ECONOMIC DISCUSSION PAPERS

Efficiency Series Paper 3/2017

Heterogeneous spillovers among Spanish provinces: A generalized spatial stochastic frontier model

Alberto Gude, Inmaculada C. Álvarez, Luis Orea

Departamento de Economía

Universidad de Oviedo

Available online at: http://economia.uniovi.es/investigacion/papers

Heterogeneous spillovers among Spanish provinces: A generalized spatial stochastic frontier model[✧]

Alberto Gude[*]
University of Oviedo

Inmaculada C. Álvarez
Universidad Autónoma de Madrid

Luis Orea
University of Oviedo

Abstract

This paper introduces new spatial stochastic frontier models to examine Spanish provinces' efficiency and its evolution over the period 2000-2013. We use a heteroscedastic version of the spatial stochastic frontier models introduced by Glass *et al.* (2016) that, in addition, allows us to identify the determinants of the spatial dependence among provinces. We contribute to the heterogeneous spatial models that have been introduced in recent years, such as Aquaro *et al.* (2015) and LeSage and Chih (2016) allowing measures of spatial dependence specific to each observation. This feature of the model lets us rank all Spanish provinces in accordance with their degree of spatial dependence, information that will aid policymakers to better allocate public resources between provinces. The period examined is of special interest given that it coincides with a break in the economic growth tendency, which leads to a deterioration in Spain´s economic situation.

Keywords: Heterogeneous spatial spillovers, spatial stochastic frontier models, Spanish provinces.

[✧] We are grateful to participants at the 15th European Workshop on Efficiency and Productivity Analysis (EWEPA) in London, June 2017, for their helpful comments and suggestions. This research was partially funded by the Government of the Principality of Asturias and the European Regional Development Fund (ERDF). Alberto Gude acknowledges financial support from the Spanish Ministry of Education, Culture and Sport (FPU13/02060).
[*] Corresponding author: Oviedo Efficiency Group, Department of Economics, School of Economics and Business, University of Oviedo, 33006 Oviedo, Spain. E-mail: gudealberto@uniovi.es

1 Introduction

Interest in the analysis of productivity at regional level has grown considerably in recent years as productivity growth is one of the most important drivers behind regional income. Thus, analysing how regional productivity evolves over time is essential to provide insights for the promotion of productivity growth in the future. In the recent literature analysing the determinants of productivity, there is a general consensus about the importance of spillover effects, understood to be the benefits obtained by a location when using productive factors from other locations. This literature has benefited from the recent advances in spatial econometric techniques. Elhorst (2003, 2010) and Lee and Yu (2010) presented the Maximum Likelihood estimators of the spatial panel models, while Kappor *et al.* (2007), Mutl and Pfaffermayr (2011) generalized the GM procedure.

During recent years important contributions have been to spatial econometric models. The article by Aquaro *et al.* (2015) introduces a Quasi Maximum Likelihood Method to consider spatial autoregressive panel data models with heterogeneous coefficients. More recently, LeSage and Chih (2016) extend the heterogeneous spatial autoregressive panel model from Aquaro *et al.* (2015) including the prior Bayesian information and deriving marginal effects.[1] The increasing availability of large panel data sets justifies the spatial model with coefficients that vary across the spatial units. Examples of such data set include large panels that cover regions, counties, states or countries in the analysis of economic variables.

With respect to the literature using frontier techniques to measure (decompose) regional productivity, Schmidt *et al.* (2009) and other papers have shown that failure to account for spatial correlation effects in frontier models may yield biased results in both direct and indirect productivity effects (see Glass *et al.*,(2013)). For this reason, it is important to use an econometric framework that allows controlling for the presence of cross-sectional dependence in the observed regional production data. Although there is extensive spatial econometric literature dealing with spatial interactions across regions, the literature on efficiency and productivity analysis does not generally take spatial effects into account.[2] Some recent exceptions are Glass *et al.* (2016) and Mastromarco *et al.* (2016). An excellent review of this -still relatively scarce- literature can be found in Ramajo and Hewings (2016).

Our research extends the spatial frontier model by Ramajo and Hewings (2016) and Glass *et al.* (2016) allowing for specific global and local spillover effects, as in Aquaro *et al.* (2015) and LeSage and Chich (2016). The empirical exercise is performed on Spanish provinces (NUTS-3 level) from 2000 to 2013, for aggregate private economic

[1] This literature has been enriched with the articles introducing Bayesian and Markov Chain Monte Carlo (MCMC) mixture estimation methods (Cornwall, (2016); Cornwall and Parent, (2017)) and Bayesian and MCMC estimation (LeSage and Chih, (2017)).

[2] Traditional growth-accounting exercises decompose economic growth into contributions due to factor accumulation and technological progress. Under this approach technical change and total factor productivity (TFP) growth are often used as synonymous because it is often assumed that all regions operate efficiently. This precludes the existence of catching-up effects among regions/countries. Färe *et al.* (1994) and Kumar and Russell (2002) among other studies used a production frontier approach to capture catching-up effects. In this sense, several measures of TFP growth are often decomposed into three basic sources: technical change, scale effects, and technical efficiency change. Applications using Spanish data are, for instance, Maudos et al (2000) and Badunenko y Romero-Ávila (2014). In addition, the latter considers sectorial interactions, showing that the aggregated productivity changes are driven by intra-sectorial productivity dynamics.

activity. The period contemplated is of special interest given that it coincides with a break in the economic growth tendency, which leads to a deterioration in Spain's economic situation.

In this context, our paper contributes to the literature on regional efficiency and productivity growth in three ways. First, this is the first paper to use a spatial stochastic frontier model to examine Spanish efficiency (productivity) growth using data at provincial level. A parallel, but still ongoing paper focusing on the EU regions, is Ramajo and Hewings (2016). As in this latter paper, we explicitly take into account spatial spillover effects by including a spatial lag of the dependent variable at the frontier. However, unlike the aforementioned paper, we use a heteroscedastic version of the spatial stochastic frontier models recently introduced by Glass *et al.* (2016) in order to examine the determinants of the Spanish provinces' efficiency.

Second, we extend the spatial stochastic frontier models introduced by Ramajo and Hewings (2016) and Glass *et al.* (2016) in the sense that we allow for province-specific degrees of spatial dependence following the heterogeneous spatial panel models by Aquaro *et al.* (2015) and LeSage and Chih (2016, 2017). Our generalized models do not only allow us to identify the determinants of spatial dependence among provinces, but also rank all Spanish provinces according to their degree of spatial dependence. This information could help policymakers to better allocate public resources within Spanish administrative units.

Finally, the suitability of some of the most traditional spatial models can be examined in our framework by allowing the degree of spatial dependence to be a function of the number of adjacent provinces or their average distance from one another, among other covariates. Moreover, our general specification of the degree of spatial dependence in our models aims to balance the suitability of several spatial weight matrices that are often used in the spatial econometric literature to capture spatial spillovers, such as simple and row-normalized binary matrices, or distance-based spatial matrices.

Our empirical strategy focuses on evaluating how public and human capital promotes economic development, introducing global and local spillovers to the econometric specification. Extensive literature exists analysing spillover effects in public and human capital (Kelejian and Robinson, (1997); Pereira and Andraz, (2013); Benos and Karagiannis, (2016), among others). Regarding the Spanish case, more recent papers show strong evidence of positive spillovers in public capital (Ramajo *et al.*, (2017) and Alvarez *et al.*, (2016a,b)) adopting the most recent econometric techniques. In the case of human capital spillovers the literature offers less conclusive results (Ramos *et al.*, (2010)). In this sense, present research sheds light on the nature of global and local spillover effects, estimating differentiated coefficients by regions. In addition, our methodological approach allows the introduction of determinants in the degree of spatial dependence following the suggestion of Autant-Bernard and LeSage (2017, p. 22).

In section 2 we present two generalizations of the spatial frontier model named Generalized Spatial Autoregressive Stochastic Frontier (GSARF) and Generalized Spatial Durbin Stochastic Frontier (SDSF), while the next section shows the estimation method. In section 4 we describe data and sources, section 5 illustrates empirical analysis. The last section discusses our main conclusions.

2 Spatial frontier models

This section first outlines the so-called Spatial Autoregressive Stochastic Frontier (SARF) and Spatial Durbin Frontier (SDF) models proposed by Glass *et al.* (2016), and then develops a generalized version of these two spatial frontier models.

2.1 Spatial autoregressive stochastic frontier (SARF)

The spatial autoregressive stochastic production frontier (SARF) model is a stochastic frontier for panel data with spatial autoregressive dependence, which may be written as follows:

$$y_{it} = \lambda(Wy)_{it} + \sum_{k=1}^{K} \beta_k x_{kit} + v_{it} - u_{it}, \tag{1}$$

where y_{it} represents the output for every cross-sectional unit $(i = 1, \dots, N)$ and time period $(t = 1, \dots, T)$; x_{kit} denotes the *kth* explanatory variable (given $k = 1, \dots, K$) and β_k is an unknown parameter, and $(Wy)_{it} = \Sigma W_i y_{it}$ stands for the endogenous spatial lag of the dependent variable, where W_i is a spatial weight vector where the weights equal one for adjacent provinces and zero for non-bordering provinces (e.g. Kelejian and Robinson, (1997)). Thus, $(Wy)_{it}$ can be viewed as a weighted measure of the output of adjacent provinces to the cross-sectional unit i. This equation includes two error terms, v_{it} and u_{it}. While the former term is a symmetric error term measuring pure random shocks, the latter term is a non-negative error term measuring provincial inefficiency. Finally, the λ parameter is the spatial autoregressive coefficient that measures the degree of spatial correlation between units.

The possible values for the λ parameter lie in an interval $(r_{min}^{-1}, r_{max}^{-1})$, in which r_{min} and r_{max} are the minimum and maximum eigenvalues of the matrix W (LeSage and Pace, (2009)). To impose these restrictions, we have parameterized the spatial autoregressive coefficient as a weighted average of the above lower and upper bounds:

$$\lambda = \left(\frac{1}{r_{min}}\right)(1-p) + \left(\frac{1}{r_{max}}\right)p, \tag{2}$$

$$0 \le p = \frac{\exp(\delta_0)}{1+\exp(\delta_0)} \le 1, \tag{3}$$

where p comes from a standard logistic function and depends on a single (constant) parameter common to all provinces. Thus this model assumes that the intensity of global spatial spillovers does not differ among provinces depending on their location, or how similar they are. If $\delta_0 \to \infty$ then the spatial autoregressive coefficient tends to the upper limit (i.e. $\lambda \to r_{max}^{-1}$), which equals one under the assumption of row standardization. In contrast, if $\delta_0 \to -\infty$, the spatial autoregressive coefficient tends to the lower bound (i.e. $\lambda \to r_{min}^{-1}$).

We first define $W = (W_1, \dots, W_N)$ as a *binary* spatial weight matrix where the weights equal one for adjacent provinces and zero for non-bordering provinces (e.g. Kelejian and Robinson, (1997)). As is customary in spatial econometric literature, we next normalize this matrix by the number of adjacent spatial units, so that each non-zero element of the matrix W equals the inverse of the number of adjacent provinces. The choice of a proper spatial weight matrix is contentious. For instance, Tiefelsdorf *et al.* (1999) point out that this standardization procedure may emphasize the prevalence of the spatial dependence on those units with fewer connections. Given the current debate about the W matrix, we propose later on in this paper a more general model that somehow nests both binary and row-standardized specifications.

2.2 *Spatial Durbin stochastic frontier (SDF)*

The spatial Durbin stochastic frontier (SDF) model introduced by (Glass *et al.*, 2016) is an extension of the spatial autoregressive SARF model that accounts for local spatial interaction, that is:

$$y_{it} = \lambda(Wy)_{it} + \sum_{k=1}^{K} \beta_k X_{kit} + \sum_{k=1}^{K} \theta_k (WX)_{kit} + v_{it} - u_{it} \tag{4}$$

where $(WX)_{kit}$ stands for the spatial lag of the *kth* explanatory variable, and θ_k is the spatial unknown parameter. Like the SARF model, the coefficients measuring the degree of local spatial spillovers are common to all provinces, and they are measured conditional on a priori definition of W. These restrictions are relaxed in our generalized version of the SDF model.

2.3 *Generalized spatial autoregressive stochastic frontier (GSARF)*

The generalized SARF model incorporates time-varying exogenous influences on the degree of global spatial interaction for every time period.

$$y_{it} = \lambda_{it}(Wy)_{it} + \sum_{k=1}^{K} \beta_k x_{kit} + v_{it} - u_{it}, \tag{5}$$

which is identical to equation (1) but with subindexes *i* and *t* added to λ, so that λ_{it} is parametrized by a vector of potential factors determining global spatial spillovers between provinces, and so we have

$$\lambda_{it} = \left(\frac{1}{r_{min}}\right)(1 - p_{it}) + \left(\frac{1}{r_{max}}\right) p_{it}, \tag{6}$$

$$0 \le p_{it} = \frac{\exp\left(\delta_0 + \sum_{m=1}^{M} \delta_m z_{mit}\right)}{1 + \exp\left(\delta_0 + \sum_{m=1}^{M} \delta_m z_{mit}\right)} \le 1, \tag{7}$$

where z_{mit} is the *mth* exogenous determinant of the autoregressive parameter (given $m = 1, \dots, M$), and δ_m is the unknown coefficient for exogenous influences, while δ_0 defines the globally persistent (common) spatial dependence.

Given that our W is a *row-standardized* binary spatial matrix, one candidate variable to be included in $z_{it} = (z_{1it}, \dots, z_{Mit})$ is the number of adjacent provinces (hereafter, n_i). Indeed, different spatial models, or at least the main feature of these models, can be examined by allowing the spatial autoregressive coefficient to be a function of several covariates, such as the number of adjacent provinces. To better understand this feature, let us assume for a moment that the spatial autoregressive coefficient is a simple linear function of n_i, that is $\lambda_{it} = \lambda_0 + \lambda_1 n_i$. In this case, the endogenous spatial lag $\lambda_{it}(Wy)_{it}$ is $[\lambda_0 + \lambda_1 n_i] \cdot \Sigma W_i y_{it}$. It is worth mentioning that if $\lambda_1 = 0$, we obtain the traditional row-normalized binary specification with a common coefficient degree of spatial correlation for all regions. If, in contrast, we assume that $\lambda_0 = 0$, we get a simple binary approach as $\lambda_{it}(Wy)_{it} = \lambda_1 \cdot \Sigma(n_i W_i) y_{it}$, and each non-zero element of $(n_i W_i)$ equals one.[3] This example shows perfectly that a spatial model with binary matrix is equivalent to a spatial model with a row-normalized binary matrix

[3] Other papers construct the weights so that they reflect the commercial relationships among regions (e.g. Álvarez *et al.*, (2003); and Cohen and Morrison Paul, (2004)). The idea behind this approach can be incorporated into our specification by adding, for instance, the (average) freight traffic with neighbouring regions divided by the freight traffic within the region as a new determinant of the spatial spillovers between regions.

where the spatial autoregressive coefficient is assumed to be a linear function of the number of adjacent units. [4]

We have illustrated above that a *heterogeneous* specification of the spatial autoregressive coefficient would capture differences in the coding scheme of the spatial matrices. Obviously, given that we are using a (non-linear) logistic function to model the autoregressive parameter in (7), our generalized models do not exactly nest the above standard spatial models. However, like the linear-based models, they do not use a unique spatial weight matrix, but a combination of several spatial weight matrices (binary and *row-standardized* binary) that are often used in spatial econometric literature to capture spatial spillovers.

On the other hand, it should be pointed out that traditional spatial econometric models do not account for the spatial spillover effect of cross-sectional invariant variables (e.g. regional dummies, time trends and time dummies) due to collinearity problems (see, for instance, Glass *et al.*, (2016), footnote #16). In contrast, our empirical approach allows us to introduce these sorts of variables at least as determinants of the spatial lags of other dependent and independent variables.

2.4 Generalized SDF (GSDF)

Our final and more comprehensive model is a generalized SDF model that incorporates time-varying exogenous influences on both the degree of global and local spatial interaction for every time period.

$$y_{it} = \lambda_{it}(Wy)_{it} + \sum_{k=1}^{K} \beta_k X_{kit} + \sum_{k=1}^{K} \theta_{kit}(WX)_{kit} + v_{it} - u_{it}, \tag{8}$$

This formulation differs from equation (4) in that the subindexes *i* and *t* are added to λ and θ. Thus λ_{it} recalls equations (6) and (7), while θ_{kit} is provided by

$$\theta_{kit} = \theta_k \exp(\textstyle\sum_{m=1}^{M} \eta_m z_{mit}) = \theta_k m_{it}, \tag{9}$$

where η_m is a parameter to be estimated, and the multiplier m_{it} is an exponential function of a set of exogenous determinants of the local spillovers. Like the GSARF model, the coefficients measuring the degree of local spatial spillovers θ_{kit} are modelled as functions of the set of covariates included in the definition of λ_{it}. This allows us to examine whether for instance, local spillovers depend on the number of adjacent provinces, or their structural inequalities. Note that to get an overall picture of this issue, we have assumed that each local spillover depends on the same set of η parameters. This implies that $\theta_{1it}/\theta_1 = \theta_{2it}/\theta_2 = \cdots = \theta_{Kit}/\theta_K$, or in other words, $m_{kit} = m_{it}$, $\forall k = 1, \dots, K$. [5]

3 Estimation

[4] In turn, the $\lambda_{it}(Wy)_{it}$ term can be viewed as a weighted average of two spatial lags of the dependent variable. Indeed, if the autoregressive parameter were a linear function, the spatial lag term $\lambda_{it}(Wy)_{it}$ can be rewritten as:

$$\lambda_{it}(Wy)_{it} = \lambda[\tau_0 \Sigma W_i y_{it} + \tau_1 \Sigma (n_i W_i) y_{it}]$$

where $\lambda = \lambda_0 + \lambda_1$, and the relative magnitude $\tau_j = \lambda_j/\lambda$ of each individual coefficient in λ_{it} allows us to tune the relative importance of each definition of the spatial weight matrix.

[5] In addition, this empirical strategy helps obtaining parameter estimates because we have found convergence problems when we allowed for input-specific multipliers (i.e. $m_{kit} \neq m_{hit}$).

The derivation of the log-likelihood of the above models is a straightforward generalization of the models proposed by Glass *et al.* (2016). First, we need to assume specific distributions for the noise and inefficiency terms. Hereafter we will assume that $v_{it} \sim N(0, \sigma_v)$ and that the variable representing inefficiency is the truncation (at zero) of a normally-distributed random variable with mean zero and standard deviation $\sigma_{uit} = f_{it}(b_{it}) \cdot \sigma_u$, where $f_{it} = e^{b_{it}'\delta} \geq 0$ is a function of exogenous variables.[6] It should be stressed here that estimating an heteroscedastic inefficiency term is also a generalization of the models proposed by Glass *et al.* (2016).

Taking into account the above two assumptions, the log-likelihood function for a sample of N provinces in period *t* can then be written as (see Glass *et al.*, 2016):

$$lnL_t = const + ln|I_N - \lambda_{it}W| - \frac{N}{2} ln[(\sigma_v^2 + f_{it}^2\sigma_u^2)] + \sum_{i=1}^{N} ln\left[\Phi\left(\frac{-(\varepsilon_{it} \cdot f_{it}\sigma_u)/\sigma_v}{(\sigma_v^2 + f_{it}^2\sigma_u^2)^{1/2}}\right)\right] - \frac{1}{2(\sigma_v^2 + f_{it}^2\sigma_u^2)} \sum_{i=1}^{N} \varepsilon_{it}^2, \quad (10)$$

where $\varepsilon_{it} = lnY_{it} - \lambda_{it}(Wy)_{it} - \sum_{k=1}^{K} \beta_k X_{kit} - \sum_{k=1}^{K} \theta_k^{(it)}(WX)_{kit}$, $\lambda^{(it)}$ and $\theta_k^{(it)}$ represent diagonal matrices with diagonal elements respectively set to the non-linear functions of other parameters λ_{it} and θ_{Kit}, while Φ is the standard normal cumulative distribution function. The final log-likelihood function of our spatial stochastic frontier models is obtained by summing the above function from *t*=1 to *t*=T, that is, $lnL = \sum_{t=1}^{T} lnL_t$. Consistent parameter estimates can be obtained by numerically maximizing lnL.

It is worth emphasizing that $ln|I_N - \lambda_{it}W|$ is the logged determinant of the Jacobian of the transformation from ε_{it} to lnY_{it} . As Glass *et al.* (2016) point out, such transformations are undertaken in ML estimations to derive the probability density function of the dependent variable from the probability density function of the disturbance.[7] For ML estimation of spatial models such as equation (10), the transformation from ε_{it} to lnY_{it} takes into account the endogeneity of the spatial lag of the dependent variable.

As is customary in the SFA literature, the error term in any of our spatial stochastic frontier models includes a noise term (v_{it}) and an inefficiency term (u_{it}). Jondrow *et al.* (1982) use the conditional distribution of u_{it} given the composed error term in differences (i.e. ε_{it}), to estimate the asymmetric random term u_{it}. Both the mean and the mode of the conditional distribution can be used as a point estimate of u_{it}. We use the conditional expectation $E(u_{it}|\varepsilon_{it})$ to estimate the asymmetric random term as it is by far the most commonly employed estimator in the stochastic frontier analysis literature (see Kumbhakar and Lovell, 2000).

[6] The defining feature of models with the scaling property is that provinces differ in their mean efficiencies, but not in the shape of the distribution of inefficiency. More details on this specification that satisfies the so-called scaling property can be found in Wang and Schmidt, (2)002; Álvarez *et al.* (2006); and Parmeter and Kumbhakar, (2014).

[7] See also Anselin (1988) and Elhorst (2009).

4 Data

The empirical strategy was performed using individual information on the Spanish provinces from 2000 to 2013 and for private economic activity. The period considered coincides with the beginning of the crisis in 2008, which allows us to analyse its influence on economic performance as well as to contrast a possible structural change. This event sharply intensified the decrease in production decrease and the decline in employment. Nevertheless, this situation did not reduce private and public investment during the whole period, which could force significant spillover effects across Spanish provinces.

Data comes from two main statistical sources. Gross value added (GVA), private labour (number of jobs, *LAB*) and population (number of inhabitants) from the Spanish National Institute of Statistics (*Instituto Nacional de Estadística*, INE). The series of productive (i.e., non-residential) private capital (*KPRI*) and public capital (*KPPUBA*) are taken from the database compiled by Mas *et al.* (2011) and Serrano-Martínez *et al.* (2017) at the *Instituto Valenciano de Investigaciones Económicas* (IVIE). Public capital includes infrastructures and social capital (public spending on health and education). Human capital (HKPRI) represents average years of schooling over total employment, taken from Bancaja Foundation and IVIE (2014). All monetary variables are expressed in 2010 constant values. As a proxy for the different production structures within each province, we use a Herfindhal index measuring the degree of specialization, computed using the percentages of employment in five different production sectors (SPEINDEX).

As determinants of efficiency we introduce human capital, and sector specialization together with a dummy variable (CRISIS) that is equal to one from 2008 onwards in order to capture structural changes associated to the economic crisis period in Spain. We also include indicators representing economies of agglomeration, based on population dispersion and the specialization index (SPEINDEX). For population dispersion we use a Herfindhal index measuring population concentration at municipal level (HER). Please note that provinces with values close to a unit are those with location advantages (Ellison and Glaeser, 1997).

The variability of global and local spillovers across provinces is explained by several determinants. First, the number of neighbours as a measure of centrality (NIP). We also consider the Theil index of inequality (Theil, 1967) adapted to compare each province with their adjacent neighbours, in order to provide a measure concerning structural differences. The Theil Index of inequality in neighbouring productive structure is also customized to differences with adjacent provinces in agriculture (TIPAGR), construction (TIPCON) and services (TIPSER).

$$TIP(k)_{it} = -\frac{1}{C}\sum_{c=1}^{C}\frac{PL_{ct}^{k}}{PL_{it}^{k}}\ln\left(\frac{PL_{ct}^{k}}{PL_{it}^{k}}\right)$$

where PL_{ct}^{k} denotes share of labor in an industry ($k = AGR, CON, SER$), for every adjacent province ($c = 1, \ldots, C$) at time period *t*. This index takes positive values if province *i* displays an average higher specialization.

The Theil index of inequality in neighbouring population (TIPOP) stems from the previous formulation:

$$TIPOP_{it} = -\frac{1}{C}\sum_{c=1}^{C}\frac{POP_{ct}}{POP_{it}}\ln\left(\frac{POP_{ct}}{POP_{it}}\right)$$

where POP_{ct} denotes total population for every adjacent province ($c = 1, \dots, C$) at time period t. By the same token, larger (smaller) neighbouring populated provinces would turn the index to negative (positive) values.

Table 1 shows the descriptive statistics corresponding to the variables used in the production function analysis.

[Insert Table 1 here]

5 Empirical analysis

The SARF, SDF, GSARF, GSDF models have been estimated by ML. The parameter estimates are shown in Table 2. Based on the random effects transformation model introduced by Mundlak (1978) and extended to the original spatial Durbin model by Debarsy (2012), the four models include the individual means of *all* time-varying variables as explanatory variables in order to control for unobserved individual effects correlated with the selected regressors.[8] This implies that the coefficients of all individual within-province means are statistically significant, indicating the need to control for unobserved and (correlated) individual effects in our application. The inclusion, in addition, of the specialization index measuring the production structure of each province as an explanatory variable has also helped to control for provincial heterogeneity.

[Insert Table 2 here]

The four models yield similar results regarding the output elasticities of labour and private capital. The direct effect of both private inputs follows conventional growth accounting, where labour elasticity is about 2/3 and capital elasticity is around 1/3, underestimating the role of capital (Romer, (1987)) according to the convention in the literature. Their elasticities are positive and significant in the four models for labour and in SDF and GSDF models in the case of private capital. Although the associated spillover effects are negative, they are indicative of Myrdal's backwash effects (Myrdal, (1957)) which may arise through competition in production factors between provinces. This would be an indirect corroboration of the existence of agglomeration economies drawing production factors to locations with greater economic activity reinforcing the theories explaining core-periphery patterns (Barbero and Zofío, (2016)). Human capital shows a positive and significant indirect effect which overcomes a negative direct effect of lower significance. In the case of Spain, this result is consistent with previous literature (see, for instance, Ramos *et al.*, 2010 and Boschma *et al.*, (2012)). During the period considered, the direct effect of public capital is not significant, possibly because it includes social capital. At the same time, the spillover effect is not significant, in line with the literature analysing spillovers effects in public capital but in contrast to previous studies including the crisis period (Pereira and Andraz, (2013)). All the models yield identical results for the estimated coefficient of the specialization index, indicating that more specialization

[8] See also Chamberlain (1980). An alternative is to estimate the models with a set of provincial dummies in the same fashion as the True Fixed Effect frontier model (TFE) introduced by Greene (2005). This empirical strategy yields however convergence problems in our iterative procedures to maximize the likelihood function.

within the province reduces home technical capabilities. The coefficient of the trend variable is significant and positive, indicating that there is technological progress.

Evidence of the existence of global spatial spillover effects from neighbouring provinces is found in all the models estimated. Therefore, the income level of neighbouring provinces positively affects economic growth, in accordance with other studies using Spanish provincial data (Arbues *et al.*, 2015). In addition, we observe that the average autoregressive parameter is more intense in GSDF models, when we expand the model taking into account the determinants of local and global spillovers. Regarding those determinants affecting global spillovers, we observe a positive effect of the number of neighbours showing the relevance of the nearest neighbouring provinces, as in Alvarez *et al.* (2016a). The latter study compares the spillovers effects of transport infrastructures using contiguity and inverse distances highlighting a situation where adjacent provinces represent most of the externalities. A population size which is greater than that of adjacent provinces significantly increases the intensity of global externalities. Regarding the Theil indices of sectorial specialization, in the GSDF model we observe a negative effect of specialization in comparison with neighbours in agriculture and services and a positive effect in the case of construction

The GSDF model also allows the introduction of determinants of local spillovers. In our case, the number of adjacent provinces and population size have a positive effect. We can observe that the number of neighbours and population size in comparison with adjacent provinces contribute positively to global and local spillovers. The number of neighbours thus reflects the degree of centrality in a better way than a contiguity (and row-normalized) weighted matrix. Therefore, these results indicate that more centrally located provinces (with more neighbours) and large local markets are richer. These findings corroborate those obtained in econometric studies suggesting that market potential is a powerful driver of economic growth (Mayer, (2008)). Finally, Theil indices of sectorial specialization have similar effects to global spillovers. Therefore, a higher specialization in agriculture and services than neighbouring provinces reduces global and local spillovers while more specialization in construction serves to increase spillovers supporting the idea that the construction bubble contributed to inflate even more the crisis in Spain. The cited negative effect of specialization and the positive influence of population size on the intensity of spatial dependencies should be relevant for policy makers and planners. They should bear in mind that cohesion policies based on Smart Specialisation may be drive in the wrong direction, whilst more recent policies such as the Urban Agenda for the EU could prove more appropriate for regional cohesion.

Overall, the statistical significance of many of the coefficients of λ_{it} and θ_{kit} allows us to conclude that the simple binary matrix,[9] the row-normalized binary matrix and a distance-based spatial matrix provide a partial picture of both global and local spatial spillovers. In other words, the above results suggest that a more comprehensive picture can be obtained using a combination of several spatial weight matrices.

The inefficiency term u_{it} is heteroskedastic both in time and cross-sectional dimension, as the half-normal distributional assumption has been extended, allowing inefficiency variance to depend on a set of determinants (i.e., human capital, the crisis period, population agglomeration and sectoral specialization). The crisis variable does not influence the efficiency level. Human capital does not contribute to increasing

[9] Note that the globally normalized binary matrices is just a scaled version of the simple (i.e. non-normalized) binary matrix. So, both models only differ in the intercept of the degree of spatial dependence.

efficiency, this being in contrast to several articles which conclude that technological knowledge embodied in physical capital requires skilled workers to become operative (Grossman and Helpman, (1991); Arrow, (1962) and Jovanovic and Rob, (1989)) and human capital endowments contribute to raise the given exogenous technology (Romer, (1986) and Lucas, (1988)). Finally, sectorial specialization with respect to neighbours affects efficiency positively.

Our results bring in new contributions to the inconclusive debate about the prevalence of Jacobian externalities (Jacobs, 1969) over Marshallian externalities (Marshall, 1890; Arrow, 1962 and Romer, 1986). The former externalities point out the beneficial impact of urban agglomerations on economic activity, and the latter approach argues that industrial specialization favours employment and productivity growth.[10] The variables population density and degree of urbanization match the stream of the literature which proxies the Jacobian externalities through city sizes and scale of local markets. The highly significant and positive coefficients for these variables goes beyond the key role of urban agglomerations in achieving efficiency and boosting convergence. In contrast, the specialization index has a negative and significant coefficient, indicating that more specialized provinces are closer to the frontier than less specialized ones. It is worth noting that an increase in the degree of specialization not only benefits a particular province but also their neighbors throughout the positive impact of global spillovers.

The descriptive statistics of the estimated λ parameters are provided in Table 3. In the first model, using the conventional spatial autoregressive SARF model, the positive estimate of the parameter associated to the spatial lagged dependent variable yields a small degree of spatial correlation (i.e. λ=0.054) common to all provinces. However, the estimated λ parameter in the SDF model is much larger (λ=0.404) once we add the spatial lags of the explanatory variables. Similar results are obtained using the generalized versions of the SARF and SDF models that allow for province-specific degrees of spatial dependency. In accordance with the lack of significance of the coefficient of the CRISIS variable, these models do not indicate notable changes in the degree of global spatial spillovers from the pre-crisis period to the crisis period.

[Insert Table 3 here]

On the other hand, the inclusion of determinants in local spillovers controls for the neighbours and structural inequalities between origin and destination for all the Theil indices. In Table 4 we show a heterogeneous multiplier on local spillovers estimates using the GSDF model and dividing the sample in line with the crisis period. The effect differs between the two sub periods indicating the negative effect of the crisis.

[Insert Table 4 here]

The differences in provincial heterogeneous multipliers on global and local spillovers estimates using both generalized models (GSARF and GSDF) are detailed in Table 5. We have remarked previously that the estimation of the GSDF model shows a more intensive and significantly positive average autoregressive parameter indicating that a neighbour's income level benefits economic growth. This result corroborates the findings in the literature analysing Spanish provinces which highlight an average positive spillover effect of the autoregressive income level (Ramajo *et al.*, (2017) and Márquez *et al.*, (2015)). However, it is worth noting that our generalized GSDF model allows us to estimate differentiated autoregressive parameters by provinces introducing its determinants.

[10] For a recent review on the current status of this topic, see Groot *et al.* (2016).

Therefore, this methodology provides some explanation as to the differences in the spillover effect of income levels by provinces, thereby extending our knowledge according to geographical location.

[Insert Table 5 here]

The maps in Figure 1 represent the average autoregressive parameter disaggregated by provinces in all our models. We observe that the generalization of the SARF and SDF models allows us to obtain a differentiated autoregressive parameter by provinces according to the determinants considered. Therefore, provinces with more neighbours (i.e., not on the coast) and higher population levels expand global spillovers, while sectorial specialization can influence positively or not, depending on the sector. This represents an interesting finding for policy makers implementing policies oriented to regional cohesion and development and especially useful to the European commission for making investment decisions.

[Insert Figure 1 here]

It should be pointed out at this stage that the above discussion on λ_{it} has to do with the *unitary* (marginal) effect on gross value added of an increase in adjacent provinces' output. In a simple SARF or SDF model, this effect is common to all observations. Therefore, the spatial-based technological (frontier) differences among provinces captured by the term $\lambda(Wy)_{it}$ in (1) and (4) only captures size differences of adjacent provinces. In contrast, the term $\lambda_{it}(Wy)_{it}$ in a generalized SARF or SDF model have to do with both differences in the unitary effect (λ_{it}) and the average size of the adjacent provinces, i.e. $(Wy)_{it}$. Although Figure 1 shows marked differences among provinces in the autoregressive parameter, it so happens that most of the spatial-based technological differences in the generalized models are mainly capturing differences in adjacent provinces' size, in the same fashion as the simple models.[11] This explains why the parameters estimates are quite robust when we move from simple to generalized models. The relatively low role of the differences in λ_{it} is caused by two things. First, while the autoregressive coefficient is a lower and upper bounded parameter, the differences in $(Wy)_{it}$ are not restricted *a priori*. And, second, we have found a slightly negative correlation between the autoregressive parameters (in deviations with respect the sample mean) and neighbors' size. So, the term capturing the effect of the differences in λ_{it}in our decomposition is partially attenuated by this empirical finding as well. Additional consequences of the above finding are that having large neighbors does not necessarily imply having large global spillovers, and, more importantly, that a size increase of 1% in large neighbors might yield a lower (marginal) spillover than an equivalent size increase in a smaller neighbour. This reinforce the next discussion (more) focused on policy implications.

We enrich the above analysis with Figure 2 that represents different situations according to the relation between the autoregressive parameters and the differences in income per capita. The values are in deviations to the average, so positive values imply that these are greater than the average. Following equity criteria, public administration

[11] We have examined this issue using the following decomposition of the spatial-based technological differences:

$$S_{it}^{*} = \lambda_{it}^{*} \cdot (Wy)_{it} + \bar{\lambda} \cdot (Wy)_{it}^{*}$$

where $S_{it} = \lambda_{it}(Wy)_{it}$, $S_{it}^{*} = S_{it} - \bar{\lambda} \cdot \overline{Wy}$, $\lambda_{it}^{*} = \lambda_{it} - \bar{\lambda}$, and $(Wy)_{it}^{*} = (Wy)_{it} - \overline{Wy}$. Thus, while the first term captures deviations due to differences among provinces in the autoregressive parameter (given their neighbors' size), the second term captures deviations in neighbors' size.

should devote all its efforts to improving the economic situation of provinces with a low level of income per capita (left side), because they are the less favoured. However, the proposed method allows us to differentiate among those provinces that are able to benefit more from their neighbours. Therefore, we observe that provinces situated in the upper quadrant to the left have larger autoregressive parameters and lower income with respect to the average. Belonging to this group are the provinces of Andalusia (Cordoba, Granada, Huelva, Jaen and Malaga), the central provinces of the Peninsula (Albacete, Ciudad Real, Toledo, Leon and Zamora), both provinces of Extremadura (Badajoz and Caceres), the provinces in Galicia (Ourense and Lugo) and Murcia. The aforementioned provinces should be concerned with the economic situations of their neighbouring provinces as they have the greatest potential to benefit from improvements in their neighbours' performances. It is worthwhile paying special attention to those neighbours belonging to the same quadrant, because of their similar characteristics (i.e. they are also poor).

[Insert Figure 2 here]

In addition, we can identify some *chains* of poor provinces in the upper quadrant to the left in Figure 2. This is the case of Caceres-Badajoz-Huelva, Lugo-Orense-Leon-Zamora and Toledo-Ciudad Real-Albacete-Murcia. Any of these provinces should be considered as target provinces by the public central administration because they are poor *and* many of their neighbours are also poor *and* they are able to benefit from their neighbours a lot (and so on).

On the right side are situated those provinces with large autoregressive parameters but better economic situations. In this group we find provinces situated in the north of Spain (Araba, Cantabria, Navarra, Tarragona, Zaragoza, La Rioja, Soria and Huesca) and close to Barcelona (i.e. the other important core economy in Spain, in addition to Madrid), Palencia, Valladolid and Burgos. These provinces exhibit higher than average income per capita, although they also have a greater potential for spillover effects. The upshot is that they should also support investment policies in their adjacent provinces. For this reason, central Government should pay attention to neighbours to those provinces, especially in the first group, because of the potential spillover effects and in accordance with equity criteria. In addition, public local administrations in these provinces could benefit from this information in order to provide more insightful advice to central Government for investment decisions. During budget negotiations they should defend the investments of their neighbouring provinces as well as their own economic situation. The remaining provinces present more reduced autoregressive parameters, making it difficult to extend spillover effects. Less potential for global spillovers is observed in the bottom part, especially in Madrid, one of the richest provinces with high economic activity that attracts most of the attention of investors and skilled labour.

In conclusion, our empirical application provides information to local administrations of their second best objectives in their budget negotiations with central Government and identifies the provinces which should receive investments so that the general economic situation can be improved. So, this is an important tool for policy makers in adopting decisions relative to the investment effort.

The maps of Figure 3 show the estimated multiplier of local spillovers (m_{it}) *common* to all the production factors considered in the production function (i.e. private capital and labour, public capital and human capital). The common multiplier for local spillovers effects is displayed by provinces, allowing us to identify which provinces benefit more intensively from neighbour's production factors. The estimated value for each province captures the positive effect of the number of neighbours and population, while sectorial specialization can favour or not depending on the sector. The map on the right shows a large variation of the local spillover multiplier when a GSDF model is estimated, in contrast to the simple SDF that imposes $m_{it} = 1$. A similar result was obtained by Alvarez *et al.* (2016b) that analyses the local spillovers of transport infrastructures using a different approach. However, while Alvarez *et al.* (2016b) find evidence that support a core-periphery model, in which the provinces that benefit more from infrastructures conform a heterogeneous star network (in particular, Madrid, Barcelona and A Coruña), we do not find such a finding because our multiplier involves both public and private factors.

[Insert Figure 3 here]

In the context of spatial econometrics, LeSage and Pace (2009) proposed a method to decompose the effect of production factors into direct and indirect effects. In the presence of spatial dependence between provinces, changes in the characteristics of province *i* can impact outcomes in province *i* (hereafter, direct effect) as well as outcomes in the set of neighbouring provinces, as well as outcomes in neighbours to those neighbouring provinces, and so on (this is the so-called indirect effect). The relative importance of both direct and indirect effect can be examined using the global multiplier$M_{it} = (I_N - \lambda_{it} W)^{-1}$, as all derivatives aiming to measure this type of diffusion of impacts include M_{it} as a common factor. Moreover, direct and indirect effects are often obtained from the average of the elements within the diagonal and outside of the diagonal of M_{it}, respectively. Also, the indirect effect can be decomposed into imported (spill-in) and exported (spill-out) effects, indicating that part of the spillover effect is imported from other provinces and that another part is exported to the rest of the provinces. Taking into account the fact that direct and (overall) indirect effects are highly correlated with parameter λ_{it} (presented in Table 5), we just present the decomposition of indirect effect into spill-in and spill-out in Table 6. Figure 4 shows the geographical distribution of these effects.

[Insert Table 6 here]

[Insert Figure 4 here]

If we focus our attention on the spill-out effects, we could qualify the policy implications inferred from both Figures 2 and 3. More concretely, the shaded provinces in Table 6 are the provinces belonging to the quadrant with high global spillover multipliers and income levels in Figure 2. All provinces in a particular provincial chain of this quadrant are initially good candidates to promote an increase in gross value added because they are poor, many of their neighbours are also poor, and they are able to benefit importantly from their neighbours. Without additional information, all these provinces are equally preferred. However, the spill-out effects in Table 6 and Figure 4 suggest that the preferred provinces should be those adjacent provinces with greater capacity to transmit their enrichment to their nearest provinces (i.e. the provinces with the highest spill-out effect). For example, if we consider the chain of provinces including Caceres, Badajoz and Huelva, more investment efforts should be put into Badajoz because of its high spill-out effect. In the case of the adjacent provinces composed by Lugo, Orense, Zamora and Leon the preferred province is the last, and so on. Therefore, information provided by this decomposition improves our knowledge about which should be the target provinces.

Other interesting results that can be obtain from the GSARF and GSDF models are the technical efficiency scores of each province. Following the method proposed by Glass *et al.* (2016), we present in Table 7 the relative efficiency of each province (RE_{it}) measured with respect the best performing unit in the sample in each time period.[12] Using different techniques, geographical disaggregation and sample period, Badunenko and Romero-Dávila (2014) and Ramajo and Hewings (2016) find out similar efficiency scores to those found in our SARF model (i.e., Extremadura and Castilla-La-Mancha regions are the less efficient regions while Madrid remains the most efficient one). Our generalized models show up that results change significantly once we incorporate determinants of both global and local spillovers (and in the inefficiency term). For instance, the efficiency score of such a singular region as Madrid considerably falls down because this province has a low capacity to capture spatial spillovers from its neighbours in the more comprehensive models (i.e. its specific autoregressive parameter is relatively low), Overall, Figure 5 suggests that we should interpret with caution the relative efficiencies of spatial frontier models with homoscedastic autoregressive coefficients.

[Insert Table 7 here]

[Insert Figure 5 here]

6 Conclusions and future research

This paper introduces new spatial stochastic frontier models to examine Spanish provinces' efficiency and its evolution over the period 2000-2013, which is of special interest given that it coincides with a break in the economic growth tendency, leading to a posterior deterioration in Spain's economic situation. In this sense, it should be stressed that this is the first paper that uses a spatial stochastic frontier model to examine Spanish efficiency growth using data at a provincial level.

This paper also has two additional contributions from a methodological point of view. Firstly, we use a heteroscedastic *and* generalized version of the spatial stochastic frontier models introduced by Glass *et al.* (2016) that simultaneously allows us to examine the determinants of Spanish provinces' efficiency and identify the determinants of the spatial dependence between provinces. Secondly, our general specification of the degree of spatial dependence is able to balance the suitability of several spatial weight matrices that are often used in spatial econometric literature to capture spatial spillovers. Our results seem to confirm this empirical strategy as two traditional (partial) spatial matrices, such as the simple binary matrix and the row-normalized binary matrix, are rejected by the data.

Evidence of the existence of global and local spatial spillover effects from neighbouring provinces is found in all the estimated models. Therefore, the income level of neighbouring provinces positively affects province' valued added and economic growth. This is not a novel finding in the literature. However, unlike previous papers, our

[12] The efficiency scores RE_{it} are calculated as the aggregation of a direct efficiency score and an indirect efficiency measure that captures imported efficiency spillovers. We do not provide an alternative relative total efficiency measure provided by Glass *et al.* (2016) that includes efficiency spillovers from each province to the rest of provinces because the results using simple and generalized models were quite similar. This similarity is due to the fact that the average autoregressive parameters of neighbouring provinces is akin in both models.

model allows us to estimate the determinants of both global and local spillovers by provinces. We observe that global and local spillovers depend on the number of neighbours as a measure of centrality and population agglomeration. In this case, the provinces with the highest multipliers for spillovers are those with the highest population agglomeration or concentrated in the centre of Spain. At the same time, differences in sectorial specialization in comparison with neighbours can affect positively or negatively depending on the sector. Agriculture and services specialization reduce global and local multipliers of spillover effects, while construction expands these effects.

In addition, the differentiated parameters of spatial dependence provide information very helpful for public administration in order to identify target provinces where it is worth it to focus public measures aiming to promote economic growth and a more regional cohesion. Moreover, regarding the inefficiency term, our findings suggest that new cohesion policies based on the *Urban Agenda for the EU* may be more appropriate for the purpose of increasing Spanish provincial efficiency rather than the *Smart Specialization Strategies* strongly promoted during recent years.

In the near future we plan to extend our empirical exercise in several ways. First, we aim to estimate productivity growth and its determinants using the spatial TFP growth decomposition proposed by Glass *et al.* (2013) that, in turns, extends the standard TFP growth decomposition by including direct and indirect components associated to the spatial autoregressive nature of the production frontier. We should, however, adapt their model to a spatial stochastic frontier model with time-varying efficiency determinants. Second, we would like to extend our empirical exercise disaggregating total economic activity by sectors, and allowing for potential inter-sectoral spillovers following a similar strategy as in Gude *et al.* (2017). Finally, we will also explore feasible specifications of the stochastic frontier model that allow for spatial correlation in both noise and inefficiency terms.

7 References

Álvarez, A., Amsler, C., Orea, L., and Schmidt, P. (2006). Interpreting and testing the scaling property in models where inefficiency depends on firm characteristics. *Journal of Productivity Analysis*, 25(3), 201-212.

Álvarez, A., Orea, L., and Fernández, J. (2003). La productividad de las infraestructuras en España. *Papeles de Economía Española*, (95), 125-136.

Álvarez, I. C., Barbero, J., and Zofío, J. L. (2016a). A spatial autoregressive panel model to analyze road network spillovers on production. *Transportation Research Part A: Policy and Practice*, 93, 83-92.

Álvarez, I., Condeço, A.M., Gutierrez, J. and Zofío, J.L. (2016b). Integrating networks analysis with the production function approach to study the spillover effects of transport infrastructure. *Regional Studies*, 50 (6), 996-1015.

Anselin, L. (1988). *Spatial econometrics: methods and models*. Kluwer Academic Publishers.

Aquaro, M., Bailey, N. and Pesaran M.H.(2015). *Quasi maximum likelihood estimation of spatial models with heterogeneous coefficients*. Working Paper Nº 15-17. USC Dornsife Institute for New Economic Thinking.

Arbues, P., J. F. Baños and M. Mayor. 2015. "The spatial productivity of transport infrastructure", *Transportation Research Part A*, 75, 166-177.

Arrow, K. J. (1962). The Economic Implications of Learning by Doing. *Review of Economic Studies*, 29(3), 155-173.

Autant-Bernard, C. and LeSage, J. P. (2017). A heterogeneous coefficient approach to the knowledge production function. Working paper.

Badunenko, O., & Romero-Ávila, D. (2014). Productivity growth across Spanish regions and industries: a production-frontier approach. *Regional Studies*, 48(7), 1242-1262.

Bancaja Foundation and IVIE (2014): "Capital Humano en España y su distribución provincial", Database available online: http://www.ivie.es/en/banco/caphum/series.php

Barbero, J., & Zofío, J. L. (2016). The Multiregional Core-periphery Model: The Role of the Spatial Topology. *Networks and Spatial Economics*, 16(2), 469-496.

Benos, N. and Karagiannis, S. (2016). "Do education quality and spillovers matter? Evidence on human capital and productivity in Greece. *Economic Modelling*, 54, 563-573.

Boschma, R., Minondo, A., and Navarro, M. (2012). Related variety and regional growth in Spain. *Papers in Regional Science*, 91(2), 241-256.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies*, 47(1), 225-238.

Cohen, J. P., and Morrison, P.C.J. (2004). Public infrastructure investment, interstate spatial spillovers, and manufacturing costs. *Review of Economics and Statistics*, 86(2), 551-560.

Cornwall G. (2016). Spatial Durbin mixture models. Working Paper.

Cornwall, G. J., and Parent, O. (2017). Embracing heterogeneity: the spatial autoregressive mixture model. *Regional Science and Urban Economics*, 64, 148-161.

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

Elhorst, J.P. (2003). *Unconditional maximum likelihood estimation of dynamic models for spatial panels. Research report 03C27*, University of Groningen, Research Institute SOM (Systems Organizations and Management).

Elhorst, J.P. (2009). *Spatial panel data models*. In: Fischer, M.M., Getis, A. (Eds.), *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*. Springer-Verlag, Berlin.

Elhorst, J.P. (2010). Applied spatial econometrics: raising the bar. *Spatial Economic Analysis*, 5(1), 9-28.

Ellison, G. and Glaeser, E.L. (1997). Geographic concentration in US manufacturing industries: a dartboard approach. *Journal of Political Economy*, 105(5), 889-927.

Färe, R., Grosskopf, S., Norris, M., and Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *American Economic Review*, 66-83.

Glass, A. J., Kenjegalieva, K., and Sickles, R. C. (2016). A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers. *Journal of Econometrics*, 190(2), 289-300.

Glass, A., Kenjegalieva, K., and Paez-Farrell, J. (2013). Productivity growth decomposition using a spatial autoregressive frontier model. *Economics Letters*, 119(3), 291-295.

Glass, A., Kenjegalieva, K., and Sickles, R. C. (2014). Estimating efficiency spillovers with state level evidence for manufacturing in the US. *Economics Letters*, 123(2), 154-159.

Greene, W. (2005). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics*, 126(2), 269-303.

Groot, H. L., Poot, J., and Smit, M. J. (2016). Which agglomeration externalities matter most and why?. *Journal of Economic Surveys*, 30(4), 756-782.

Grossman, G. M., and Helpman, E. (1991). Quality ladders in the theory of growth. *Review of Economic Studies*, 58(1), 43-61.

Gude, A., Dall'erba, S., and Álvarez, I. (2017) *Inter-industry and dynamic spatial spillovers on regional productivity: Evidence from Spanish panel data*, Working Paper.

Jacobs, J. (1969) The Economy of Cities. *Vintage*, New York, NY.

Jondrow, J., Lovell, C. K., Materov, I. S., and Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics,* 19(2-3), 233-238.

Jovanovic, B., and Rob, R. (1989). The growth and diffusion of knowledge. *Review of Economic Studies*, 56(4), 569-582.

Kappor, M., Kelejian, H.H. and Prucha, I.R. (2007). Panel data models with spatially correlated error components. *Journal of econometrics*, 140, 97-130.

Kelejian, H. H., and Robinson, D. P. (1997). Infrastructure productivity estimation and its underlying econometric specifications: a sensitivity analysis. *Papers in Regional Science*, 76(1), 115-131.

Kumar, S., and Russell, R. R. (2002). Technological change, technological catch-up, and capital deepening: relative contributions to growth and convergence. *American Economic Review*, 92(3), 527-548.

Kumbhakar, S. C., and Lovell, C. A. K. (2000). *Stochastic Frontier Analysis*. Cambridge University Press.

Lee, L. and Yu, J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154, 165-185.

LeSage, J. P., and Chih, Y. Y. (2016). Interpreting heterogeneous coefficient spatial autoregressive panel models. *Economics Letters*, 142, 1-5.

LeSage, J. P., and Chih, Y. Y. (2017). A Bayesian spatial panel model with heterogeneous coefficients. *Regional Science and Urban Economics*. In Press.

LeSage, J. P., and Pace, R. K. (2009). *Introduction to Spatial Econometrics*. CRC Press.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22(1), 3-42.

Márquez, M. A., Ramajo, J., & Hewings, G. J. (2015). Regional growth and spatial spillovers: Evidence from an SpVAR for the Spanish regions. *Papers in Regional Science*, 94(S1), S1-S18.

Marshall, A. (1890). Principles of economics. Macmillan, London.

Mas, M., F. Pérez and E. Uriel. 2011. “El stock y los servicios del capital en España y su distribución territorial (1964-2009)”, Madrid: Fundación BBVA. (http://www.ivie.es/news/2011/ws_sto01.php?idioma=EN).

Mastromarco, C., Serlenga, L., and Shin, Y. (2016). Modelling Technical Efficiency in Cross Sectionally Dependent Stochastic Frontier Panels. *Journal of Applied Econometrics*, 31(1), 281-297.

Maudos, J., Pastor, J. M., and Serrano, L. (2000). Efficiency and productive specialization: an application to the Spanish regions. *Regional Studies*, 34(9), 829-842.

Mayer, T. (2008). Market potential and development, CEPR Discussion Papers 6798.

Mundlak, Y. (1978). On the pooling of time-series and cross-section data. *Econometrica* 64, 69-85.

Mutl, J. and Pfaffermayr, M. (2011). The Hausman test in a Cliff and Ord panel model. *The Econometrics Journal*, 14(1), 48-76.

Myrdal, G. (1957). *Economic Theory and under-developed regions*. Duckworth, London.

Parmeter, C. F., and Kumbhakar, S. C. (2014). *Efficiency analysis: a primer on recent advances. Foundations and Trends in Econometrics*, Now Publishers.

Pereira, A. M., and Andraz, J. M. (2013). On the economic effects of public infrastructure investment: A survey of the international evidence. *Journal of Economic Development*, 38(4), 1.

Ramajo, J., and Hewings, G. J. (2016). Modeling Regional Productive Performance Using a Spatial Stochastic Frontier Approach: New Evidence for Europe (1995-2007). Available at SSRN: http://ssrn.com/abstract=2739612

Ramajo, J., Márquez, M. A., and Hewings, G. J. (2017). Spatiotemporal Analysis of Regional Systems: A Multiregional Spatial Vector Autoregressive Model for Spain. *International Regional Science Review*, 40(1), 75-96.

Ramos, R., Suriñach, J., and Artís, M. (2010). Human capital spillovers, productivity and regional convergence in Spain. *Papers in Regional Science*, 89(2), 435-447.

Romer, P. (1987). Growth based on increasing returns due to specialization. *American Economic Review*, 72, 56-62.

Romer, P.M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94 (5), 1002-1037.

Schmidt, A. M., Moreira, A. R., Helfand, S. M., and Fonseca, T. C. (2009). Spatial stochastic frontier models: accounting for unobserved local determinants of inefficiency. *Journal of Productivity Analysis*, 31(2), 101-112.

Serrano-Martínez, L., Más-Ivars, M., Pérez-García, F., Uriel-Jiménez, E., Benages-Candau, E., & Robledo-Domínguez, J.C. (2017). Acumulación y productividad del capital en España y sus comunidades autónomas en el siglo XXI. *Documentos de trabajo (Fundación BBVA)*, (1), 1-182.

Theil, H. (1967). *Economics and Information Theory*, North Holland.

Tiefelsdorf, M., Griffith, D. A., and Boots, B. (1999). A variance-stabilizing coding scheme for spatial link matrices. *Environment and Planning A*, 31(1), 165-180.

Wang, H. J., and Schmidt, P. (2002). One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis*, 18(2), 129-144.

Table 1. Descriptive statistics

Variable	Mean	Std. Dev.	Min.	Max.
Gross Value Added (thousand Euros)	145616.97	238448.57	12734.21	1483591.5
Labour (thousand jobs)	275.88	402.81	28.4	2489.5
Private Capital (thousand Euros)	24365543.4	38291117.7	2472534.58	255462960
Public Capital (thousand Euros)	5936357.62	5307152.57	1258057.61	37922917.8
Human Capital (schooling years)	10.91	0.83	8.64	13.07
Specialization index (Herfindahl)	0.36	0.07	0.26	0.69
Population dispersion (Herfindahl)	0.14	0.11	0.04	0.58
Theil Index of agricultural specialization	-3.97	20.77	-219.47	0.36
Theil Index of construction	-0.01	0.19	-0.93	0.9
Theil Index of services	-0.02	0.53	-2.07	3.38
Theil Index of population	-3.07	6.01	-25.45	0.36

Table 2. Maximum likelihood estimates

	SARF		SDF		GSARF		GSDF	
	Coef.	**t-stat**	**Coef.**	**t-stat**	**Coef.**	**t-stat**	**Coef.**	**t-stat**
const	-0.081 **	-2.637	-0.011	-0.615	-0.045 **	-2.739	-0.008	-0.480
LAB	0.597 **	16.694	0.668 **	17.819	0.600 **	17.584	0.668 **	19.728
KPRI	0.070	1.808	0.108 **	2.910	0.005	0.109	0.084 *	2.167
HKPRI	-0.017	-1.347	-0.022 *	-2.120	-0.014	-1.271	-0.021 *	-2.183
KPUBA	-0.010	-0.290	-0.009	-0.294	0.019	0.542	0.011	0.343
SPEINDEX	-0.294	-1.804	-0.351 *	-2.384	-0.428 **	-3.256	-0.313 *	-2.476
T	0.012 **	4.867	0.005	1.877	0.014 **	6.873	0.007 **	2.917
Local spillovers								
WLAB			-0.316 **	-4.948			-0.300 **	-4.513
WKPRI			-0.102 **	-2.955			-0.119 **	-3.836
WHKPRI			0.050 **	5.114			0.030 **	4.187
WKPUBA			0.029	1.099			0.011	0.390
Determinants of global spillovers								
WY	0.4931 **	36.003	1.158 **	7.374	0.499 **	42.764	1.271 **	8.119
WYNIP					0.013 *	2.377	0.092 **	4.538
WYTIPAGR					0.003 **	6.306	-0.004 *	-2.111
WYTIPCON					-0.090	-1.878	0.363	1.788
WYTIPSER					0.036	1.567	-0.360 **	-5.336
WYTIPOP					0.002	0.712	0.041 **	4.156
Determinants of local spillovers								
WXNIP							0.087 **	4.681
WXTIPAGR							-0.008 **	-6.138
WXTIPCON							0.513 **	3.395
WXTIPSER							-0.473 **	-5.918
WXTIPOP							0.057 *	2.583
Log of std. dev. of disturbances								
LNSV	-2.835 **	-65.468	-2.983 **	-57.290	-3.310 **	-31.095	-3.314 **	-50.096
Log of std. dev. of half normal								
LNSU	-4.929	-1.485	-3.640 **	-10.844	-2.788 **	-20.208	-3.300 **	-17.792
HKPRI	0.623	1.207	0.439	1.894	0.193	1.824	0.302 *	2.368
CRISIS	-2.121	-0.328	-0.242	-0.634	0.220	1.875	0.264	1.845
HER	-23.200	-0.605	-6.772 *	-2.007	-1.938 *	-2.249	-3.322 *	-2.046
SPEINDEX	-5.950	-0.923	-6.805 *	-2.144	-8.633 **	-5.429	-13.551 **	-5.980
Information criteria								
Log-likelihood	1.415		1.503		1.469		1.610	
Observations	658		658		658		658	

*Notes: ** denotes statistical significance at the 1% level and * at 5% level.*

Table 3. Heterogeneous autoregressive parameter estimates (λ_{it})

Model	Period	Mean	Std. Dev.	Min.	Max.
SARF					
	2000-2013	0.054	-	-	-
SDF					
	2000-2013	0.404	-	-	-
GSARF					
	2000-2007	0.058	0.024	-0.076	0.112
	2008-2013	0.057	0.029	-0.165	0.112
	2000-2013	0.058	0.027	-0.165	0.112
GSDF					
	2000-2007	0.449	0.103	0.114	0.611
	2008-2013	0.443	0.118	0.041	0.648
	2000-2013	0.447	0.109	0.041	0.648

Table 4. Heterogeneous multiplier (m_{it}) on local spillovers estimates (GSDF Model)

Period	Mean	Std. Dev.	Min.	Max.
2000-2007	1.045	0.269	0.368	1.667
2008-2013	1.047	0.32	0.374	2.003
2000-2013	1.046	0.292	0.368	2.003

Table 5. Autoregressive parameters and local spatial multipliers

NUTS-III	Province	Autoregressive parameters (λ_{it})		Local spatial multipliers (m_{it})
		GSARF	**GSDF**	**GSDF**
ES111	Coruña, A	0.047 (37)	0.351 (41)	0.824 (40)
ES112	Lugo	0.064 (17)	0.526 (12)	1.258 (9)
ES113	Ourense	0.056 (26)	0.476 (23)	1.099 (23)
ES114	Pontevedra	0.057 (25)	0.425 (30)	0.972 (29)
ES120	Asturias	0.058 (23)	0.392 (36)	0.889 (35)
ES130	Cantabria	0.073 (12)	0.531 (11)	1.242 (12)
ES211	Araba/Álava	0.063 (19)	0.599 (1)	1.585 (1)
ES212	Gipuzkoa	0.053 (31)	0.429 (29)	0.988 (28)
ES213	Bizkaia	0.066 (15)	0.324 (42)	0.737 (41)
ES220	Navarra	0.064 (18)	0.534 (9)	1.290 (7)
ES230	Rioja, La	0.064 (16)	0.515 (14)	1.203 (15)
ES241	Huesca	0.054 (28)	0.452 (26)	1.018 (27)
ES242	Teruel	0.042 (40)	0.440 (28)	0.892 (34)
ES243	Zaragoza	0.105 (1)	0.539 (8)	1.188 (18)
ES300	Madrid	-0.072 (47)	0.352 (40)	1.141 (20)
ES411	Ávila	0.034 (45)	0.198 (46)	0.452 (46)
ES412	Burgos	0.083 (5)	0.590 (2)	1.475 (2)
ES413	León	0.087 (4)	0.541 (7)	1.249 (10)
ES414	Palencia	0.049 (35)	0.513 (15)	1.248 (11)
ES415	Salamanca	0.074 (10)	0.397 (35)	0.864 (38)
ES416	Segovia	0.037 (44)	0.141 (47)	0.393 (47)
ES417	Soria	0.046 (39)	0.502 (19)	1.136 (21)
ES418	Valladolid	0.098 (2)	0.507 (18)	1.118 (22)
ES419	Zamora	0.042 (41)	0.532 (10)	1.329 (6)
ES421	Albacete	0.077 (8)	0.550 (6)	1.260 (8)
ES422	Ciudad Real	0.066 (14)	0.579 (3)	1.468 (3)
ES423	Cuenca	0.047 (38)	0.366 (39)	0.685 (43)
ES424	Guadalajara	0.053 (30)	0.255 (44)	0.530 (45)
ES425	Toledo	0.061 (20)	0.560 (5)	1.355 (5)
ES431	Badajoz	0.081 (7)	0.523 (13)	1.196 (17)
ES432	Cáceres	0.047 (36)	0.510 (17)	1.242 (13)
ES511	Barcelona	0.055 (27)	0.307 (43)	0.737 (42)
ES512	Girona	0.031 (46)	0.242 (45)	0.603 (44)
ES513	Lleida	0.051 (33)	0.442 (27)	0.939 (30)
ES514	Tarragona	0.071 (13)	0.491 (20)	1.080 (24)
ES521	Alicante/Alacant	0.058 (22)	0.368 (38)	0.839 (39)
ES522	Castellón/Castelló	0.053 (29)	0.402 (34)	0.900 (32)
ES523	Valencia/València	0.082 (6)	0.407 (33)	0.879 (37)
ES611	Almería	0.051 (32)	0.382 (37)	0.880 (36)
ES612	Cádiz	0.057 (24)	0.409 (32)	0.927 (31)
ES613	Córdoba	0.074 (9)	0.576 (4)	1.449 (4)
ES614	Granada	0.073 (11)	0.461 (24)	1.025 (26)
ES615	Huelva	0.037 (43)	0.484 (21)	1.183 (19)
ES616	Jaén	0.060 (21)	0.510 (16)	1.223 (14)
ES617	Málaga	0.041 (42)	0.482 (22)	1.199 (16)
ES618	Sevilla	0.088 (3)	0.419 (31)	0.896 (33)
ES620	Murcia	0.051 (34)	0.458 (25)	1.075 (25)

Table 6. Indirect global spillover multipliers

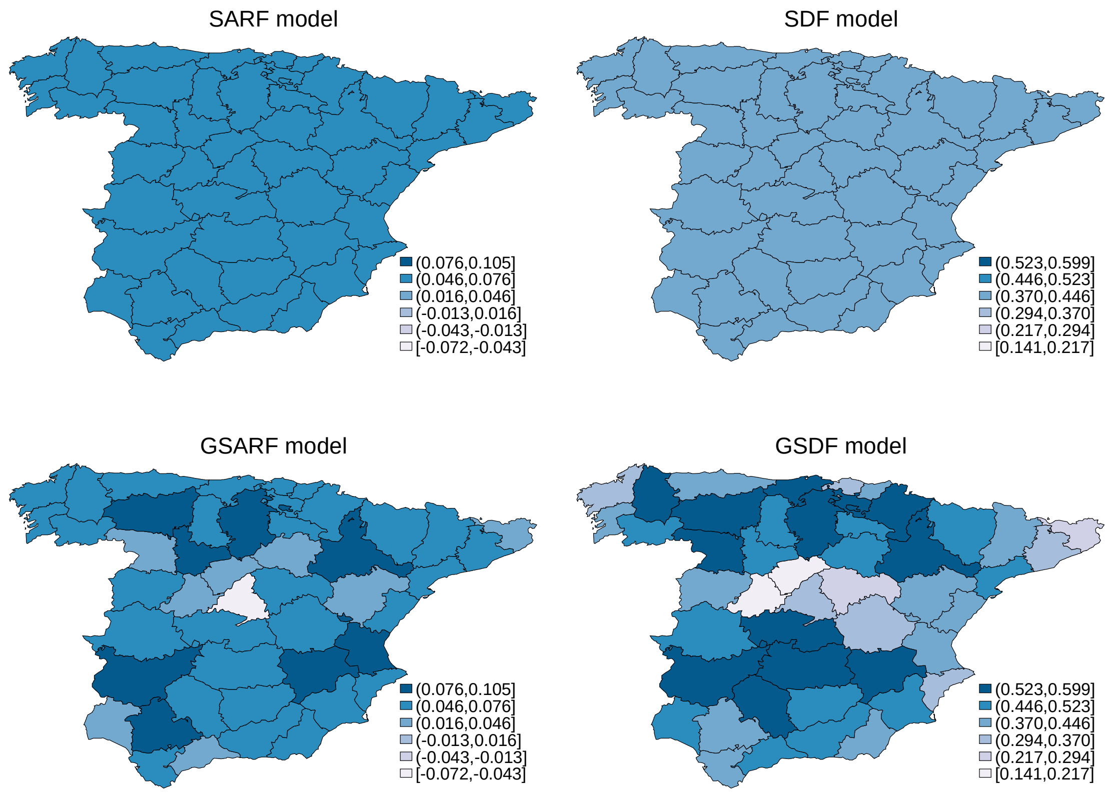
NUTS-III*	Provinces	Spill-in	Spill-out
ES413	León	0,972	1,367
ES421	Albacete	0,942	1,247
ES243	Zaragoza	0,922	1,233
ES431	Badajoz	0,96	1,148
ES613	Córdoba	1,056	1,13
ES130	Cantabria	0,974	1,126
ES412	Burgos	1,055	1,121
ES211	Araba/Álava	1,086	1,109
ES614	Granada	0,833	1,045
ES618	Sevilla	0,745	1,034
ES418	Valladolid	0,852	1,033
ES112	Lugo	0,875	1,028
ES422	Ciudad Real	1,075	1,025
ES423	Cuenca	0,622	0,98
ES514	Tarragona	0,806	0,923
ES220	Navarra	0,985	0,92
ES230	Rioja, La	0,99	0,873
ES411	Ávila	0,32	0,815
ES425	Toledo	0,936	0,811
ES113	Ourense	0,858	0,81
ES242	Teruel	0,717	0,793
ES523	Valencia/València	0,673	0,763
ES513	Lleida	0,707	0,761
ES617	Málaga	0,854	0,744
ES616	Jaén	0,983	0,729
ES416	Segovia	0,229	0,726
ES424	Guadalajara	0,408	0,684
ES419	Zamora	0,949	0,666
ES414	Palencia	0,987	0,665
ES241	Huesca	0,817	0,64
ES114	Pontevedra	0,712	0,629
ES417	Soria	0,847	0,612
ES415	Salamanca	0,668	0,602
ES213	Bizkaia	0,608	0,599
ES612	Cádiz	0,705	0,588
ES620	Murcia	0,755	0,584
ES120	Asturias	0,75	0,562
ES212	Gipuzkoa	0,785	0,552
ES432	Cáceres	0,855	0,544
ES521	Alicante/Alacant	0,636	0,534
ES615	Huelva	0,844	0,532
ES511	Barcelona	0,477	0,446
ES300	Madrid	0,519	0,433
ES111	Coruña, A	0,604	0,409
ES522	Castellón/Castelló	0,678	0,406
ES611	Almería	0,662	0,395
ES512	Girona	0,369	0,285

*Provinces are sorted according to the values of the indirect spill-out. The shaded provinces belong to the upper quadrant on the left of Figure 2.

Table 7. Total relative efficiency estimates (RE_{it})

NUTS-III	Province	SARF	SDF	GSARF	GSDF
ES111	Coruña, A	0.975 (42)	0.974 (33)	0.911 (31)	0.750 (40)
ES112	Lugo	0.987 (31)	0.974 (34)	0.932 (26)	0.885 (15)
ES113	Ourense	0.990 (26)	0.977 (29)	0.904 (34)	0.867 (20)
ES114	Pontevedra	0.990 (27)	0.970 (38)	0.885 (41)	0.796 (32)
ES120	Asturias	0.992 (21)	0.982 (22)	0.954 (14)	0.822 (27)
ES130	Cantabria	0.991 (25)	0.977 (30)	0.964 (10)	0.933 (6)
ES211	Araba	0.999 (2)	1.000 (1)	0.968 (7)	1.000 (1)
ES212	Gipuzkoa	0.989 (28)	0.972 (35)	0.937 (25)	0.839 (23)
ES213	Bizkaia	0.980 (39)	0.970 (37)	0.962 (11)	0.760 (39)
ES220	Navarra	0.986 (33)	0.967 (40)	0.931 (27)	0.924 (10)
ES230	Rioja, La	0.997 (9)	0.988 (10)	0.950 (18)	0.930 (7)
ES241	Huesca	0.970 (43)	0.962 (45)	0.899 (38)	0.828 (25)
ES242	Teruel	0.981 (38)	0.975 (32)	0.904 (35)	0.788 (33)
ES243	Zaragoza	0.999 (5)	0.999 (3)	0.999 (1)	0.909 (12)
ES300	Madrid	1.000 (1)	0.999 (2)	0.856 (43)	0.707 (42)
ES411	Ávila	0.986 (32)	0.967 (42)	0.841 (44)	0.579 (46)
ES412	Burgos	0.997 (8)	0.990 (6)	0.988 (2)	0.966 (2)
ES413	León	0.987 (30)	0.979 (25)	0.973 (5)	0.938 (5)
ES414	Palencia	0.997 (11)	0.989 (7)	0.945 (21)	0.928 (9)
ES415	Salamanca	0.999 (4)	0.989 (8)	0.951 (15)	0.775 (37)
ES416	Segovia	0.992 (22)	0.966 (43)	0.841 (45)	0.531 (47)
ES417	Soria	0.994 (18)	0.971 (36)	0.857 (42)	0.823 (26)
ES418	Valladolid	0.998 (7)	0.996 (4)	0.968 (8)	0.870 (19)
ES419	Zamora	0.995 (14)	0.986 (12)	0.943 (22)	0.912 (11)
ES421	Albacete	0.998 (6)	0.984 (18)	0.901 (37)	0.902 (13)
ES422	Ciudad Real	0.976 (41)	0.976 (31)	0.930 (28)	0.964 (4)
ES423	Cuenca	0.981 (37)	0.965 (44)	0.889 (40)	0.720 (41)
ES424	Guadalajara	0.995 (13)	0.986 (14)	0.939 (24)	0.649 (44)
ES425	Toledo	0.967 (44)	0.967 (41)	0.906 (33)	0.884 (16)
ES431	Badajoz	0.953 (46)	0.938 (46)	0.840 (46)	0.876 (17)
ES432	Cáceres	0.959 (45)	0.937 (47)	0.810 (47)	0.805 (30)
ES511	Barcelona	0.994 (16)	0.986 (13)	0.961 (12)	0.699 (43)
ES512	Girona	0.953 (47)	0.968 (39)	0.904 (36)	0.637 (45)
ES513	Lleida	0.992 (20)	0.980 (24)	0.940 (23)	0.796 (31)
ES514	Tarragona	0.984 (35)	0.983 (21)	0.974 (4)	0.855 (22)
ES521	Alicante	0.981 (36)	0.983 (19)	0.949 (19)	0.770 (38)
ES522	Castellón	0.991 (24)	0.978 (26)	0.919 (29)	0.778 (36)
ES523	Valencia	0.992 (23)	0.983 (20)	0.979 (3)	0.785 (34)
ES611	Almería	0.994 (15)	0.988 (9)	0.951 (16)	0.785 (35)
ES612	Cádiz	0.985 (34)	0.985 (15)	0.957 (13)	0.814 (29)
ES613	Córdoba	0.999 (3)	0.985 (16)	0.918 (30)	0.965 (3)
ES614	Granada	0.988 (29)	0.985 (17)	0.970 (6)	0.874 (18)
ES615	Huelva	0.994 (17)	0.982 (23)	0.906 (32)	0.862 (21)
ES616	Jaén	0.978 (40)	0.978 (27)	0.947 (20)	0.930 (8)
ES617	Málaga	0.997 (10)	0.995 (5)	0.950 (17)	0.888 (14)
ES618	Sevilla	0.997 (12)	0.988 (11)	0.964 (9)	0.831 (24)
ES620	Murcia	0.993 (19)	0.978 (28)	0.896 (39)	0.816 (28)

Figure 1. Autoregressive spatial parameters (λ_{it})



Source: Own elaboration

Figure 2. Cluster of provinces based on income per capita and autoregressive parameters

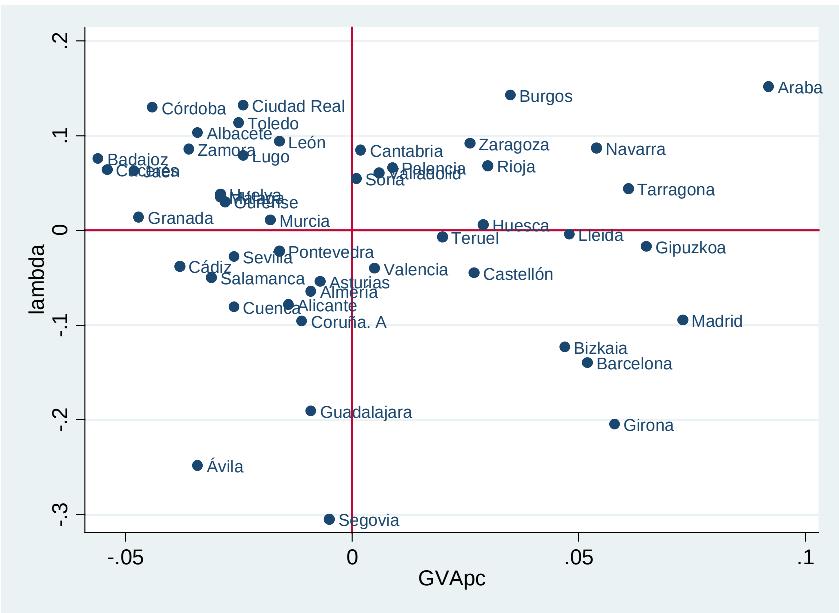
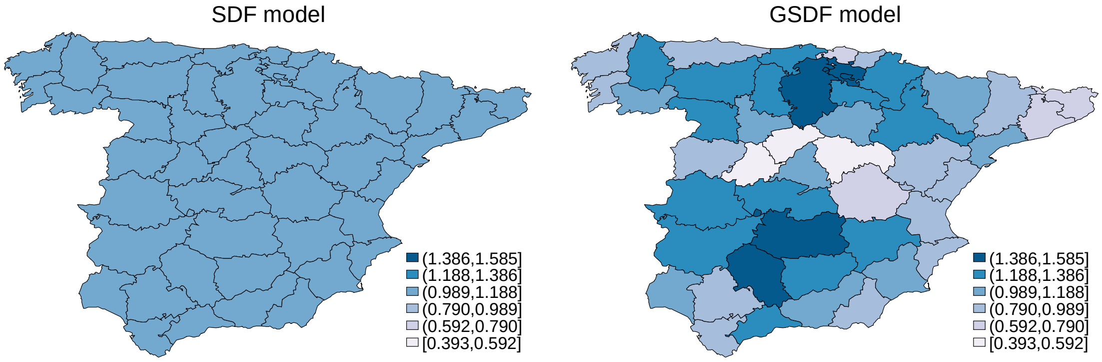
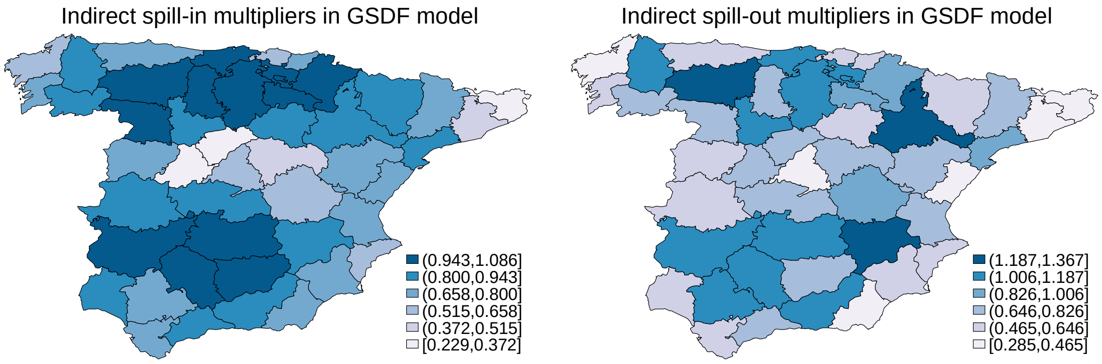


Figure 3. Local spatial multipliers (m_{it})



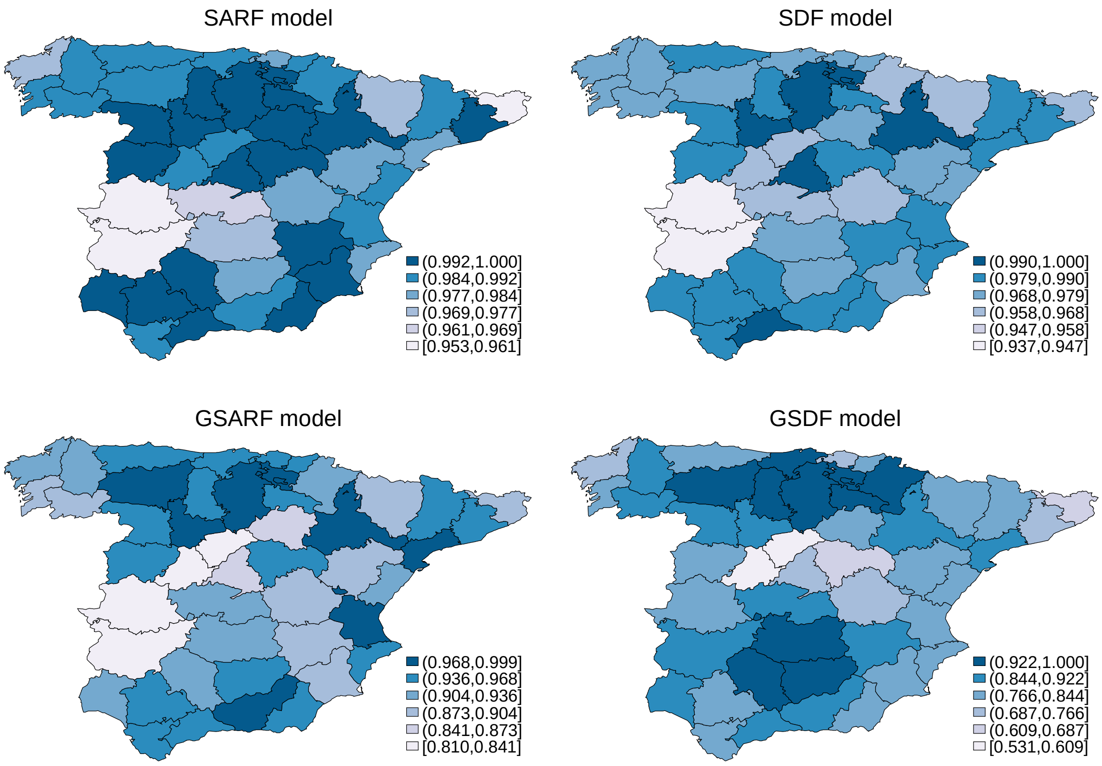
Source: Own elaboration

Figure 4. Indirect global multipliers



Source: Own elaboration

Figure 5. Total relative efficiency estimates (RE_{it})



Source: Own elaboration