Regional productivity growth in Europe: a Schumpeterian perspective

Roberto Basile*

June 20, 2014

Abstract

Using data for the European regions at NUTS-2 level, we test the predictions of a micro-founded Schumpeterian growth model with technological interdependence recently developed by Ertur and Koch (2011, EK11). Spatial interdependence is identified by means of a *semiparametric geoadditive spatial autoregressive model* which permits us to disentangle the effect of nonlinearities, spatial heterogeneity and spatial dependence. A control function approach is applied to estimate this particular SAR-type model using the spatial lag of the *quality of regional governance* and the spatial lags of *various social capital measures* as instrumental variables for the endogenous term $\mathbf{Wy}$. The results corroborates the predictions of EK11's model: R&D investments and R&D spillovers are important divers of regional growth in Europe. However, spillover effects are much lower after controlling for spatial unobserved heterogeneity. Moreover, important nonlinearities in the effect of physical capital investments emerge, putting into question the strong homogeneity assumption and suggesting a threshold effect in growth behavior.

Keywords: Regional growth, spatial dependence, nonlinearities, semiparametric models.
Jel codes: R11, R12, C14

*Department of Economics, Second University of Naples, Corso Gran Priorato di Malta, 1 - 81043, Capua (CE), Italy. *Email*: roberto.basile@unina2.it

1 Introduction

The empirical literature on regional growth is quite large (Magrini, 2004; Rey and Le Gallo, 2009; Ertur and Le Gallo, 2009) and it has increased a lot since the seminal paper by Barro and Sala-i-Martin (1995). The most recent contributions for the case of European regions provide clear evidence of club convergence and strong spatial interdependence (Basile, 2009; Fiaschi and Lavezzi 2007; Fotopoulos 2008). These results have led scholars to question the explanatory power of neoclassical exogenous growth models and to look at endogenous growth theories as suitable frameworks to interpret the actual regional development (Basile and Usai, 2014). Particularly appealing are those models which identify a large set of self-reinforcing mechanisms that can cause poverty traps (Azariadis and Stachurski, 2005) as well as those Schumpeterian models which emphasize the role of technology transfer as a driving force for economic growth and club convergence (Howitt, 2000; Howitt and Mayer-Foulkes, 2005; Acemoglu et al., 2006). However, this group of studies lacks the necessary micro-foundations to model interregional knowledge diffusion. Specifically, it does not properly take the issues related to spatial proximity into account.

First attempts to "regionalize" endogenous growth theory have been essentially non-analytical approaches, focusing on the issue of the boundaries of knowledge spillovers (Doring and Schnellenbach, 2006). These studies have questioned how geographically limited knowledge diffusion may help explain clusters of regions with persistently different levels of growth. The intrinsic limitations of these frameworks have raised the need for theoretical studies focusing on the explicit incorporation of space into growth models.

Over the past few years, there has been some work in this direction. Specifically, a group of authors has proposed extensions of multi-country growth models that include technological interdependence across regions to take account of neighborhood effects in growth and convergence processes. The most recent contribution is the one by Ertur and Koch (2011) (EK11), that is an extension of Howitt's (2000) multi-country endogenous Schumpeterian growth model. The growth equation predicted by this model can be classified as a Spatial Durbin specification of Howitt's (2000) model; this equation also encompasses other widely used empirical growth equations, such as the multi-region neoclassical growth model with and without technological interdependence.

Capturing interregional growth spillovers is certainly an important task and, in this context, EK11's model represents one of the most important contributions. However, other potential sources of misspecification bias in the empirical regional growth analysis should also be considered, especially those concerning the functional form specification and the unobserved spatial heterogeneity. First, the linearity assumption implicit in Howitt's and in EK11's models may lead to strong biases in presence of threshold effects and non-monotonicity in growth behavior. Indeed, strong *nonlinearities* have been already detected in several regional growth studies (Azomahou, Ouardighi, Nguyen-Van, and Pham, 2011; Basile and Gress, 2005; Basile, 2008, 2009; Basile, Capello, and Caragliu, 2012; Ertur and Gallo, 2009; Fotopoulos, 2012). Controlling for *unobserved heterogeneity* is another fundamental challenge in empirical growth analysis, as failing to do so can introduce omitted-variable biases and preclude causal inference. To complicate the analysis, spatial interdependence may simply be the consequence of (spatially

correlated) omitted variables rather than being the result of spillovers.

In a nutshell, we can say that modeling regional economic growth requires the adoption of complex econometric tools, which allow us to deal with important methodological issues, such as spatial dependence, unknown functional form and unobserved heterogeneity. Recently, some authors have proposed new econometric approaches (namely, *Penalized Spline Spatial Autoregressive and Spatial Durbin Models*, or more simply PS-SAR and PS-SDM) which address these issues simultaneously (see Basile et al. 2014, for an overview of this literature).

Using data for the European regions at NUTS-2 level, in this paper we apply a PS-SAR model to empirically test the predictions of multi-country endogenous growth models with technological interdependence without imposing a functional form and controlling for the effect of unobserved spatial heterogeneity. Specifically, a control function approach is applied to estimate this particular SAR-type model using the spatial lag of *quality of regional governance* and spatial lags of various measures of *social capital* (instead of higher-order spatial lags of $\mathbf{x}$) as instrumental variables for the endogenous term $\mathbf{Wy}$. The results corroborates the predictions of EK11's model, with R&D investments generating both a direct and an indirect effect on regional growth. However, R&D effects are much lower after controlling for spatial unobserved heterogeneity. Important nonlinearities also emerge, especially in the effect of physical capital investments, putting into question the strong homogeneity assumption and suggesting a threshold effect in growth behavior.

The rest of the paper is organized as follows. In the next section we discuss the main issues concerning the specification, the estimation and the identification of empirical regional growth models. Section 3 reports the results of the econometric analysis. Section 4 concludes and suggests new directions for future research in this field.

2 Modeling regional growth: econometric specification, identification and estimation

2.1 Econometric specification: neighboring effects

During the second half of the 1990's several empirical studies have provided evidence of spatial contagion in regional growth behavior, thus challenging the cross-region independence assumption implicitly adopted by previous works (for a review, see Abreu et al. 2005; Rey and Janikas, 2005; and Fingleton and Lopez-Bazo, 2006). Using spatial econometrics techniques, these studies have shown that regional growth rates depend crucially on the growth rates and initial (and structural) conditions of nearby economies, rather than just on any one region's own initial (and structural) conditions. When interpreting their results, these authors make reference to the notion of (geographically bounded) interregional knowledge spillovers (or spatial technological interdependence), without formally demonstrating the linkage between the two.

More recent studies (Lopez-Bazo et al., 2004; Egger and Pfaffermayer, 2006; Pfaffermayr, 2009a, b; Ertur and Koch, 2007, 2011) have instead shown that spatial technological interdependence can be explicitly modeled in multi-country (or multi-region) exogenous and endogenous growth frameworks to account for neighborhood effects in growth and convergence processes. These studies have provided sound theoretical foundations for the specific form taken by spatial

autocorrelation in econometric growth models. Thus, they have further stimulated the empirical assessment of the existence of neighboring effects in regional growth (Rey and Le Gallo, 2009).

The most recent study in this direction is the one by EK11, which is an extension of the multi-country Schumpeterian growth model with technology transfer elaborated by Howitt (2000). As well known, the Schumpeterian framework is characterized by a technological gap (convergence) approach, where the vertical R&D productivity of each region i is a positive function of the technological gap of this region from the technological frontier. Thus, the further away a region is from the technological frontier, the higher its productivity in the research sector, because it can benefit from the accumulated knowledge in other regions (*advantage of backwardness* conferred to technological laggards). According to Howitt (2000), however, all regions share the same global technological frontier, since each region diffuses the same quantity of knowledge. EK11, instead, assume that the access of region i to the accumulated knowledge of any other region j is conditional to the proximity relationship between i and j. In other words, in EK11 the technological frontier becomes *local* (i.e. it is specific to each region), being defined as a geometric weighted average of knowledge levels in all regions, with weights given by some measure of bilateral distance between the regions. With this change in the theoretical framework, the authors introduce the assumption of spatial technological interdependence within the Schumpeterian model. Because of the direct relationship between R&D productivity and the region specific technological gap, all regions undertaking R&D activity converge to the same steady-state (world) growth rate and, thus, to parallel growth paths, like in Howitt (2000) and Solow (1956).

Given the assumption of spatial technological interdependence, the (empirical counterpart of the) growth equation of per worker income of each region i (with $i = 1,...,N$) can be written as:

$$\begin{aligned} \gamma_i &= \beta_0 + \beta_1 \ln y_{0,i} + \beta_2 \ln \frac{s_{k,i}}{n_i + g_w + \delta} + \beta_3 \ln s_{A,i} + \beta_4 \ln n_i \\ &\quad + \theta_1 \sum_{j \neq i}^{N} w_{ij} \ln y_{0,j} + \theta_2 \sum_{j \neq i}^{N} w_{ij} \ln \frac{s_{k,j}}{n_j + g_w + \delta} + \rho \sum_{j \neq i}^{N} w_{ij} \gamma_j + \varepsilon_i \\ &\quad \varepsilon_i \sim iid\mathcal{N}\left(0, \sigma_\varepsilon^2\right) \qquad i = 1,...,n \end{aligned} \tag{1}$$

where γ_i is the average annual growth rate of income per worker of region i computed for the period between time 0 and time T and $y_{0,i}$ is its income per worker at time 0; $s_{k,i}$ is the regional rate of investment in physical capital; $(n_i + g_w + \delta)$ is the regional effective depreciation rate, with n_i the regional employment growth rate, g the common exogenous technology growth rate and δ the rate of depreciation of physical capital assumed identical in all economies; and $s_{A,i}$ measures the regional intensity of investment is R&D. The other variables ($\sum_{j \neq i}^{N} w_{ij} \ln y_{0,j}$, $\sum_{j \neq i}^{N} w_{ij} \ln \frac{s_{k,j}}{n_j + g_w + \delta}$ and $\sum_{j \neq i}^{N} w_{ij} \gamma_j$) are spatial lag terms measuring the weighted average values of the neighbors of region i, with elements w_{ij} indicating the connectivity between a region i and all regions belonging to its neighborhood. The intensity of spillover effects is captured by the parameter ρ, identical for all regions. Finally, $\beta_k's$ and $\theta_l's$ are unknown parameters to be estimated and ε_i is an error term assumed to be identically and independently distributed (*iid*).

In matrix form we have

$$\mathbf{y} \quad = \quad \mathbf{X}\beta + \theta \mathbf{W}_N \mathbf{Z} + \rho \mathbf{W}_N \mathbf{y} + \varepsilon \tag{2}$$

where $\mathbf{y}$ is a $(N \times 1)$ vector of the regional growth rates, $\mathbf{X}$ the $(N \times 7)$ matrix of the explanatory variables, including the constant term, the logarithms of per worker income, the logarithms of the investment rates in physical capital divided by the effective depreciation rate, the logarithms of the working-age population growth rates and the logarithms of expenditures in the research sector. $\mathbf{W}_N$ is a $(N \times N)$ spatial weights matrix, $\mathbf{W}_N\mathbf{Z}$ is the $(N \times 2)$ vector of the spatial lags of the logarithms of the per worker income and the logarithms of the investment rates in physical capital divided by the effective depreciation rate and $\mathbf{W}_N\mathbf{y}$ is the endogenous spatial lag term. β and θ are vectors of parameters associated to $\mathbf{X}$ and $\mathbf{W}_N\mathbf{Z}$, respectively. ρ is the spatial autoregressive parameter and ε is the $(N \times 1)$ vector of *iid* errors.

The reduced form of equation 2 can be easily derived:

$$\mathbf{y} \quad = \quad \mathbf{AX}\beta + \mathbf{A}\theta \mathbf{W}_N \mathbf{Z} + \mathbf{A}\varepsilon \tag{3}$$

where $\mathbf{A} = (\mathbf{I}_N - \rho \mathbf{W}_N)^{-1}$. The growth rate of per worker income in a location i is therefore influenced not only by its own characteristics (initial conditions, saving rate, employment growth rate and R&D intensity), but also by initial conditions, saving rates, employment growth rates in all other locations through the inverse spatial transformation $\mathbf{A}$, the so-called '*spatial multiplier effect*' (Anselin, 2004). Equation 3 also suggests that there are spatial externalities in un-modeled effects: a random shock (or disturbance) in a specific location i does not only affect the outcome in that region, but it has also an impact on the outcome in all other locations through ('*spatial diffusion process of random shocks*').

Equation 1 can be classified as a Spatial Durbin Model (SDM) specification of Howitt's Schumpeterian growth model. As already observed, Howitt (2000) assumes that each region diffuses the same amount of knowledge to other regions, so that the last term of equation 1 is identical to each region and can be incorporated in the constant term and also θ_1 and θ_2 are equal to zero. In other words, Howitt (2000) excludes the possibility of specific technological interdependence across regions. Moreover, if β_3 and β_4 are both equal to zero, equation 1 encompasses also the multi-region Solow growth model with imperfect technological interdependence developed by the same authors (Ertur and Koch, 2007), which can be classified as Spatial Durbin Model (SDM) specification of the Solow growth model. Indeed, in addition to factor accumulation, equation 1 shows that innovation caused by R&D investment plays a major role in explaining the growth process. Finally, also the Solow growth model constitutes a particular case of the multi-region Schumpeterian growth model when R&D expenditures have no effect on growth and there is no technological interdependence between regions.

2.2 Further misspecification biases: nonlinearities and unobserved heterogeneity

Equation 1 allows us to capture the effect of spatial interdependence. However, this equation entails a strong linearity assumption on the growth behavior. This assumption has been considered as particularly inappropriate for the analysis of complex heterogeneous regions. For example, it has been observed that regional growth behavior in the West and in the East European Union

may greatly differ (Ertur and Koch, 2007) and, more generally, that the evidence of regional *club convergence* (parameter heterogeneity or multiple regimes) is the rule rather than the exception in regional growth analysis (Ertur and Le Gallo, 2009).

As it is well known, club convergence can be generated by the original Solow-Swan model by simply assuming that either the saving rate or the population growth rate is a function of income per worker. Masanjala and Papageorgiou (2004) have also shown how replacing the commonly used Cobb-Douglas aggregated production specification with the more general Constant-Elasticity-of-Substitution specification can generate parameter heterogeneity in the Solow growth equation. Obviously, also the linear functional form imposed in equation 1, deriving from the assumption of an aggregate Cobb-Douglas production function, can be relaxed. Indeed, more flexible specifications of equation 1 are needed to assess the hypothesis of parameter heterogeneity and, more specifically, to identify possible threshold effects in the relationship between the savings rate or the R&D intensity and growth. More generally, we claim that the functional form remains always unknown and the linear form imposed in the traditional growth analysis may represent a source of mis-specification bias. Thus, a semiparametric framework which recognizes the uncertainty in the functional form is often recommended (see, e.g., Liu and Stengos, 1999; Durlauf et al., 2001; Fiaschi and Lavezzi, 2003; Kalaitzidakis et al, 2001; Basile, 2008, 2009; Basile et al., 2012).

Controlling for *unobserved heterogeneity* is another fundamental task in empirical growth analysis, as failing to do so can introduce omitted-variable biases and preclude causal inference. To complicate the analysis, spatial interdependence may simply be the consequence of (spatially correlated) omitted variables rather than being the result of spillovers. If this is the case, there are no compelling reasons for using traditional parametric models, like the SAR or SEM. As shown by McMillen (2012), a simple semiparametric model, with a smooth interaction between latitude and longitude (the so-called *Geoadditive Model*), can remove unobserved heterogeneity. However, the aim of our empirical study is to assess the impact of spillover effects (for example the global effect of a localized shock in R&D investment) rather than simply compensate for unobserved heterogeneity. Thus, we need to capture the effect of spatial spillover through the inclusion of spatial interaction terms, besides controlling for unobserved heterogeneity and functional form mis-specifications. This is a complex objective that the traditional parametric spatial econometric approach can hardly attain. In fact, several prominent scholars have recently called for a review of the methodological basis of the traditional spatial econometrics. McMillen (2010, 2012) points that there is a fundamental contradiction between the severity of the unknowns in the specification (functional form and spatially correlated omitted variables) and the overwhelming use of maximum likelihood methods (which heavily depend on the assumption of a correct specification). Pinkse and Slade (2010) recognize the intrinsic complexity of spatial data which suffer from so many problems (irregular distribution, varying density, aggregation, etc.) that precludes the use of naively parsimonious specifications, like the family of SAR models. Their advice can be summarized in avoiding overparameterized specifications and letting the application guide the theory; this has a clear parallel with the position of McMillen. According to Gibbons and Overman (2012), the dominant approach in spatial econometrics is not convincing because of the many, sometimes unjustified, hypotheses made about the functional form, the presence of omitted factors, the spatial weights, and so on. A further critical

issue raised by Gibbons and Overman (2012) concerns the use of lagged values of the regressors as instrument variables (IV) for the spatial lag of the endogenous variable in the SAR-type models. The arguments are somewhat familiar with Pinkse and Slade (2010): the first is the unconvincing exclusion of these terms (spatial lagged regressors) from the structural equation; the second is the unjustified claim of exogeneity for the X's variables in a typical spatial model (contrary, they are expected to be endogenous and correlated with the unobserved determinants to the endogenous variable). The last deficiency can be treated more efficiently by using, once again, less structured models.

Given these limitations, it is important to pay attention to other, more flexible, approaches, which help us to overcome part of the deficiencies encountered in the parametric framework. In particular, here we focus on *Penalized Spline Spatial Autoregressive Geoadditive Models* developed, among others, by Basile and Gress (2005), Su and Jin (2010), Su (2012), Basile, Capello, and Caragliu (2012) and Montero, Mínguez, and Durbán (2012). A review of these contributions and some more methodological advances in the literature are reported in Basile et al. (2014).

With respect to standard parametric spatial econometric models, these semiparametric approaches offer a more convenient way of addressing simultaneously the three problems mentioned above (substantive spatial dependence, unobserved heterogeneity and unknown functional form). The most general specification of a semiparametric growth model we propose in this paper is:

$$\begin{aligned}\gamma_i \; = \; & \beta_0 + f_1\left(\ln y_{0,i}\right) + f_2\left(\ln \frac{s_{k,i}}{n_i + g_w + \delta}\right) + f_3\left(\ln s_{A,i}\right) + f_4\left(\ln n_i\right) \\ & + m_1\left(\sum_{j \neq i}^{N} w_{ij} \ln y_{0,j}\right) + m_2\left(\sum_{j \neq i}^{N} w_{ij} \ln \frac{s_{k,j}}{n_j + g_w + \delta}\right) + h\left(no_i, e_i\right) + \rho \sum_{j \neq i}^{N} w_{ij} \gamma_j + \varepsilon_i \end{aligned} \tag{4}$$

$$\varepsilon_i \sim iid \mathcal{N}\left(0, \sigma_\varepsilon^2\right) \qquad i = 1, ..., n$$

This semiparametric framework allows us to relax the linearity assumption and, simultaneously, model spatial dependence and unobserved heterogeneity. Indeed, $f_1(.) - f_4(.)$ and $m_1(.) - m_2(.)$ are unknown smooth functions of univariate continuous covariates, capturing nonlinear effects of exogenous variables. The term $h\left(no_i, e_i\right)$ is a smooth spatial trend surface, i.e. a smooth interaction between latitude (*northing*) and longitude (*easting*). It allows us to control for unobserved spatial heterogeneity. Finally, ε_i are *iid* normally distributed random shocks. [1]

Model 5 reflects the notion of spatial dependence made of two parts: (*i*) a spatial trend due to unobserved regional characteristics, which is modeled by the smooth function of the coordinates, and (*ii*) global spatial spillover effects, which are modeled by including the spatial lag terms. This model can be termed as Penalized-spline Spatial Durbin Schumpeterian Model (PS-SDM-Schumpeterian). Excluding th two terms $m_1(.) - m_2(.)$, we get the so-called PS-SAR-Schumpeterian.

[1] This assumption can be relaxed by a more general specification, such as $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2 \Lambda)$ being Λ a covariance matrix reflecting cross-sectional dependence in the errors as, for example, in Pinheiro and Bates (2000)

Each univariate smooth term in 5 can be approximated by a linear combination of q_k known basis functions $b_{q_k}(x_k)$:

$$f_k(x_k) = \sum_{q_k} \beta_{q_k} b_{q_k}(x_k)$$

with β_{q_k} unknown parameters to be estimated. To reduce mis-specification bias, $q_k' s$ should be large enough, which results in a danger of over-fitting. The smoothness of the functions can be controlled by penalizing 'wiggly' functions when fitting the model. A measure of 'wiggliness', $J_k \equiv \beta_k' \mathbf{S}_k \beta_k$, is associated with each k smooth function, with $\mathbf{S}_k$ a positive semidefinite matrix. Equivalently, each univariate smooth spatial lag term can be approximated by $m_k(Wx_k) = \sum_{q_k} \delta_{q_k} b_{q_k}(Wx_k)$. The penalized spline base–learners can be extended to the two dimensions to handle interactions by using tensor products (Currie, Durbán, and Eilers, 2006). In this case, smooth bases are built up from products of 'marginal' bases functions. For example,

$$h(no, e) = \sum_{q_{no}} \sum_{q_e} \beta_{q_{no}, q_e} b_{q_{no}}(x_{no}) b_{q_e}(x_e)$$

Given the bases for each smooth term, equation (5) can be rewritten in matrix form as:

$$\begin{aligned} \mathbf{y} &= \rho \mathbf{W}\mathbf{y} + \Sigma_{q_1} \beta_{1q_1} b_{1q_1}(x_1) + \Sigma_{q_2} \beta_{2q_2} b_{2q_2}(x_2) + \ldots + \varepsilon \\ &= \rho \mathbf{W}\mathbf{y} + \mathbf{X}\beta + \mathbf{W}\mathbf{X}\delta + \varepsilon \end{aligned} \tag{5}$$

where matrix $\mathbf{X}$ ($\mathbf{WX}$) includes all the basis functions evaluated at the $x's$ ($Wx's$) covariate values, while β, δ and ρ represent the coefficients to be estimated.

As in the parametric SDM, also in the PS-SDM, the interpretation of econometric results must be based on the computation of direct, indirect (spillover) and total effects. The indirect (spillover) smooth effect of x_k ca be written as

$$\widehat{f}_k^I(x_k) = [\mathbf{I}_n - \widehat{\rho}\mathbf{W}_n]_{ij}^{-1} \left[x_k \widehat{\beta} + W x_k \widehat{\delta}_k \right] \tag{6}$$

The direct smooth effect can also be computed as:

$$\widehat{f}_k^D(x_k) = [\mathbf{I}_n - \widehat{\rho}\mathbf{W}_n]_{ii}^{-1} x_k \widehat{\beta}_k \tag{7}$$

Finally, the total smooth effect is:

$$\widehat{f}_k^T(x_k) \;=\; \widehat{f}_k^D(x_k) + \widehat{f}_k^I(x_k) \tag{8}$$

Similar expressions can be provided for the direct, indirect and total effects of the PS-SAR (see Basile et al., 2014).

2.3 Estimation and identification: a control function approach

As well known, any semiparametric model can be expressed as a mixed model and, thus, it is possible to estimate all the parameters (including the smoothing parameters) using restricted maximum likelihood methods (REML) (Wood, 2011). A complication with the PS-SAR and

the PS-SDM is given by the presence of the endogenous spatial lag term $\mathbf{W}_n\mathbf{y}$ on the right hand side. Basile et al. (2014) show how the REML methodology can be extended to estimate the parameters of PS-SAR and PS-SDM either in a single step or in a 2-step "control function" (CF) approach (Blundell and Powell, 2003).[2] In this study, we use the 2-step approach.

In the case of nonparametric and semiparametric additive models, the CF approach imposes extra identification assumptions – i.e. conditional mean restrictions $E(\mathbf{u}|\mathbf{Q}) = 0$ and $E(\mathbf{u}|\mathbf{X},\mathbf{Q}) = E(\mathbf{u}|\mathbf{X},\mathbf{v}) = E(\mathbf{u}|\mathbf{v})$ – not imposed in a standard IV approach. However, in the case of nonparametric additive models, the CF approach offers a critical advantage over the IV method (Wooldridge, 2010). In particular, the application of the standard 2-SLS method to nonparametric additive models (i.e. the substitution of the fitted values from the first-stage nonparametric regression of $\mathbf{X}$ on $\mathbf{Q}$, into nonlinear structural functions) generally yields inconsistent estimates of the structural parameters. Instead, alternative procedure involving the use of the residuals $\mathbf{v}$ from the first-stage regression to control for the endogeneity of the regressors $\mathbf{X}$ do yield identification of the ASF (Blundell and Powell, 2003).

Using the CF approach to estimate the PS-SAR model implies to run the following first–step semiparametric regression

$$\mathbf{W}_n\mathbf{y} \quad = \quad \beta_0 + \sum_m g_m(\mathbf{Q}) + \nu$$

where ν is a random vector satisfying conditional mean restrictions $E(\nu|\mathbf{Q}) = 0$ and $\mathbf{Q}$ is a set of m conformable instruments. The functions g_m define generic representations of different types of covariate effects, including both linear and nonparametric smooth components.

The residuals from the first step are then included in the original PS-SDM or PS-SAR equation to control for the endogeneity of $\mathbf{W}_n\mathbf{y}$.[3] Obviously, the endogeneity of any other continuously distributed regressor in the PS-SDM or in the PS-SAR model can also be addressed via the CF approach if valid instruments are available.[4] Since the second-step regression contains generated regressors (i.e. the first-step residuals), a bootstrap procedure is recommended to compute the p-values (see Basile et al., 2014).

In line with Kelejian and Prucha (1997), $\mathbf{Q}$ might contain all exogenous terms included in the model and some of their spatial lags. However, the use of the spatial lags of the X variables as instruments for $\mathbf{Wy}$ has been recently criticized by Gibbons and Overman (2012). They suggest, instead, to use alternative instruments. Following this suggestion, in order to identify the causal effect of $\mathbf{Wy}$ we use the spatial lag of a measure of the quality of regional Government developed by Charron et al. (2012) and the spatial lags of various measures of social capital. The use of

[2]Generally speaking, the CF approach is an alternative to standard instrumental variable (IV) methods (either two-stage-least squares – 2SLS or GMM). It is a two–step procedure: in the first step the endogenous explanatory variables ($\mathbf{X}$) are regressed on a set of instrumental variables ($\mathbf{Q}$); the residuals from the first step are then included in the original equation to "control" for the endogeneity bias. In linear models ($\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$), the CF approach relies on the same identification (orthogonality) conditions – i.e. unconditional moment restriction $E(\mathbf{Q}'u) = 0$ – as the IV methods and leads to the usual 2SLS estimator. The CF approach treats endogeneity as an omitted variable problem, where the inclusion of estimates of the first-stage errors $\mathbf{v}$ (the part of the regressors $\mathbf{X}$ that is correlated with $\mathbf{Q}$) as a covariate corrects the inconsistency of least-squares regression of $\mathbf{y}$ on $\mathbf{X}$.

[3]Both first and second step equations can be estimated by using the REML estimator.

[4]The requirement that the endogenous regressor be continuously distributed is the most important limitation for the applicability of the CF approach in this context.

these variables as instruments can be justified on the basis of the argument that the quality of the institutional arrangement of its neighbors may affect the growth performance (**y**) of region i only through the **Wy**.

3 Econometric analysis

Using regional data for the European regions, in this section we investigate the performance of the *semiparametric geoadditive Schumepterian growth model* presented in equation 17 and compare it to several more restricted specifications. In Section 3.1 we describe the dataset, the variables included in the model, and the choice of the spatial weights matrix. In Section 3.2 we report the results of the model selection. Finally, in Section 3.3 we report the results of a semiparametric geoadditive Schumepterian growth model.

3.1 Data, descriptive analysis and the spatial weights matrix

Using Cambridge Econometrics data and a sample of 248 NUTS2 regions for the 1991-2011 period, we compute the income per worker (y_{it}) of region i (with $i = 1,...,N$) at time t (with $t = 0,...,T$) as the ratio between gross value added at constant prices 2000 and total employment. Income levels are normalized with respect to the EU-27 average in order to remove co-movements due to the European wide business cycle and trends in the average values. Average annual productivity growth rates are then computed as $\gamma_y = \frac{\ln y_T - \ln y_0}{T}$.

To give an idea of the existing scale of regional imbalances in Europe, we report some evidence on the distribution dynamics of regional income per worker using the continuous state-space intra-distribution dynamics (IDD) approach developed by Quah (1997), based on the estimation of conditional density functions and ergodic distributions (see Fotopoulos, 2008; Fiaschi and Lavezzi, 2007; Basile, 2009, for previous applications of this approach to European regional data).[5] To estimate the conditional density function we choose $\tau = 20$, so that y is the vector of labor productivity levels in 2011 (the last year) and x the vector of productivity levels in 1991 (the first year). The function is estimated using a local linear density estimator with variable bandwidth (see Basile, 2010, for a thorough discussion of conditional density estimators applied for the IDD analysis). Finally, we compute the ergodic distribution of regional income per worker using the transition matrix extracted from the estimated conditional density function. The shape of this ergodic distribution suggests the existence of convergence clubs: different groups of regions tend to converge to different long-run parallel growth paths (Figure 1). Specifically,

[5] Given the distribution of regional income per worker (or labor productivity) at time t and its associated probability measure, ϕ_t, the IDD approach consists of describing the law of motion of the stochastic process, $\{\phi_t, t \geq 0\}$. If this process is assumed to be first-order Markov, than the law of motion for $\{\phi_t, t \geq 0\}$ can be modeled as an autoregressive process: $\phi_{t+\tau}(y) = \int_0^\infty f_\tau(y|x)\phi_t(x)dx$, where $f_\tau(y|x)$ is the expected density of y (the productivity levels at time $t+\tau$) conditional upon x (the productivity levels at time t). In other words, the conditional density $f_\tau(y|x)$ describes the probability that a region moves to a certain state of relative productivity given that it has a certain relative productivity level in the initial period. If the transition density function is time-invariant, then the erogodic distribution can be computed as: $\phi_\infty(y) = \int_0^\infty f_\tau(y|x)\phi_\infty(x)dx$ (Johnson, 2005). This function describes the long-term behavior of the productivity distribution: it is the density of what the cross-region productivity distribution tends towards, should the system continue along its historical path (Quah, 2007).

a main peak at a level just above the EU average indicates that a large group of regions tend to converge to that level, while a group of regions appears to be entrapped in low productivity levels (some of them at about 0.35 times the EU average productivity level; some others at about 0.6 times the EU average productivity level).

Figure 1

We also provide information on the spatial distribution of income per worker over the sample period. Using nonparametric regression spline methods, we regress labor productivity levels in 1991 and in 2011 as well as the growth rates of labor productivity on the smooth interaction between latitude and longitude (or *northing* and *easting*), $y = \alpha + f(no, e) + \varepsilon$. Figures 2a-2c plot the geographical components (the so-called spatial trend surface) of this model, showing a Core-Periphery pattern in the spatial distribution of productivity levels and growth rates. Specifically, higher incomes per worker are clustered in the Center of the Continent, while lower incomes are concentrated in two peripheral areas: the first one includes Portuguese and Greek regions, while the lowest income levels are clustered in Eastern regions (Poland, Hungary, Slovakia, Romania, Lithuania and Czech Republic). The latter, however, show the highest growth rates.

Figure 2

The variables included in the r.h.s. of the growth model are the initial conditions (log of per worker income in 1991, $\ln y_0$), the ratio between the investment rate in physical capital and the net depreciation rate $\left(\ln \frac{s_k}{n+0.05}\right)$, the R&D investment rate ($\ln s_A$), computed as the ratio between R&D investments and total GDP, the employment growth rate ($\ln n$). Figures 2d-f provide evidence of the spatial distribution of these variables.

Some specification includes also the spatial lags of initial conditions ($W \ln y_0$), of the net investment rate $\left(W \ln \frac{s_k}{n+g_w+\delta}\right)$, and of the growth rate ($W\gamma$). The spatial weights matrix (W) used in the present study is a combination of a binary spatial weights matrix based on the cut-off criterion and the negative exponential distance function ($exp(-d_{ij})$):

$$w_{ij} = \begin{cases} exp(-d_{ij}) & if \quad 100km < d_{ij} < 700km \\ 0 & if \quad d_{ij} < 100km \quad or \quad if \quad d_{ij} > 700km \end{cases}$$

3.2 Model selection

We use the data and the variables described above to compare the performance of different competing parametric and semiparametric models. The most restricted specification is the *parametric* linear Solow model without spatial effects, relating the regional productivity growth rate (γ_i) of region i to the logarithm of the initial conditions ($\ln y_{0,i}$) and the log of investment rate in physical capital divided by the effective depreciation rate of region i $\left(\ln \frac{s_{k,i}}{n_i+0.05}\right)$ (see equation 9 in the appendix). This model does not contain a spatial trend and is based on the assumption of spatial independence of the residuals. OLS coefficients of this model are all significant and have

the expected sign (Table 1): regional growth rates are positively affected by effective physical capital accumulation rates and negatively influenced initial per-worker income levels. Thus, the conditional convergence hypothesis cannot be rejected, but the speed of convergence (equal to 1.066%) is rather slow and the corresponding half life is 65 years. (See the appendix for a list of all parametric and semiparametric models estimated.)

Table 1 about here

The second parametric a-spatial model is the Howitt (2000) Schumpeterian linear model including the log of the R&D investment rate ($\ln s_{A,i}$) and the log of the employment growth rate ($\ln n_i$) as further determinants of the productivity growth rate (see equation 10 in the appendix). OLS results confirm the predictions of the Schumpeterian Howitt's model, with both R&D investments and employment growth rates positively affecting productivity growth. The implied speed of convergence of this model (1.519%) is higher than the Solow's one, and the corresponding half life is 46 years. The introduction of R&D investments and employment growth rates leads to a reduction of Akaike information criterion (AIC) and Bayesian information criterion (BIC) statistics (see Table 2).

Table 2 about here

These two parametric models are compared to their nonparametric counterparts: the *nonlinear a-spatial neoclassical model* (equation 11) and the *nonlinear a-spatial Schumpeterian model* (equation 12). These two specifications include nonparametric smooth additive terms and are estimated by *REML*. We firstly observe that the a-spatial nonlinear models outperform (in terms of AIC and BIC) their parametric counterparts, indicating that the functional form imposed in the parametric models is too restrictive. This evidence is reinforced by the F tests reported in Table 3, where we compare the a-spatial models with models imposing a restricted functional form for each term. In the case of the a-spatial neoclassical model, the F tests reject the restricted specifications for both terms ($\ln y_0$) and $\left(\ln \frac{s_k}{n_i + 0.05}\right)$. Specifically, an inverted U-shaped relationship between growth and initial conditions emerges (Figure 3). Thus, it seems that a diverging behavior characterizes the group of Eastern regions, while Western regions maintain a conditional predicted convergence path. The assumption of identical speed of convergence is therefore rejected for the neoclassical growth model. Nonlinearities in the effects of gross physical investment are also clearly detected (Figure 3). Specifically, an increase in the saving rate is associated with an increase in growth rates only when the net investment rate is above the EU average. These results are in line with Azariadis and Drazen's (1990) theoretical prediction. However, in the case of the a-spatial Schumepterian model the linearity assumption is strongly rejected only for $\left(\ln \frac{s_k}{n_i + 0.05}\right)$ and ($\ln n$). In particular, the hump-shaped relationship between growth and initial conditions disappear with a Schumpeterian specification of the model.

Table 3 about here

Significant gains in model performance (AIC and BIC values) are obtained once spatial lag terms and a geoadditive component are included in the model, thus highlighting the importance of controlling for spatial dependence and unobserved spatial heterogeneity (Table 2). Nevertheless, using the BIC criterion, it emerges that the best specification is the semiparametric Schumpeterian SAR geoadditive model rather than the semiparametric Schumpeterian SDM geoadditive model. Since the BIC penalizes more the degrees of freedom, we conclude that the geoadditive-SDM is over-specified. It is also worth noticing that the estimate of the ρ parameter strongly drops when a spatial trend is included in the model (Table 2). This is an expected result, since part of the spatial dependence is now captured by the spatial trend surface.

Following Augustin, Musio, Wilpert, Kublin, Wood, and Schumacher (2009), we check for the gross violation of distributional assumptions of the residuals of the different models by using the empirical semi-variogram of the residuals. For the estimation of the semivariogram, we use the `variogram` function of the `geoR` package in `R`. We calculate the empirical semivariogram of the observed residuals. Then, these residuals are permuted 99 times and envelopes are computed by taking, at each spatial lag, the maximum and minimum values of the semi-variograms for the permuted residuals. Figure 4 shows that the spatial dependence, underlying the a-spatial linear and nonlinear models, is well captured by both parametric and semiparametric spatial econometric models. In fact, the semivariograms for the residuals of the a-spatial linear and nonlinear models are Gaussian-type and do not stabilize even for large distances. However, once spatial econometric models are used, the experimental semivariograms turn in pure nugget semivariogram (the theoretical semivariogram representing the situation of absence of spatial correlation), which means that all these specifications are able to account for the spatial correlation. Table 2 summarizes these results.

Figure ?? about here

Finally, we check for the presence of a spatial trend in the residuals. To this scope, we regress the residuals of each model on the smooth interaction between latitude and longitude by using a P-spline method. The results reported in Table 2 show that, while parametric spatial econometric models capture spatial autocorrelation, they do not capture the residual spatial trend. On the contrary, semiparametric spatial autoregressive models are able to capture both spatial dependence and spatial trends. All in all, our results clearly display the superiority of semiparametric Schumpeterian SAR geoadditive model estimated using the 2-stage control function approach. Thus, in the next section we focus on the results of this model and report the relative evidence on direct and indirect (spillover) effects.

3.3 Evidence from the Schumpeterian PS-SAR model

The estimation results of the Schumpeterian geoadditive SAR model appear in (Table 4). First, we run a semiparametric regression of the endogenous term $\mathbf{W}_n\mathbf{y}$ on the exogenous variables and the external instruments.[6] Then, we insert the first-stage residuals in the original semiparametric

[6]Unfortunately there is not a well-known and widely accepted test for the validity of the conditional mean restrictions imposed by the CF approach. A practically feasible way of testing such restrictions consists of including some of the excluded instruments in the control function (CF) estimate and check for the significance of their coefficients.

regression to correct the inconsistency due to the endogeneity problem. In line with the results of linearity tests, only physical capital investments and employment growth rates are introduced as smooth terms. The model also includes a spatial trend surface, $h(no,e)$, constructed by using the spatial coordinates in re-scaled form. All smooth terms are specified using P-spline basis functions. Both stages are estimated using the REML method. Second-stage results show that all smooth terms have an edf higher than 1, confirming that they enter nonlinearly the model.

Table 4 about here

The average marginal total effect of initial conditions ($\ln y_0$) is -1.342, while the average marginal total effect of ($\ln s_A$) is 0.361. Since the model is specified in logged levels, we can interpret the impact estimates as elasticities. Thus, we would conclude that a 1% increase in $\ln y_0$ ($\ln S_A$) would result in a 1.342% decrease (0.361% increase) in growth . Around 7/10 of this impact comes from the direct effect magnitude of -0.991 (0.266), and 3/10 from the indirect or spatial spillover impact based on its scalar impact estimate of 0.351 (0.094).

Figure 5 reports the plots of total, direct and indirect smooth effects for $\ln \frac{s_k}{n+g_w+\delta}$ and $\ln n$ computed using equations (7), (6) and (8).[7] The point-wise 95% confidence bands (obtained using the bootstrap procedure described in Basile et al., 2014) show that all effects are also significant in most part of the variable domain. Nonlinearities in the effects of physical capital investments are also clearly detected. Specifically, an increase in $\ln \frac{s_k}{n+g_w+\delta}$ is associated with an increase in growth rates only when $\ln \frac{s_k}{n+g_w+\delta}$ is above the EU average. These results are in line with Azariadis and Drazen's (1990) theoretical prediction. The influence of the employment growth rate on regional growth is positive, although the effect is not homogeneous across the sample. As expected, indirect effects are always lower that the direct effects.

Figure 5 about here

Finally, a picture of the estimated spatial trend surface — $h(no,e)$ — is reported in Figure **??**. It emerges quite clearly that, even after having controlled for the effect of the characteristics of the regions (in terms of initial conditions, physical capital investment rate, employment growth rates and R&D investments and for the spillover effects (through the spatial lag term), there remain significant growth differences over space. In particular, the growth rates are significantly higher in North-Eastern and North-Western regions (mainly the UK and Ireland).

Figure ?? about here

These coefficients should not be significant because the CF should pick up all of the correlation between the structural error term and $(\mathbf{W}_n\mathbf{y},\mathbf{Z})$, where $(\mathbf{Z}=(\mathbf{X},\mathbf{W}_n\mathbf{X}))$. In particular, if $\mathbf{y}=\rho\mathbf{W}_n\mathbf{y}+f(\mathbf{X})+u_1$ and $\mathbf{W}_n\mathbf{y}=\mathbf{X}\delta+v_2$, then $\mathbf{W}_n\mathbf{X}$ can be added to the CF estimation. This means that we can regress $\mathbf{y}$ on $\mathbf{W}_n\mathbf{y},\mathbf{X},\hat{v}_2$ and $\mathbf{W}_n\mathbf{X}$, using whatever model/estimation method, and test coefficients on $\mathbf{W}_n\mathbf{X}$. In our case, all the external instruments turned out to be strictly exogenous.

[7]Actually, total, direct and indirect effects are not smooth over the domain of variable x_k due to the presence of the spatial multiplier matrix in the algorithms. A wiggly profile of direct, indirect and total effects would appear even if the model were linear. Therefore, in the spirit of this paper, we have applied a spline smoother to obtain smooth curves.

4 Conclusions

References

AUGUSTIN, N., M. MUSIO, K. V. WILPERT, E. KUBLIN, S. WOOD, AND M. SCHUMACHER (2009): "Modeling Spatio-Temporal Forest Health Monitoring Data," *Journal of the American Statistical Association*, 104, 899–911.

AZOMAHOU, T., J. E. OUARDIGHI, P. NGUYEN-VAN, AND T. PHAM (2011): "Testing Convergence of European Regions: A semiparametric approach," *Economic Modelling*, 28, 1202–1210.

BASILE, R. (2008): "Regional Economic Growth in Europe: a Semiparametric Spatial Dependence Approach," *Papers in Regional Science*, 87, 527–544.

——— (2009): "Productivity Polarization across Regions in Europe: The Role of Nonlinearities and Spatial Dependence," *International Regional Science Review*, 32, 92–115.

BASILE, R., R. CAPELLO, AND A. CARAGLIU (2012): "Technological Interdependence and Regional Growth in Europe," *Papers in Regional Science*, 91, 697–722.

BASILE, R., AND B. GRESS (2005): "Semi-parametric Spatial Auto-covariance Models of Regional Growth Behavior in Europe," *Region et Developement*, 21, 93–118.

BLUNDELL, R., AND J. POWELL (2003): *Advances in Economics and Econometrics Theory and Application*chap. Endogeneity in Nonparametric and Semiparametric Regression Models. Cambridge University Press.

CURRIE, I., M. DURBÁN, AND P. EILERS (2006): "Generalized Array Models with Application to Multidimensional Smoothing," *Journal of the Royal Statistical Society B*, 68, 259–280.

EILERS, P., AND B. MARX (1996): "Flexible Smoothing with B-Splines and Penalties," *Statistical Science*, 11, 89–121.

ERTUR, C., AND J. L. GALLO (2009): *Handbook of Regional Growth and Development Theories*chap. Regional Growth and Convergence: Heterogenous Reaction versus Interaction Spatial Econometric Approaches, pp. 374–388. Edward Elgar, Cheltenham.

ERTUR, C., AND W. KOCH (2011): "A Contribution to the Schumpeterian Growth Theory and Empirics," *Journal of Economic Growth*, 16, 215–255.

FOTOPOULOS, G. (2012): "Nonlinearities in Regional Economic Growth and Convergence: the Role of Entrepreneurship in the European Union Regions," *The Annals of Regional Science*, 48, 719–741.

GIBBONS, S., AND H. G. OVERMAN (2012): "MOSTLY POINTLESS SPATIAL ECONOMETRICS?*," *Journal of Regional Science*, 52(2), 172–191.

KELEJIAN, H., AND I. PRUCHA (1997): "Estimation of Spatial Regression Models with Autoregressive Errors by Two-Stage Least Squares Procedures: A Serious Problem," *International Regional Science Review*, 20, 103–111.

McMillen, D. P. (2010): "Issues in Spatial Data Analisis," *Journal of Regional Science*, 50, 119–141.

——— (2012): "Perspectives on spatial econometrics: Linear smoothing with structured models," *Journal of Regional Science*, 52(2), 192–209.

Montero, J., R. Mínguez, and M. Durbán (2012): "SAR Models with Nonparametric Spatial Trends. A P-Spline Approach," *Estadística Española*, 54, 89–111.

Pinheiro, J., and D. Bates (2000): *Mixed-effects Models in S and S-PLUS*. Springer-Verlag, New York.

Pinkse, J., and M. Slade (2010): "The Future of Spatial Econometrics," *Journal of Regional Science*, 50, 103–117.

Su, L. (2012): "Semiparametric GMM Estimation of Spatial Autoregressive Models," *Journal of Econometrics*, 167, 543–560.

Su, L., and S. Jin (2010): "Profile Quasi-Maximum Likelihood Estimation of Partially Linear Spatial Autoregressive Models," *Journal of Econometrics*, 157, 18–33.

Wood, S. (2011): "Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 73, 3–36.

Wooldridge, J. M. (2010): "Econometric Analysis of Cross Section and Panel Data, Vol. 1 of MIT Press Books," .

Appendix: models specification

1. Linear a-spatial neoclassical (Solow) model:

$$\gamma_i = \beta_0 + \beta_1 \ln y_{0,i} + \beta_2 \ln \frac{s_{k,i}}{n_i + g + \delta} + \varepsilon_{S,i} \tag{9}$$

$$\varepsilon_{S,i} \sim iid\mathcal{N}\left(0, \sigma^2_{\varepsilon_S}\right) \qquad i = 1, ..., n$$

2. Linear a-spatial Schumpeterian (Howitt, 2000) model:

$$\gamma_i = \beta_0 + \beta_1 \ln y_{0,i} + \beta_2 \ln \frac{s_{k,i}}{n_i + g_w + \delta} + \beta_3 \ln s_{A,i} + \beta_4 \ln n_i + \varepsilon_{H,i} \tag{10}$$

$$\varepsilon_{H,i} \sim iid\mathcal{N}\left(0, \sigma^2_{\varepsilon_H}\right) \qquad i = 1, ..., n$$

3. Nonlinear a-spatial neoclassical model:

$$\gamma_i = \beta_0 + f_1\left(\ln y_{0,i}\right) + f_2\left(\ln \frac{s_{k,i}}{n_i + g + \delta}\right) + \varepsilon_{NLS,i} \tag{11}$$

$$\varepsilon_{NLS,i} \sim iid\mathcal{N}\left(0, \sigma^2_{\varepsilon_{NLS}}\right) \qquad i = 1, ..., n$$

4. Nonlinear a-spatial Schumpeterian model:

$$\gamma_i = \beta_0 + f_1\left(\ln y_{0,i}\right) + f_2\left(\ln \frac{s_{k,i}}{n_i + g_w + \delta}\right) + f_3\left(\ln s_{A,i}\right) + f_4\left(\ln n_i\right) + \varepsilon_{NLH,i} \tag{12}$$

$$\varepsilon_{NLH,i} \sim iid\mathcal{N}\left(0, \sigma^2_{\varepsilon_{NLH}}\right) \qquad i = 1, ..., n$$

5. Linear Spatial Durbin neoclassical (Erthur and Koch, 2007) model:

$$\begin{aligned} \gamma_i = {} & \beta_0 + \beta_1 \ln y_{0,i} + \beta_2 \ln \frac{s_{k,i}}{n_i + g_w + \delta} \\ & + \theta_1 \sum_{j \neq i}^{N} w_{ij} \ln y_{0,j} + \theta_2 \sum_{j \neq i}^{N} w_{ij} \ln \frac{s_{k,j}}{n_j + g_w + \delta} + \rho \sum_{j \neq i}^{N} w_{ij} \gamma_j + \varepsilon_{EK07,i} \end{aligned} \tag{13}$$

$$\varepsilon_{EK07,i} \sim iid\mathcal{N}\left(0, \sigma^2_{\varepsilon_{EK07}}\right) \qquad i = 1, ..., n$$

6. Linear Spatial Durbin Schumpeterian (Erthur and Koch, 2011) model:

$$\begin{aligned} \gamma_i = {} & \beta_0 + \beta_1 \ln y_{0,i} + \beta_2 \ln \frac{s_{k,i}}{n_i + g_w + \delta} + \beta_3 \ln s_{A,i} + \beta_4 \ln n_i \\ & + \theta_1 \sum_{j \neq i}^{N} w_{ij} \ln y_{0,j} + \theta_2 \sum_{j \neq i}^{N} w_{ij} \ln \frac{s_{k,j}}{n_j + g_w + \delta} + \rho \sum_{j \neq i}^{N} w_{ij} \gamma_j + \varepsilon_{EK11,i} \end{aligned} \tag{14}$$

$$\varepsilon_{EK11,i} \sim iid\mathcal{N}\left(0, \sigma^2_{\varepsilon_{EK11}}\right) \qquad i = 1, ..., n$$

7. Nonlinear SDM neoclassical model:

$$\begin{aligned}\gamma_i &= \beta_0 + f_1\left(\ln y_{0,i}\right) + f_2\left(\ln \frac{s_{k,i}}{n_i + g_w + \delta}\right) + \\ & +m_1\left(\sum_{j\neq i}^{N} w_{ij}\ln y_{0,j}\right) + m_2\left(\sum_{j\neq i}^{N} w_{ij}\ln \frac{s_{k,j}}{n_j + g_w + \delta}\right) + \rho\sum_{j\neq i}^{N} w_{ij}\gamma_j + \varepsilon_{NLEK07,i}\end{aligned} \tag{15}$$

$$\varepsilon_{NLEK07,i} \sim iid\mathscr{N}\left(0, \sigma^2_{\varepsilon_{NLEK07}}\right) \qquad i = 1,...,n$$

8. Nonlinear SDM Schumpeterian model:

$$\begin{aligned}\gamma_i &= \beta_0 + f_1\left(\ln y_{0,i}\right) + f_2\left(\ln \frac{s_{k,i}}{n_i + g_w + \delta}\right) + f_3\left(\ln s_{A,i}\right) + f_4\left(\ln n_i\right) \\ & +m_1\left(\sum_{j\neq i}^{N} w_{ij}\ln y_{0,j}\right) + m_2\left(\sum_{j\neq i}^{N} w_{ij}\ln \frac{s_{k,j}}{n_j + g_w + \delta}\right) + \rho\sum_{j\neq i}^{N} w_{ij}\gamma_j + \varepsilon_{NLEK11,i}\end{aligned} \tag{16}$$

$$\varepsilon_{NLEK11,i} \sim iid\mathscr{N}\left(0, \sigma^2_{\varepsilon_{NLEK11}}\right) \qquad i = 1,...,n$$

9. Geoadditive SDM neoclassical model:

$$\begin{aligned}\gamma_i &= \beta_0 + f_1\left(\ln y_{0,i}\right) + f_2\left(\ln \frac{s_{k,i}}{n_i + g_w + \delta}\right) + \\ & +m_1\left(\sum_{j\neq i}^{N} w_{ij}\ln y_{0,j}\right) + m_2\left(\sum_{j\neq i}^{N} w_{ij}\ln \frac{s_{k,j}}{n_j + g_w + \delta}\right) + h\left(no_i, e_i\right) + \rho\sum_{j\neq i}^{N} w_{ij}\gamma_j + \varepsilon_{GeoNLEK07,i}\end{aligned} \tag{17}$$

$$\varepsilon_{GeoNLEK07,i} \sim iid\mathscr{N}\left(0, \sigma^2_{\varepsilon_{GeoNLEK07}}\right) \qquad i = 1,...,n$$

10. Geoadditive SDM Schumpeterian model:

$$\begin{aligned}\gamma_i &= \beta_0 + f_1\left(\ln y_{0,i}\right) + f_2\left(\ln \frac{s_{k,i}}{n_i + g_w + \delta}\right) + f_3\left(\ln s_{A,i}\right) + f_4\left(\ln n_i\right) \\ & +m_1\left(\sum_{j\neq i}^{N} w_{ij}\ln y_{0,j}\right) + m_2\left(\sum_{j\neq i}^{N} w_{ij}\ln \frac{s_{k,j}}{n_j + g_w + \delta}\right) + h\left(no_i, e_i\right) + \rho\sum_{j\neq i}^{N} w_{ij}\gamma_j + \varepsilon_{GeoNLEK11,i}\end{aligned} \tag{18}$$

$$\varepsilon_{GeoNLEK11,i} \sim iid\mathscr{N}\left(0, \sigma^2_{\varepsilon_{GeoNLEK11}}\right) \qquad i = 1,...,n$$

11. Geoadditive SAR neoclassical model:

$$\begin{aligned}\gamma_i &= \beta_0 + f_1\left(\ln y_{0,i}\right) + f_2\left(\ln \frac{s_{k,i}}{n_i + g_w + \delta}\right) + \\ & h\left(no_i, e_i\right) + \rho\sum_{j\neq i}^{N} w_{ij}\gamma_j + \varepsilon_{GeoNLEK07,i}\end{aligned} \tag{19}$$

$$\varepsilon_{GeoNLEK07,i} \sim iid\mathscr{N}\left(0, \sigma^2_{\varepsilon_{GeoNLEK07}}\right) \qquad i = 1,...,n$$

12. Geoadditive SAR Schumpeterian model:

$$\begin{aligned}\gamma_i &= \beta_0 + f_1\left(\ln y_{0,i}\right) + f_2\left(\ln \frac{s_{k,i}}{n_i + g_w + \delta}\right) + f_3\left(\ln s_{A,i}\right) + f_4\left(\ln n_i\right) \\ & +h\left(no_i, e_i\right) + \rho\sum_{j\neq i}^{N} w_{ij}\gamma_j + \varepsilon_{GeoNLEK11,i}\end{aligned} \tag{20}$$

$$\varepsilon_{GeoNLEK11,i} \sim iid\mathscr{N}\left(0, \sigma^2_{\varepsilon_{GeoNLEK11}}\right) \qquad i = 1,...,n$$

TABLE 1
Results linear models

Variable	Solow (1956)	Howitt (2000)	Ertur-Koch (2007)	Ertur-Koch (2011)
Intercept	-0.099	-2.362	-0.059	-1.236
	0.082	0.000	0.268	0.001
$\ln y_0$	-0.960	-1.310	-1.015	-1.130
	0.000	0.000	0.000	0.000
$\ln \frac{s_k}{n+g_w+\delta}$	0.141	0.660	0.166	0.655
	0.075	0.005	0.032	0.003
$\ln s_A$		0.510		0.261
		0.000		0.001
$\ln n$		0.538		0.488
		0.017		0.021
$W \ln y_0$			0.936	0.706
			0.000	0.001
$W \ln \frac{s_k}{n+g_w+\delta}$			-0.167	-0.187
			0.028	0.008
$W\gamma$			0.882	0.736
			0.000	0.000
Diagnostic tests				
Weak instruments			11.189	8.378
			0.000	0.000
Wu-Hausman			13.786	7.555
			0.000	0.006
Sargan			12.848	9.621
			0.117	0.293

TABLE 2
Model comparison

Model	AIC	BIC	Spatial trend	Spatial dependence	Rho
Neoclassical a-spatial linear	5.188	5.230	Trend	Yes	
Schumpeterian a-spatial linear	4.994	5.065	Trend	Yes	
Neoclassical a-spatial nonlinear	5.119	5.210	Trend	Yes	
Schumpeterian a-spatial nonlinear	4.920	5.061	Trend	Yes	
Neoclassical SDM linear	5.004	5.089	Trend	No	0.882
Schumpeterian SDM linear	4.858	4.971	Trend	No	0.736
Neoclassical SDM nonlinear	4.753	4.925	Trend	No	0.809
Schumpeterian SDM nonlinear	4.628	4.841	Trend	No	0.707
Neoclassical SDM Geoadditive	4.597	4.942	No trend	No	0.463
Schumpeterian SDM Geoadditive	4.387	4.817	No trend	No	0.324
Neoclassical SAR Geoadditive	4.642	4.867	No trend	No	0.450
Schumpeterian SAR Geoadditive	4.410	4.767	No trend	No	0.303

Notes: The parametric a-spatial linear models are estimated by Ordinary Least Squares (*OLS*). Parametric spatial linear models are estimated through Instrumental Variable methods (*IV*), using ***policy*** as instruments. All semiparametric and geoadditive models are estimated by Restricted Maximum Likelihood (*REML*). A Control Function (CF) approach is applied in the case of spatial nonlinear models and Geoadditive models to control for the endogeneity of the spatial lag of the dependent variable, using the REML estimator for the estimation of the smoothing parameters in each step.

FIGURE 1
Regional distribution of per worker income

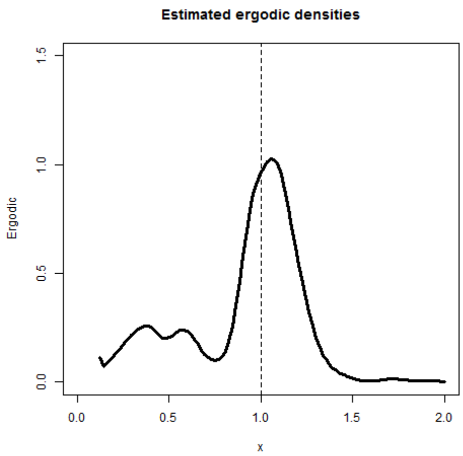

TABLE 3
F Tests for the smooth terms of the a-spatial nonlinear models

Smooth terms	Restricted form	F	*p-value*
A-spatial nonlinear Solow			
$f_1(\ln y_0)$	lny_0	5.630	0.004
$f_2\left(\ln \frac{s_k}{n+g_w+\delta}\right)$	$\ln \frac{s_k}{n+g_w+\delta}$	5.306	0.005
A-spatial nonlinear Howitt			
$f_1(\ln y_0)$	lny_0	3.719	0.055
$f_2\left(\ln \frac{s_k}{n+g_w+\delta}\right)$	$\ln \frac{s_k}{n+g_w+\delta}$	9.601	0.000
$f_3(\ln s_A)$	$\ln s_A$	2.737	0.099
$f_4(\ln n)$	$\ln n$	13.065	0.000
Schumpeterian SDM nonlinear			
$f_1(\ln y_0)$	lny_0	2.056	0.153
$m_1(W \ln y_0)$	$W lny_0$	1.663	0.199
$f_2\left(\ln \frac{s_k}{n+g_w+\delta}\right)$	$\ln \frac{s_k}{n+g_w+\delta}$	11.649	0.000
$m_2\left(W \ln \frac{s_k}{n+g_w+\delta}\right)$	$W \ln \frac{s_k}{n+g_w+\delta}$	10.794	0.001
$f_3(\ln s_A)$	$\ln s_A$	6.727	0.010
$f_4(\ln n)$	$\ln n$	5.876	0.016
Schumpeterian SDM Geoadditive			
$f_1(\ln y_0)$	lny_0	0.215	0.643
$m_1(W \ln y_0)$	$W lny_0$	1.092	0.338
$f_2\left(\ln \frac{s_k}{n+g_w+\delta}\right)$	$\ln \frac{s_k}{n+g_w+\delta}$	5.078	0.000
$m_2\left(W \ln \frac{s_k}{n+g_w+\delta}\right)$	$W \ln \frac{s_k}{n+g_w+\delta}$	0.462	0.497
$f_3(\ln s_A)$	$\ln s_A$	3.743	0.054
$f_4(\ln n)$	$\ln n$	0.374	0.541
Schumpeterian SAR Geoadditive			
$f_1(\ln y_0)$	lny_0	0.088	0.767
$f_2\left(\ln \frac{s_k}{n+g_w+\delta}\right)$	$\ln \frac{s_k}{n+g_w+\delta}$	10.281	0.000
$f_3(\ln s_A)$	$\ln s_A$	2.906	0.090
$f_4(\ln n)$	$\ln n$	5.410	0.005

TABLE 4

Control function estimates of the Schumpeterian Geoadditive SAR Model

	First stage	Second stage
Parametric terms	*Estimate (Bootstrap p-value)*	
(Intercept)	0.0686	-1.212
	(0.052)	(0.000)
$W\gamma$		0.298
		(0.000)
$\ln y_0$		-0.941
		(0.000)
$\ln s_A$		0.253
		(0.000)
Smooth terms	*edf*	*edf*
$f_1(\ln y_0)$	1.000	
$f_2\left(\ln \dfrac{s_k}{n+g_w+\delta}\right)$	1.455	3.020
$f_3(\ln s_A)$	1.000	
$f_2(\ln n)$	1.000	2.542
$h(no, e)$	14.674	13.720
$l(res(first.step))$		1.000
$g(W(gov))$	2.505	
$s_1(W(SC_1))$	2.500	
$s_2(W(SC_2))$	2.965	
$s_3(W(SC_3))$	3.075	
$s_4(W(SC_4))$	1.814	
$s_5(W(SC_5))$	1.854	
$s_6(W(SC_6))$	2.968	
$s_7(W(SC_7))$	3.629	
$s_8(W(SC_8))$	2.222	
Average direct effect of $\ln y_0$		-0.991
		(0.000)
Average indirect effect of $\ln y_0$		-0.351
		(0.000)
Average total effect of $\ln y_0$		-1.342
		(0.000)
Average direct effect of $\ln s_A$		0.266
		(0.000)
Average indirect effect of $\ln s_A$		0.094
		(0.000)
Average total effect of $\ln s_A$		0.361
		(0.000)

Notes: Bootstrap p-values for the significance of the parametric coefficients are reported in parenthesis. Smooth terms are specified using P-spline basis functions. Smoothing parameters are estimated using the REML.

FIGURE 2
Regional distribution of growth rates and their determinants

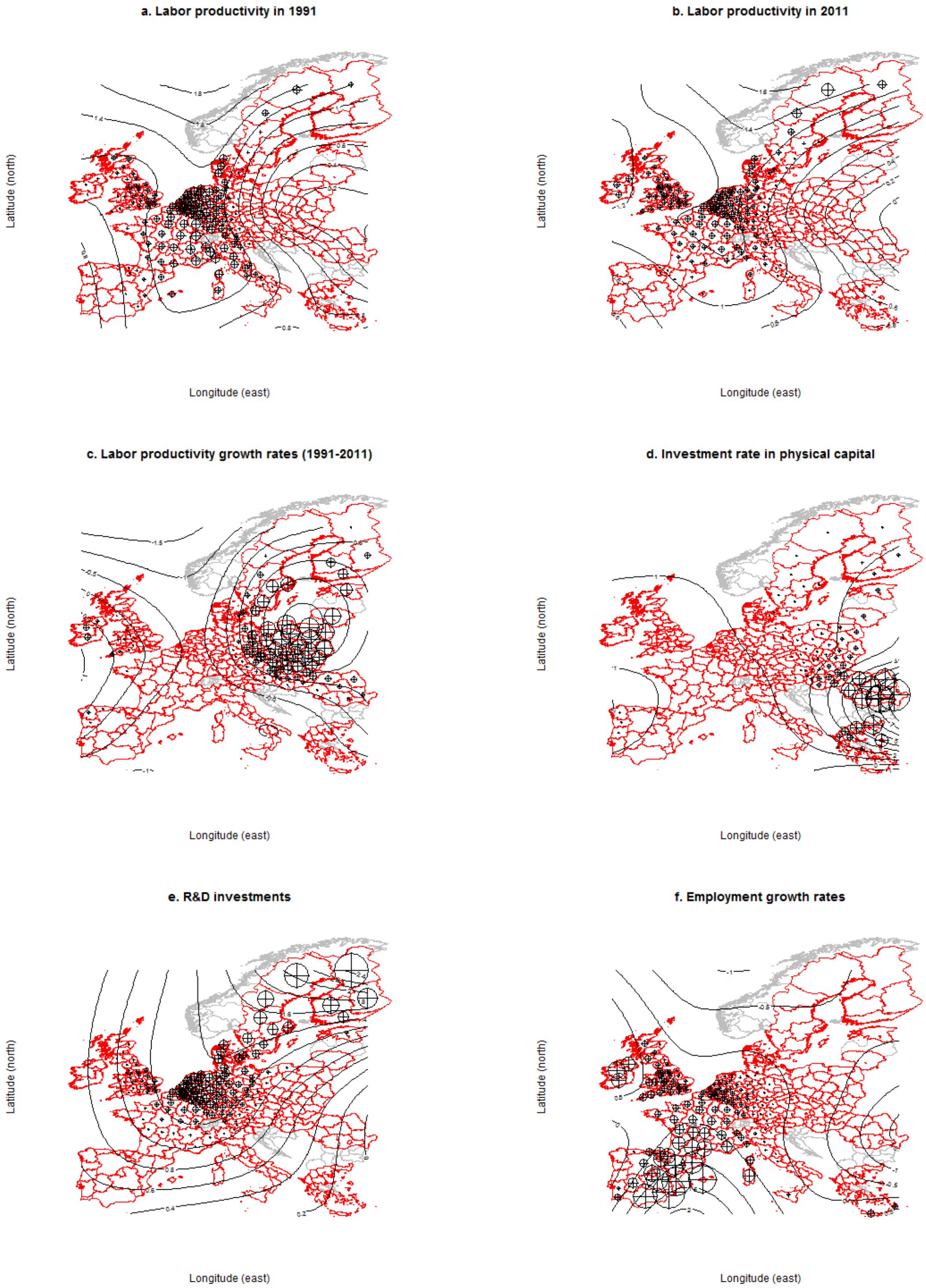

FIGURE 3
Smooth effects from nonlinear a-spatial Solow and Howitt models

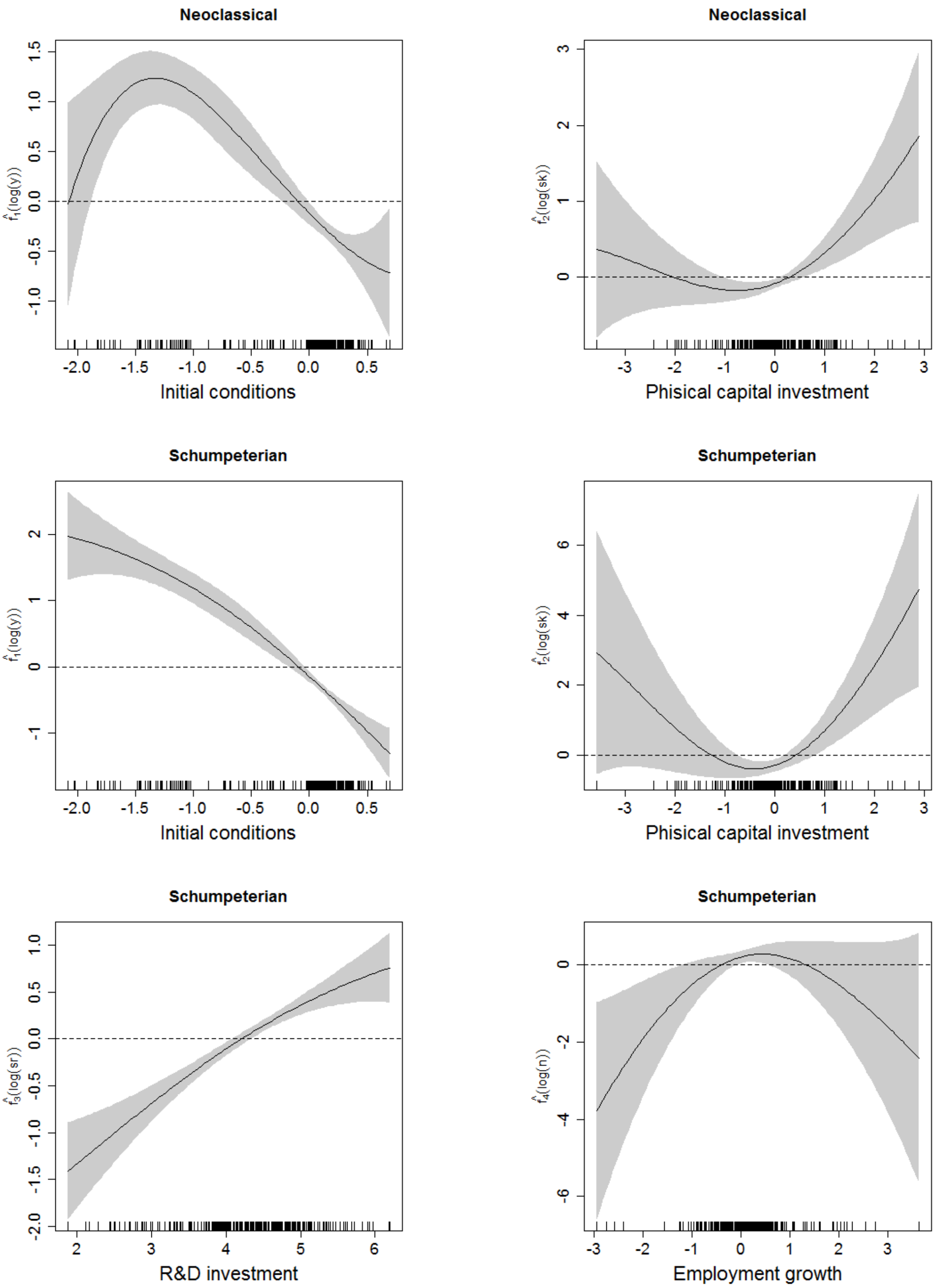

FIGURE 4
Semivariogram of the residuals

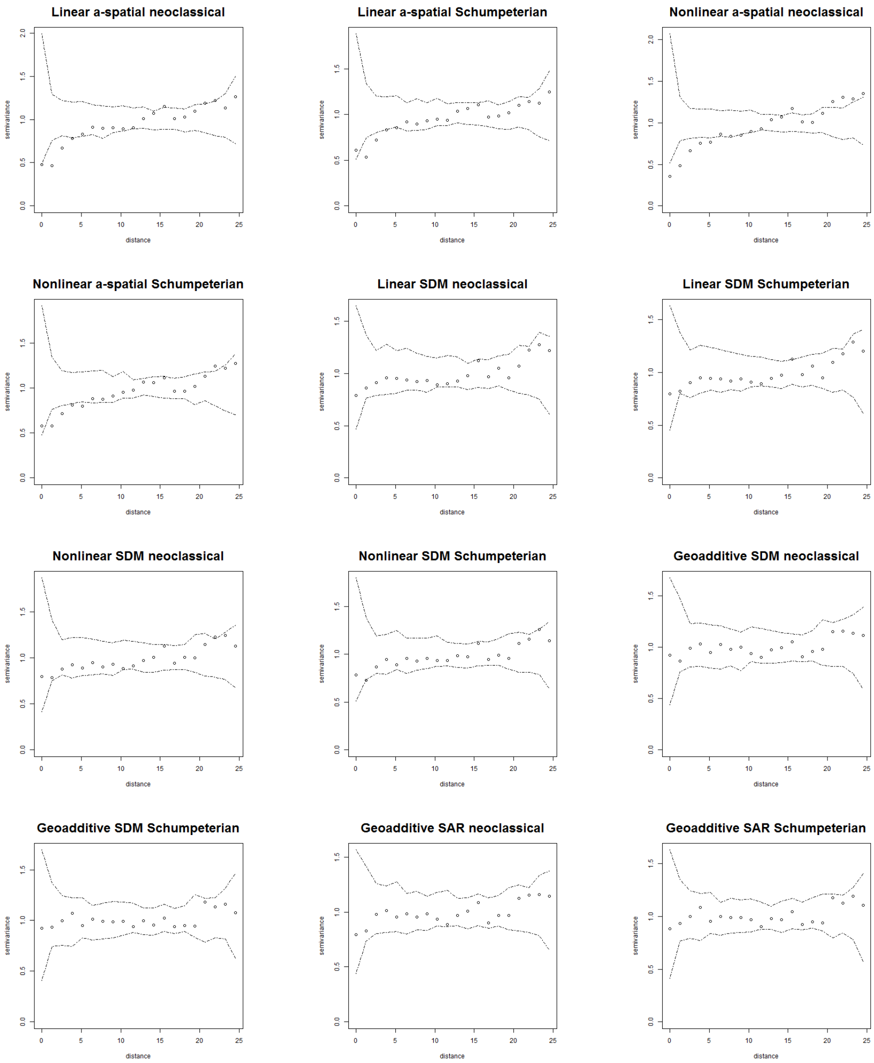

FIGURE 5

Direct, indirect and totale smooth effects from nonlinear SAR EK11 model