuals. Then, from Equation (20), the marginal effects are given by the following $n \times n$ matrices:

$$\begin{aligned}\frac{\partial E(Y|y60, s, n)}{\partial \ln(y60^T)} &\equiv (I_n - G_n)^{-1} [M_n + \text{diag}\{\theta_2(\text{lopen})\}] \\ \frac{\partial E(Y|y60, s, n)}{\partial \ln(s^T)} &\equiv (I_n - G_n)^{-1} \text{diag}\{\theta_3(\text{lopen})\} \\ \frac{\partial E(Y|y60, s, n)}{\partial \ln(n^T + 0.05)} &\equiv (I_n - G_n)^{-1} \text{diag}\{\theta_4(\text{lopen})\}\end{aligned}$$

where we define $\text{diag}\{a\}$ as an $n \times n$ diagonal matrix with the elements of an $n \times 1$ vector, a , on the main diagonal. Following LeSage and Pace [10], we label the diagonal elements of each matrix given above as the direct impacts and off-diagonal elements as the indirect impacts.

In Table 2, we report the estimated average direct impact (ADI) and average indirect impact (AII) of the explanatory variables, where the latter can be easily defined as the difference between average total impact⁴ and the average direct impact. Average direct and indirect impacts from the parametric model of [4] are denoted as ADI_{EK} and AII_{EK} , respectively. The interpretation of Table 2 is as follows. Firstly, we observe that a 1% increase in the real initial GDP per capita of an economy, holding other factors fixed, results in a decrease by 0.5% in its own real GDP per capita growth rate. However, this change increases the rest of the economies' economic growth rates by 0.01% on average due to the spatial dependence. From another point of view, a 1% increase in all of the regions'/nations' initial real GDP per capita speeds up this economy's real GDP per capita growth by 0.01%. This result indicates a positive spillover effect of the initial level of income. Secondly, a 1% increase in this economy's real investment rate increases its own real GDP growth rate by 2.08% on average. However, this change slows down the rest of the nations' real GDP growth by 0.08% on average. Thirdly, we see that a 1% increase in the population growth rate of this economy retards its own economic growth by 3.8% on average, but helps to improve the rate of economic growth of the rest of the countries by 0.16% on average.

Fourthly, when comparing our results to [4]'s results, we find that both the nonparametric and the parametric model give almost the same average direct effects of the initial per capita income, investment rate and population growth rate on the economic growth rate and that both models result in the same signs in the average direct and indirect effects. However, the AII values from the nonparametric model are much smaller than the results from the parametric model in absolute value, especially for the initial per capita income and the population growth rate. This is not surprising, as the parametric model assumes that all of the spatial weights take non-negative values, while our nonparametric spatial weights are estimated from the data without such a restriction. Although it is popular practice to assume non-negative spatial weights, this is an assumption imposed without support from econometric or economic theory. For example, trade treaties and monetary policies are both double-edged swords that may bring opposite impacts to different national economies, and non-negative spatial weights may not be able to capture the opposite interactions among different economies. As both [4]'s parametric SDM and our proposed semiparametric SDM approximate the unknown true relationships in their own best capacity, however, our model imposes less restrictions than the parametric SDM and is believed to bring a better fit to the data and more reliable inference. Although the numbers are different, both models give the same sign in estimated direct and indirect

⁴ As it is stated by LeSage and Pace [10] (p. 37), average total impact can be expressed in two different ways, however, which give the same numerical results. The first viewpoint states an influence from a change in the initial real GDP per capita of an economy on all of the regions, while the second viewpoint states an impact of changes in the initial real GDP per capita of the entire economy on a region/nation.

effects. Overall, we observe that the average direct and indirect effects can take opposite signs, and the effect of the former is much stronger than that of the latter in absolute magnitude.

Table 2. Average direct and indirect impact estimates.

	ln(y60)	ln(s)	ln(n + 0.05)
ADI	−0.0050 (−0.0122, −0.0005)	0.0208 (0.0089, 0.0341)	−0.0381 (−0.0782, −0.0054)
ADI _{EK}	−0.0119 (−0.0159, −0.0078)	0.0184 (0.0139, 0.0229)	−0.0336 (−0.0585, −0.0094)
AII	0.0001 (−0.0077, 0.0056)	−0.0008 (−0.0174, 0.0190)	0.0016 (−0.0347, 0.0289)
AII _{EK}	0.0140 (0.0052, 0.0244)	−0.0018 (−0.0169, 0.0124)	0.0275 (−0.0321, 0.0860)

Note: A 95% bootstrap percentile confidence interval is given in the parenthesis.

LeSage and Pace [10] (p. 39) explain how to obtain the standard errors for the ADI and AII estimates via a simulation method. In the parametric setup, as the spatial weight matrices are known, theoretically, one can apply the delta method to obtain the standard errors, and the simulation method tends to provide at best an approximation as one does not know the exact distribution of the estimated coefficients in finite samples; however, this method is feasible as the average direct and indirect impacts only depend on a finite number of unknown parameters. As our proposed semiparametric model contains both unknown spatial weights and unknown coefficient curves, the simulation method would involve simulating from a joint distribution with dimension equal to $2n(n - 1)$ (two spatial weight matrix estimates) plus $4n$ (four coefficient curve estimates) or 16,744 in our empirical application. Therefore, the simulation method is infeasible for our empirical interest here. As both the ADI and AII are in the form of sample averages and it is well known that the bootstrap method can be used to estimate the sample average and its standard error well (e.g., [49]), we decide to report our bootstrap estimates of the confidence intervals for the ADI and AII.

Below, we explain a residual-based bootstrap method to test whether the ADI and AII are significantly different from zero at the 5% significance level. We use the nonparametric bootstrap percentile method to construct a 95% confidence interval. Following [50,51], we first estimate the functional coefficients using an oversmoothed bandwidth, which tends to zero at a slower speed than the optimal bandwidth. Then, we obtain estimated residuals. The rest of the bootstrap procedure is given below.

1. Resample the estimated residuals and obtain the bootstrap errors, U^b .
2. Calculate:

$$Y^b = (I_n - \hat{G}_n)^{-1} [\hat{M}_n \ln(y60) + \tilde{\theta}_1^*(lopen) + \tilde{\theta}_2^*(lopen) \odot \ln(y60) + \tilde{\theta}_3^*(lopen) \odot \ln(s) + \tilde{\theta}_4^*(lopen) \odot \ln(n + 0.05) + U^b],$$

where $\tilde{\theta}_k^*(\cdot)$, $k = 1, 2, 3, 4$ are coefficient estimates using a larger bandwidth than the optimal bandwidth. Call $\{\ln(y60_i), \ln(s_i), \ln(n_i + 0.05), Y^b\}_{i=1}^n$ the bootstrap sample.

3. Estimate Model (19) from the bootstrap sample, and record $\tilde{\theta}_1^b(\cdot)$, $\tilde{\theta}_2^b(\cdot)$, $\tilde{\theta}_3^b(\cdot)$ and $\tilde{\theta}_4^b(\cdot)$, the bootstrap estimates of the functional coefficients.
4. Calculate the bootstrap value of average direct and indirect impacts of the explanatory variables, ADI^b and AII^b , respectively.
5. Repeat Steps 1–4 for 999 times.
6. Find the 0.025th and 0.975th empirical percentile of the 999 bootstrap values of ADI and AII and the point estimates given in Table 2 to establish the 95% bootstrap percentile confidence interval.

The confidence intervals are reported in Table 2. We see that the ADI values are statistically significantly different from zero at the 5% significance level. Moreover, we find that there is no significant effect on average from neighbouring countries per capita initial income, savings rate and population growth rate on economic growth rate of the country of interest. The same inference is obtained for the parametric model, except that the parametric model sees a significant average indirect impact of initial per capita income. Note that insignificant AII's do not imply that the indirect impact from economy i on economy j is insignificant for all (i, j) .

Figure 2 presents estimated spatial weighting functions. We plot both estimated spatial weighting functions, $\hat{g}(\cdot)$ and $\hat{m}(\cdot)$, for the geographic distances ranging from zero to 20 in 100 km, as the estimated spatial weights in the absolute value have an average of 5.069×10^{-7} and 3.893×10^{-9} , respectively, when $z > 20$. In Figure 2a,b, firstly, we see negative spatial weights, which greatly contradicts traditional parametric spatial regression models, which often assume non-negative spatial weights. Negative spatial interactions are indeed common in practice, especially in social networks; see [52,53] for strategic interactions within the monetary policy committee of the Bank of England. Secondly, both spatial weight functions are not strictly monotonic and exhibit convexity among nations that are not very far apart, concavity among nations with moderately far distances and a zero spatial weight function among nations that are far away. Moreover, the spatial weight functions take bigger absolute values among nations with smaller distance apart and smaller absolute values among far-away nations, which implies a relatively larger economic interaction among nearby nations than among far-away nations. Lastly, in Figure 2b, we observe positive estimated spatial weights when the distance ranges between 0.229 and 1.894, which correspond to 22.9 and 189.4 km, respectively. For the distances greater than 189.4 km, negative spatial weights are getting closer to zero as the distance between two countries increases. As the turning point 1.894 is really small and occurs for countries with a small area, our results imply that spatial interaction is very strong and different between two nearby small countries with small areas than between two countries with longer distances when at least one country has a large area.

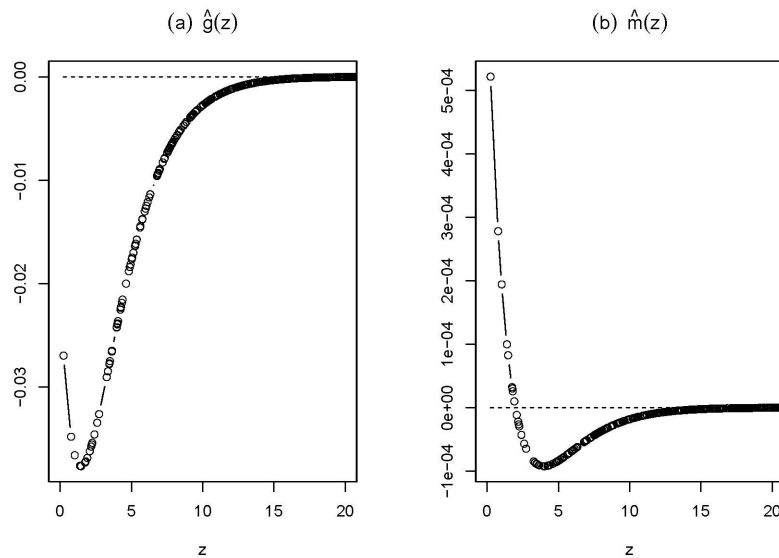


Figure 2. Estimated spatial weighting functions.

6. Conclusions

We employ a spatial Durbin model combined with the nonparametric spatial weighting functions, as well as the unknown functional coefficients to estimate the augmented Solow growth model with a sample of 91 countries over the period 1960 to 1995. We find a negative spatial lag

effect of neighbouring country's GDP per capita growth rate and initial GDP per capita on the economic growth rate of country *i*. These effects are declining in magnitude as the geographical distance between the two countries increases. Finally, allowing coefficients as a function of trade openness of a country enables us to see the true country-specific effect of each determinant of economic growth. Moreover, we find significant average direct impact from each production factors. However, our findings show that the average indirect impact of these variables is insignificant at the 5% significance level.

Acknowledgements: We would like to thank two anonymous referees for their insightful and instructive comments. We are also grateful for comments by participants at the 8th Midwest Graduate Student Summit hosted by Purdue University at Lafayette in April 2015, the 49th Annual Conference of the Canadian Economics Association hosted by Ryerson University in Toronto in May 2015 and the 32nd Meeting of the Canadian Econometric Study Group hosted by University of Guelph at Guelph in September 2015.

Author Contributions: The authors contributed equally to the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mankiw, N.G.; Romer, D.; Weil, D.N. A contribution to the empirics of economic growth. *Q. J. Econ.* **1992**, *107*, 407–437.
2. Backus, D.; Henriksen, E.; Storesletten, K. Taxes and the global allocation of capital. *J. Monet. Econ.* **2008**, *55*, 48–61.
3. Baltagi, B.H.; Eggerb, P.; Pfaffermayr, M. Estimating regional trade agreement effects on FDI in an interdependent world. *J. Econom.* **2008**, *145*, 194–208.
4. Ertur, C.; Koch, W. Growth, technological interdependence and spatial externalities: Theory and evidence. *J. Appl. Econom.* **2007**, *22*, 1033–1062.
5. Ertur, C.; Koch, W. A contribution to the theory and empirics of Schumpeterian growth with worldwide interactions. *J. Econ. Growth* **2011**, *16*, 215–255.
6. Cassette, A.; Creel, J.; Farvaque, E.; Paty, S. Governments under influence: Country interactions in discretionary fiscal policy. *Econ. Model.* **2013**, *30*, 79–89.
7. Brueckner, J.K. Strategic interaction among governments: An overview of empirical studies. *Int. Reg. Sci. Rev.* **2003**, *26*, 175–188.
8. Cliff, A.; Ord, J.K. *Spatial Processes: Models & Applications*; Pion: London, UK, 1981.
9. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1988.
10. LeSage, J.P.; Pace, R.K. *Introduction to Spatial Econometrics*; Chapman and Hall/CRC: New York, NY, USA, 2009.
11. Anselin, L.; Florax, R. *New Directions in Spatial Econometrics*; Springer: Berlin, Germany, 2011.
12. Abreu, M.; de Groot, H.L.F.; Florax, R.J.G.M. Space and growth: A survey of empirical evidence and methods. *Région et Développement* **2005**, *21*, 12–43.
13. Durlauf, S.N.; Kourtellos, A.; Minkin, A. The local Solow growth model. *Eur. Econ. Rev.* **2001**, *45*, 928–940.
14. Mamuneas, T.P.; Savvides, A.; Stengos, T. Economic development and the return to human capital: A smooth coefficient semiparametric approach. *J. Appl. Econom.* **2006**, *21*, 111–132.
15. Kourtellos, A. Modeling parameter heterogeneity in cross-country regression models. In *Economic Growth and Development (Frontiers of Economics and Globalization)*; de la Grandville, O., Ed.; Emerald Group Publishing Limited: Bingley, UK, 2011; pp. 367–387.
16. De Long, J.B.; Summers, L. H. Equipment investment and economic growth. *Q. J. Econ.* **1991**, *106*, 445–502.
17. Moreno, R.; Trehan, B. Location and the growth of nations. *J. Econ. Growth* **1997**, *2*, 399–418.
18. Chua, H.B. On Spillovers and Convergence. Ph.D. Thesis, Harvard University, Cambridge, MA, USA, 1993.
19. Ades, A.; Chua, H.B. Thy neighbour's curse: Regional instability and economic growth. *J. Econ. Growth* **1997**, *2*, 279–304.
20. Barro, R.J.; Sala-i-Martin, X. *Economic Growth*; McGraw-Hill: New York, NY, USA, 1995.
21. Ertur, C.; le Gallo, J.; Baumont, C. The European regional convergence process, 1980–1995: Do spatial regimes and spatial dependence matter? *Int. Reg. Sci. Rev.* **2006**, *29*, 3–34.

22. Basile, R. Regional economic growth in Europe: A semiparametric spatial dependence approach. *Pap. Reg. Sci.* **2008**, *87*, 527–544.
23. Murdoch, J.C.; Rahmatian, M.; Thayer, M.A. A spatially autoregressive median voter model of recreation expenditures. *Public Financ. Q.* **1993**, *21*, 334–350.
24. LeSage, J.P. What regional scientists need to know about spatial econometrics. *Rev. Reg. Stud.* **2014**, *44*, 13–32.
25. LeSage, J.P.; Pace, R.K. The biggest myth in spatial econometrics. *Econometrics* **2014**, *2*, 217–249.
26. Case, A.C.; Rosen, H.S. Budget spillovers and fiscal policy interdependence. *J. Public Econ.* **1993**, *52*, 285–307.
27. Pinkse, J.; Slade, M.E.; Brett, C. Spatial price competition: A semiparametric approach. *Econometrica* **2002**, *70*, 1111–1153.
28. Conley, T.G.; Ligon, E. Economic distance and cross-country spillovers. *J. Econ. Growth* **2002**, *7*, 157–187.
29. Lee, L.F. Identification and estimation of econometrics models with group interactions, contextual factors and fixed effects. *J. Econom.* **2007**, *140*, 333–374.
30. Sun, Y. *Functional-Coefficient Spatial Autoregressive Models with Nonparametric Spatial Weights*; Working Paper; University of Guelph: Guelph, ON, Canada, 2015.
31. Ahrens, A.; Bhattacharjee, A. Two-step Lasso estimation of the spatial weights matrix. *Econometrics* **2015**, *3*, 128–155.
32. Basile, R.; Durbán, M.; Mínguez, R.; Montero, J.M.; Mur, J. Modeling regional economic dynamics: Spatial dependence, spatial heterogeneity and nonlinearities. *J. Econ. Dyn. Control* **2014**, *48*, 229–245.
33. Seber, G.A.F. *A Matrix Handbook for Statisticians*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2008.
34. Su, L. Semiparametric GMM estimation of spatial autoregressive models. *J. Econom.* **2012**, *167*, 543–560.
35. Kelejian, H.H.; Piras, G. Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. *Reg. Sci. Urban Econ.* **2014**, *46*, 140–149.
36. Qu, X.; Lee, L.F. Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *J. Econom.* **2015**, *184*, 209–232.
37. Ord, K. Estimation methods for models of spatial interaction. *J. Am. Stat. Assoc.* **1975**, *70*, 120–126.
38. Lee, L.F. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* **2004**, *72*, 1899–1925.
39. Bao, Y.; Ullah, A. Finite sample properties of maximum likelihood estimator in spatial models. *J. Econom.* **2007**, *137*, 396–413.
40. Kelejian, H.H.; Prucha, I.R. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *J. Real Estate Financ. Econ.* **1998**, *17*, 99–121.
41. Lee, L.F. Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. *Econom. Rev.* **2003**, *22*, 307–335.
42. Chen, X. Large sample sieve estimation of semi-nonparametric models. In *Handbook of Econometrics 6B*; Heckman, J.J., Leamer, E.E., Eds.; Springer: New York, NY, USA, 2007; pp. 5549–5632.
43. Li, Q.; Racine, J.S. *Nonparametric Econometrics*; Princeton University Press: Princeton, NJ, USA, 2007.
44. Anselin, L.; Bera, A. Spatial dependence in linear regression models with an introduction to spatial econometrics. In *Handbook of Applied Economic Statistics*; Ullah, A., Giles, D.E.A., Eds.; Marcel Dekker: New York, NY, USA, 1998; pp. 237–289.
45. Lee, L.F. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *J. Econom.* **2007**, *137*, 489–514.
46. Andrews, D.W. Asymptotic normality of series estimators for nonparametric and semiparametric regression models. *Econometrica* **1991**, *59*, 307–345.
47. Newey, W.K. Convergence rates and asymptotic normality for series estimators. *J. Econom.* **1997**, *79*, 147–168.
48. Fan, J.; Gijbels, I. *Local Polynomial Modelling and Its Applications*; Chapman and Hall/CRC: New York, NY, USA, 1996.
49. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman & Hall: New York, NY, USA, 1993.
50. Härdle, W.; Marron, J.S. Bootstrap simultaneous error bars for nonparametric regression. *Ann. Stat.* **1991**, *19*, 778–796.
51. Sun, Y. A consistent nonparametric equality test of conditional quantile functions. *Econom. Theory* **2006**, *22*, 614–632.

52. Bhattacharjee, A.; Holly, S. Structural interactions in spatial panels. *Empir. Econom.* **2011**, *40*, 69–94.
53. Bhattacharjee, A.; Holly, S. Understanding interactions in social networks and committees. *Spat. Econ. Anal.* **2013**, *8*, 23–53.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).