On the estimation of spatial stochastic frontier models: an alternative skew-normal approach

Thomas de Graaff¹

Received: 19 October 2018 / Accepted: 8 July 2019 / Published online: 6 August 2019
© The Author(s) 2019

Abstract

This paper deals with an alternative approach to combine spatial dependence and stochastic frontier models using a large statistical literature on skew-normal distribution functions. I show how to combine a spatial dependence structure with a stochastic frontier model, that is, (1) straightforward to estimate, (2) able to combine spatial dependence and a technical efficiency term in a single error term, and (3) produce consistent estimates. With smaller sample sizes estimation of the parameter, governing technical efficiencies becomes imprecise. The consistency of parameter estimation is shown using simulations, and I provide an empirical application to estimate spatially correlated technical efficiencies within an European regional production function context.

JEL Classification R11 · R15

1 Introduction

One of the most distinct features of European regions is that they differ widely in their economic performance, even when controlling for regional characteristics, such as sector structure and population size. Obviously, countries differ in terms of institutions, culture, stability, and so forth, which determine for a large part the international differences in economic performance. However, wealth and income are sometimes even more dispersed *within* countries than across countries. To illustrate this, Fig. 1 shows the dispersion of relative regional GDP per capital across European countries.

European income seems to be concentrated within large metropolitan areas (most notably—the capital cities) such as Paris, London, Luxembourg, Oslo, and Stockholm. Apart from this urban–rural divide, large differences are present as well

✉ Thomas de Graaff
t.de.graaff@vu.nl

¹ Department of Spatial Economics, School of Business and Economics, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Fig. 1 GDP per capita across Europe in 2007—mean EU27 = 100. *Source:* Own calculation based on Cambridge Econometrics (2015)

between subnational regions. The most well-known example is the North–South division within Italy, but these differences can as well be noticed within almost all other countries in Europe. Most notable examples are Spain (North–South division as well), France (with a relatively poorer central area), the UK (where a North–South divide is, albeit less notable, visible as well), and Germany (with the former division between East Germany and West Germany). Thus, some regions within countries are significantly more successful than others—even when faced with similar national institutions.

What makes these regions economically successful? This is perhaps the most crucial and complex research question for both regional policy makers and regional scientists to answer. Policy makers would like to have policies able to steer regions to success, and regional scientists are especially interested in the (nature of the) determinants that drive this success. At the heart, this research question deals with the absolute and relative location advantages of regions.¹

¹ This question is part of a larger debate about the drivers of economic success, both on national and regional levels. Some scientists favour the proposition that regions prosper because of historical events and the associated path dependencies (e.g. Landes 1998), while others emphasize absolute (e.g. Diamond 1998) and relative (e.g. Fujita et al. 1999) locational advantages. Obviously, these different drivers require different policy instruments—if any at all.

A strongly related research question deals with the exact nature of economic performance and how to measure it. To do so, endowment levels should be taken into account. In the economics literature, this can be reflected by the use of regional production functions (see, e.g. Rodríguez-Pose and Crescenzi 2008; Basile et al. 2012). Given the size of production factors, such as labour and capital, regions should attain a certain production level, but usually produce suboptimal. The distance between the optimal and actual production level is usually measured by technical (in)efficiencies and, stochastically, modelled by a stochastic frontier approach.

There is already a sizeable literature dealing with benchmarking regions using regional technical efficiencies modelled by a stochastic frontier approach (see, amongst others, Driffield and Munday 2001; Brock 1999; Puig-Junoy 2001; Puig-Junoy and Pinilla 2008; Alvarez 2007; Otsuka 2017).² This literature usually deals with the relative (sectoral) performance of regions, and this is the approach this paper takes as well. The production factors are then usually constituted of the aggregates of various forms of labour (high skilled and low skilled) and capital (both physical and human) within a region.

However, taking only local endowments into account boils down to an absolute location approach: it does not matter where the region is located with respect to its neighbours. However, the relative location of the region matters as well as regions are intrinsically connected to each other in networks formed by trade, knowledge spillovers, commuting, and migration (Thissen et al. 2016). It is crucial to control for this spatial dependence as omitting it might lead to bias—at least in the estimation of technical efficiencies (Anselin 1988).

The literature that combines spatial dependence and stochastic production frontiers is, although relatively recent, already sizeable. Most studies employ a parametric approach, and the enumeration that follows is definitely not conclusive. One of the first parametric studies was Barrios and Ladado (2010), who uses an iterative back-fitting algorithm to find consistent parameter estimates although they do not allow for correlation between the technical efficiency and spatial dependence structure. Pavlyuk (2010) uses as well a parametric approach, but does not report how he estimates consistently both the spatial dependence process and technical efficiencies. Fusco and Vidoli (2013) and Vidoli et al. (2016) separate out the error term in a spatial lag structure and technical efficiencies, with an application to the Italian wine sector. Kinfu and Sawhney (2015) apply a spatial stochastic frontier analysis to maternal care in India. Glass et al. (2013) decompose productivity growth using a spatial autoregressive model, whereafter Glass et al. (2016) extend the analysis to a spatial panel setting. Finally, Jiang et al. (2017) apply a fixed effects stochastic frontier model to energy efficiency in Chinese Provinces. In addition, there is a smaller literature that resorts to a Bayesian approach and simulation techniques, i.e. Schmidt et al. (2009), Areal et al. (2010), and Tsionas and Michaelides (2016).

² As one referee remarked there are a related macroeconomics and finance literature dealing with a non-parametric approach (e.g. Favero and Papi 1995; Resti 1997). This approach has, however, not yet permeated in the regional science literature.

A specific feature that applies to most of the studies above is that they model the spatial dependence and efficiency processes separately (see, e.g. Fusco and Vidoli 2013). Then, as I will argue below, the error term is by definition multivariate as it is a combination of a normal and truncated normal distribution, where one of them or even both are multivariate due to the involved spatial correlation structure, which makes estimation cumbersome.

In contrast, this study applies a alternative approach firmly rooted in the *statistical* literature. Using a relatively straightforward skew-normal distribution approach, I show how to combine a spatial error structure with a stochastic frontier model, that is, (1) straightforward to estimate, (2) able to combine spatial dependence and a frontier model in a single error term, and (3) produce consistent estimates. The latter is shown by a simulation study, where—although all parameters are consistent—it is clear that the parameter measuring technical inefficiencies is very inefficient (i.e. large standard errors) with small amounts of observations. Skew-normal distribution is not often applied in the econometric stochastic frontier literature with as notable exception Chen et al. (2014), although they are looking at fixed effects panel models instead of spatial dependence models.

The remainder of this paper is structured as follows. The next section introduces the concept of regional technical efficiencies and discusses some measurement issues. Consecutively, it treats the modelling (and its associated estimation) of technical efficiencies in two ways: a mainly econometric and a more statistical one.³ The last subsection deals with the introduction of spatial dependence in stochastic production frontiers. Section 3 provides simulation results to indicate the performance of the proposed estimation methods, within small and realistic samples as usually encountered when benchmarking (European) regions. Section 4 provides an application of spatial stochastic frontier modelling and gives an estimation of the average technical efficiencies of European NUTS2 regions in the period 2000–2010. The last section concludes by indicate how in the proposed framework, more complex spatial dependence structures can be incorporated in stochastic frontier models. Estimation of these models, however, requires complex multivariate likelihood or simulation techniques.

2 Estimating regional technical inefficiencies

Since the late 1970s, economists increasingly recognized that although having access to the same set of production factors, firms do not necessarily produce the same output—and that there was consequently a need to econometrically correct for that (see the seminal papers of Aigner et al. 1977; Meeusen and van den Broek 1977). To explain this variation in output, it was argued that firms do not deploy production factors with the same efficiency. For example, labour might be less productive because of firm-specific lack of monitoring, which opens the possibility for

³ Actually, both modelling approaches date back to one common source, namely Weinstein (1964).

Fig. 2 Production possibility frontiers in a standard (left) and a spatial structure (right)

various forms of shirking on the work floor. Or firms do not have access to the same technology and have therefore different output levels.

This, however, creates a problem. If most firms do not produce according to profit maximization, but systematically lower than that, then traditional production function estimates are *biased*.⁴ Namely, not being able to optimize profits or costs leads to the fact that firms end up beneath an estimated ideal profit level. Consequently, in the literature associated with stochastic production functions, the error terms are usually composed error terms: the traditional error term reflecting noise and a new error term—being strictly positive—measuring a firm's inefficiency.

Analogously to firms, regions with similar inputs do not necessarily attain the same production level as well. Partly, this may be due to missing covariates (such as not being able to correctly measuring human and social capital, but partly this may be caused by the fact that inputs are not always deployed as *efficient* as possible—due to local or national institutions, social structures, etc.

To control for this, the regional science literature has borrowed from the firm-specific efficiency literature the concept of regional stochastic production frontier analysis. As Fig. 1 clearly shows, some regions are probably more efficient than others—even within the same country. And this should be reflected when benchmarking those regions.

To illustrate the concept of a regional stochastic production frontier, the left production isoquant in Fig. 2 shows how a region's technical efficiency can be measured. Denote y as the *given* maximum attainable production a region can get using the production factors x_1 and x_2 , say capital and labour. Regions A_1, A_2, A_3, B_1, B_2 , and B_3 are all producing inefficiently. With the production factors x_1 and x_2 that theoretically enable them to produce y , they produce on average $\hat{y}$. The distance between

⁴ A similar line of reasoning could be held for minimizing cost functions. The remainder of this section deals with production functions, but note that the same arguments hold for costs functions as well.

Y and $\hat{y}$ is then a measure for the average efficiency. More precisely, average efficiency is defined as the ratio $\hat{y}/y$ where y is the length of the line between the origin and y . As a result, technical efficiencies must be smaller than 1.

Estimation of technical efficiencies may, however, be biased in the presence of spatial dependence or unobserved spatial heterogeneity amongst regions. Namely, when one assumes that only neighbouring regions benefit from other regions' technological knowledge through the traditional Marshallian channels of shared customers and suppliers, shared labour pools or spillover mechanisms (or just through unobserved spatial heterogeneity), then straightforward estimation of technical efficiencies is biased. This can be seen in the right production isoquant depicted in Fig. 2. Assume that regions A_1, A_2 , and A_3 belong to country A and regions B_1, B_2 , and B_3 belong to country B , then it is quite well conceivable because of spatial unobserved heterogeneity or spatial dependence that the technical efficiencies of the regions in country A and B are related. In Fig. 2, region B_3 might not produce inefficiently at all *given* the fact that neighbouring regions in country B produce less efficiently. This works as well the other way around. Regions very central in a network and surrounded by very efficiently producing regions have besides a strong economic structure probably a very favourable *relative* location as well. In this context, efficiency can be related to advantages related to the *absolute* location, while spatial dependence relates to the *relative* location. To control for the inefficiency in both production and the geographical location, this paper incorporates a spatial correlation structure in stochastic production frontiers.

To show how one can incorporate spatial dependence in regional cross-sectional stochastic production frontier models, I first revisit concisely the non-spatial stochastic frontier model in Sect. 2.1.⁵ Thereafter, in Sect. 2.2, an alternative estimation and not commonly used estimation method is introduced.⁶ Finally, I show how one can readily incorporate spatial dependence in stochastic production function frontier models in Sect. 2.3.

2.1 Stochastic production frontiers

To start with, assume for simplicity that the production, y_i of firm i ($i \in \{1, \dots, N\}$) can be modelled in a cross section as a Cobb–Douglas production function, thus (using vector notation)⁷:

⁵ See, amongst many others, Kumbhakar and Knox Lovell (2000), Kumbhakar and Tsionas (2006), Wang and Schmidt (2009) and Wang and Ho (2010) for some recent contributions to the econometric literature.

⁶ That is, for econometricians, not for statisticians. Dominguez-Molina et al. (2003, 2007) show how instead of a composed error structure a singular error structure in the form of a skew-normal distribution function can be used.

⁷ For practical purposes, such a production function is in its simplest form denoted as:

$$Y = AL^\alpha K^{1-\alpha}$$

where L stands for labour, K for capital, and A for the level of technology (also known as: labour augmenting technology).

$$y = f(X; \beta')TE, \quad (1)$$

where X is the matrix of production factors, β the vector parameters of the Cobb–Douglas production function and TE denotes the so-called firm-specific technical efficiency. Thus, TE is a distance measure of the firm to the (maximum) production of the best production firm there is—*within* the sample of firms. As a consequence, TE must be smaller or equal to one for each firm. Aigner et al. (1977) and Meeusen and van den Broek (1977) already specified (1) by assuming that $TE = \exp(-u)$, where u represents a stochastic variable. Assuming a logarithmic specification yields:

$$\ln(y) = \ln(X)\beta - u + v, \quad (2)$$

where u being a stochastic variable as well, where $u \sim N(0, \sigma_u^2)$ and $v \sim N(0, \sigma_v^2)$ and with the explicit condition that $u > 0$.

For likelihood purposes, one usually considers the composite stochastic variable $\epsilon = v - u$. Further, usually both u and v are conveniently considered independent. This enables us to find the marginal density of ϵ , namely:

$$\begin{aligned} f(\epsilon) &= \int_0^\infty f(u, \epsilon) du \\ &= \frac{2}{\sigma} \phi\left(\frac{\epsilon}{\sigma}\right) \Phi\left(-\frac{\epsilon\lambda}{\sigma}\right) \end{aligned} \quad (3)$$

where $\sigma_\epsilon = \sqrt{\sigma_u^2 + \sigma_v^2}$ and $\lambda = \frac{\sigma_u}{\sigma_v}$. Note that the marginal distribution of ϵ is a conditional distribution of u and v and that u and v are intertwined by this conditional nature.⁸ Note further that an estimate of the technical efficiency can now be obtained by finding the distribution of $f(u|\epsilon)$.

Obviously, estimation of this model with ordinary least squares regression creates a bias because of the simultaneous appearance of the two stochastic variables with one being truncated at zero. The traditional estimation procedure uses a likelihood procedure based on the density in Eq. (3). However, introducing a more complex error structure in specification (2) is rather cumbersome and not very intuitive. The next subsection proposes therefore an alternative specification and corresponding estimation procedure, which is more straightforward to adapt.

2.2 A skew-normal approach

The stochastic error structure, ϵ , in specification (2) dates back to Weinstein (1964) and can be rewritten in its most simple form as the sum of a normal and a truncated normal distributed variable:

$$\epsilon = \delta|\mu| + \sqrt{1 - \delta^2}v, \quad (4)$$

⁸ A different way of denoting this is to observe that we are interested in the probability $\pi(v|u' > 0)$ where u' is now a normally distributed variable.

Fig. 3 Several examples of the skew-normal distribution function with varying α

where μ and ν are independent $N(0, 1)$ variables, and $\delta \in (-1, 1)$. Here, the stochastic variable ϵ is generated by means of convolution.

A different genesis of ϵ can be realized by conditioning:

$$\epsilon = (\nu | \mu > 0), \quad (5)$$

where (μ, ν) is distributed as a bivariate normal random variable with correlation δ . From here, it is quite straightforward to show that both geneses (4) and (5) of ϵ lead to the same density function:

$$\epsilon Z \sim SN(\alpha) = 2\phi(x)\Phi(\alpha x), \quad (6)$$

which is called the skew-normal density function.⁹ The parameter α in density (6) is a skewness parameter and determines the shape of the density function.

Density (6) is shown in Fig. 3 for some values of the parameter α . When α is positive, the density is skewed to the right and when it is negative, it is skewed to the left. When α is zero, the density becomes a standard normal density function and if $\alpha \rightarrow \infty(-\infty)$, then the density converges to the half-normal density; $2\phi(x)$ for $z \geq 0(\leq 0)$.

Conveniently, if $\epsilon \sim SN(\alpha)$ and $\ln(y) = \ln(X\beta) + \sigma\epsilon$, then the affine transformation $\ln(y) \sim SN(\ln(X\beta), \sigma^2, \alpha)$ holds, which can be expressed as:

$$\epsilon \sim 2\phi((\ln y) - \ln(X\beta); \sigma^2) \Phi(\alpha(\ln(y) - \ln(X\beta))). \quad (7)$$

⁹ The seminal paper in this field is by Azzalini (1985). Other relevant references with respect to the skew-normal distribution are, among others, Azzalini and DallaValle (1996), Azzalini and Capitanio (1999), Azzalini (2005), and Arellano-Valle and Azzalini (2006, 2008).

Note that in this case $\ln X\beta$, σ^2 and α can be seen as a location parameter, a scale parameter and a skewness parameter, respectively.

The direct relation between specification (2) and (7) can be seen as well through stating $\ln(y_i) - \ln(X_i)\beta = \pi(v|u > 0) = \epsilon$, where

$$\epsilon = \begin{pmatrix} u \\ v \end{pmatrix} \sim N(0, \Omega^*), \quad \Omega^* = \begin{pmatrix} 1 & \delta' \\ \delta' & \sigma^2 \end{pmatrix}, \quad (8)$$

and where $\alpha = \delta/\sqrt{1 - \delta^2}$, $\delta = \sigma_u$, and $\sqrt{1 - \delta^2}\sigma_\epsilon = \sigma_v$. The latter equality signifies the intrinsic relation between u and v which is implicit in specification (2). Note that specification (2) only holds when $\delta < 0$. I do not explicitly impose this condition on the model, but choose to leave this as an empirical test.

Estimation of the density in (7) is rather straightforward. When using maximum likelihood, the log likelihood can be denoted as:

$$\ell\ell = -\frac{n}{2} \ln(\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2} \epsilon' \epsilon + \sum_i \ln((2\Phi(\alpha\epsilon_i))), \quad (9)$$

where ϵ_i is the i th observation of the vector ϵ .

Skew-normal distributions are not much used in econometrics, but for this purpose, they will do very nicely.¹⁰ They allow us to use a single error term instead of a composite one, which has some benefits (such as clarity) when working with multivariate distributions. Moreover, the interpretation of the parameters seems as well more intuitive (using scale, location, and skewness parameters). A disadvantage of using skew-normal distributions is the need to use a re-parametrization of the parameters in order to estimate them properly.

The next subsection introduces a spatial variant of this distribution function and applies it to both spatial lag and spatial error models.

2.3 Spatial dependence in stochastic production frontier models

Adopting the skew-normal distribution enables us to directly adopt a spatial lag modelling approach [or SAR model as defined by LeSage and Pace (2009)] as follows:

$$y = \rho Wy + X\beta + \epsilon, \quad (10)$$

where ϵ is defined by (8). Likewise for the spatial error model:

$$\begin{aligned} y &= X\beta + \mu \\ \mu &= \lambda W\mu + \epsilon \end{aligned}$$

¹⁰ Interestingly, the statistics literature mentions as a possible application of skew-normal distribution function the area of stochastic production frontier models. However, this has yet not permeated fully in the econometrics literature (being the exception Chen et al. 2014)—although there is a nice R package that is able to deal with various forms of skew-normal and skew-t distribution functions, see <http://pbil.univ-lyon1.fr/library/sn/html/sn.html>. For more information about the skew-normal distribution function, see <http://azzalini.stat.unipd.it/SN>.

Some authors, such as Fusco and Vidoli (2013) choose to directly separate the spatial efficiency term and the spatial error structures as, e.g.:

$$\epsilon = |\mu| + [I - \lambda W]^{-1} \nu \quad (11)$$

corresponding with model (4). Note that apart from interpretation issues, this actually raises additional difficulties with the fact that μ and ν should now be related to each other in a very intricate multivariate way.

If we now adopt the standard notation that $A = [I - \rho W]$ and $B = [I - \lambda W]$, then the log likelihood specified in (9) is straightforwardly adapted to a spatial lag or spatial error model. For instance, the log likelihood for a stochastic frontier model with spatial dependence in the error term resolves to:

$$\ell \ell = -\frac{n}{2} \ln(\pi) - \frac{n}{2} \ln(\sigma^2) + \ln |B| - \frac{1}{2} \epsilon' \epsilon + \sum_i \ln((2\Phi(\alpha \epsilon_i))), \quad (12)$$

where $\epsilon = \frac{1}{\sigma} B[\ln(Y) - \ln(X\beta)]$.

Finally, I need to calculate the technical efficiencies as resulting from the vector ϵ . For this, I need an expression for $\mathbb{E}(u|\epsilon)$. Dominguez-Molina et al. (2003) give a generic expression for $u|\epsilon$ ¹¹, namely being a normal distribution with mean and variance equal to:

$$\text{Mean} = \frac{1}{\frac{\delta^2}{\sigma^2} + 1} \frac{\delta}{\sigma^2} \epsilon \quad (13)$$

$$\text{Variance} = \frac{1}{\frac{\delta^2}{\sigma^2} + 1} \quad (14)$$

Estimation of, e.g. the likelihood of (12) yields $\hat{\epsilon}$, $\hat{\delta}$ (using $\hat{\alpha}$) and $\hat{\sigma}$ which can then be used to draw from $u|\epsilon$ simulation wise and derive the expectation for each region.

3 Simulation

Already preluding the estimation results in the next section, I set up the simulation with the following cross-sectional structure (in vector notation):

$$Y = 2 + 0.35 \ln(K) + 0.65 \ln(L) + \epsilon, \quad (15)$$

where $\epsilon = \delta|u| + \sqrt{1 - \delta^2}\nu$ and

$$\ln(K) \sim \text{Uniform}(5, 7)$$

$$\ln(L) \sim \text{Uniform}(5, 7)$$

$$u \sim \text{Normal}(0, 0.3)$$

$$\nu \sim \text{Normal}(0, 0.3)$$

¹¹ They actually do this for the multivariate setting where both u and ν and their correlation are multivariately distributed. Our case is a special case of their result.

Table 1 Mean and standard deviation (between parentheses) of frontier model estimation results for various δ 's and number of observations

Obs.	Variable	Simulated δ at		
		– 0.2	– 0.5	– 0.8
250	Constant	2.05 (0.29)	2.00 (0.28)	1.98 (0.22)
	ln(Capital)	0.35 (0.03)	0.35 (0.03)	0.35 (0.02)
	ln(Employment)	0.65 (0.03)	0.65 (0.03)	0.65 (0.03)
	$\hat{\sigma}$	0.32 (0.04)	0.31 (0.04)	0.29 (0.03)
	$\hat{\delta}$	– 0.35 (0.32)	– 0.41 (0.32)	– 0.73 (0.21)
1000	Constant	2.03 (0.16)	1.98 (0.14)	1.99 (0.11)
	ln(Capital)	0.35 (0.02)	0.35 (0.02)	0.35 (0.01)
	ln(Employment)	0.65 (0.02)	0.65 (0.02)	0.65 (0.01)
	$\hat{\sigma}$	0.31 (0.02)	0.30 (0.02)	0.30 (0.02)
	$\hat{\delta}$	– 0.31 (0.26)	– 0.41 (0.25)	– 0.78 (0.08)
10,000	Constant	2.03 (0.06)	1.99 (0.05)	2.00 (0.04)
	ln(Capital)	0.35 (0.01)	0.35 (0.00)	0.35 (0.00)
	ln(Employment)	0.65 (0.01)	0.65 (0.00)	0.65 (0.00)
	$\hat{\sigma}$	0.31 (0.01)	0.30 (0.01)	0.30 (0.01)
	$\hat{\delta}$	– 0.29 (0.15)	– 0.47 (0.11)	– 0.79 (0.05)

I draw $\ln(K)$, $\ln(L)$, u and v 1000 times, where I vary the number of observations—so the length of these vectors—as well (with lengths 250, 1000, and 10,000) and then estimate the model parameters, where after I calculate the distributional mean and standard deviation of each parameter allowing me to infer consistency and efficiency for each of the model parameters.

Table 1 gives the results of a simulation exercise with only a frontier model. Here, the number of observations (250, 1000, 10,000) is varied as well as the simulated value of δ (– 0.2, – 0.5, – 0.8). All variables, except for $\hat{\delta}$, behave as expected and conform theory. They converge to their true values as the sample size gets bigger. $\hat{\sigma}$ also converges to its true value, but only for large sample sizes and large true δ 's. Moreover, its standard deviation is much larger than the other parameters, making this parameter relatively imprecise to estimate.

To simulate a spatial stochastic frontier model, I use the spatial weight matrix, W , from the empirical application (see Sect. 4).¹² So, the model I now estimate is the same as model 15, but now I have in addition the following specification for the error term μ :

¹² I deliberately choose for a realistic weight matrix as these types and sizes typically occur in the literature. Artificial weight matrices are slightly cumbersome to make, and one typically needs to resort to rook or queen matrices as has been done in the simulations in Anselin and Florax (1995). Moreover, first-order contiguity does not really make sense as we have islands and thus disjointed dependencies and fully specified distance matrices have disadvantages as well (see LeSage and Pace 2009). Four-nearest neighbour matrices as is used here are often used, but obviously you can vary the type of W -matrix as well in the simulation. For reasons of conciseness, I refrain from this option.

Table 2 Mean and standard deviation (between parentheses) of a frontier model estimation results with a spatial error structure for various δ 's and λ 's

Variable		Simulated δ at		
		– 0.2	– 0.5	– 0.8
$\lambda = 0.2$	Constant	2.03 (0.33)	1.96 (0.21)	1.96 (0.27)
	ln(Capital)	0.35 (0.03)	0.35 (0.03)	0.35 (0.02)
	ln(Employment)	0.65 (0.03)	0.65 (0.03)	0.65 (0.02)
	$\hat{\lambda}$	0.17 (0.09)	0.18 (0.09)	0.18 (0.09)
	$\hat{\sigma}$	0.32 (0.04)	0.30 (0.04)	0.29 (0.04)
	$\hat{\delta}$	– 0.27 (0.33)	– 0.33 (0.35)	– 0.68 (0.30)
$\lambda = 0.5$	Constant	2.07 (0.36)	1.96 (0.35)	1.95 (0.28)
	ln(Capital)	0.35 (0.03)	0.35 (0.03)	0.35 (0.02)
	ln(Employment)	0.65 (0.03)	0.65 (0.03)	0.65 (0.02)
	$\hat{\lambda}$	0.48 (0.07)	0.48 (0.07)	0.48 (0.07)
	$\hat{\sigma}$	0.32 (0.04)	0.31 (0.04)	0.30 (0.04)
	$\hat{\delta}$	– 0.31 (0.32)	– 0.41 (0.32)	– 0.73 (0.22)
$\lambda = 0.8$	Constant	2.17 (0.59)	1.89 (0.57)	1.88 (0.44)
	ln(Capital)	0.35 (0.03)	0.35 (0.03)	0.35 (0.02)
	ln(Employment)	0.65 (0.03)	0.65 (0.03)	0.65 (0.02)
	$\hat{\lambda}$	0.78 (0.04)	0.78 (0.04)	0.79 (0.06)
	$\hat{\sigma}$	0.32 (0.04)	0.31 (0.04)	0.30 (0.04)
	$\hat{\delta}$	– 0.33 (0.32)	– 0.42 (0.32)	– 0.74 (0.20)

$$\mu = \lambda W\mu + \epsilon, \quad (16)$$

with ϵ as before.

The size of the weights matrix is 256×256 , so the sample size is restricted. Therefore, I vary now δ (with -0.2 , -0.5 , and -0.8) and λ (with 0.2 , 0.5 , and 0.8). For time constraints, I now draw $\ln(K)$, $\ln(L)$, u and v 100 times.

Table 2 presents the simulation results of the corresponding frontier model with an error structure. Conform Table 1, it is clear that with small sample sizes (in this case being 256), the parameter measuring technical efficiencies (α) is not very precisely estimated, whereas all other parameters are very close to their true value. When the true δ is closer to one or, to a lesser extent, when λ gets higher, estimation becomes slightly more efficient, but not by much.

4 Empirical application: the efficiency of European regions

In this section, I apply the concept of spatial stochastic frontiers to European NUTS-2 regional production functions. The next subsection first describes concisely the data, and the subsequent subsection gives the estimation results.

4.1 Data and specification

NUTS-2 (Nomenclature of Units for Territorial Statistics) is a geocode standard for referencing the subdivisions of European countries for statistical purposes, where the addition 2 stands for the geographical level of more or less provinces. I use two databases. For labour, I use the European regional database by Cambridge Econometrics: a database containing detailed sectoral information about the regional provision of labour (see Cambridge Econometrics 2015). For regional gross value added and capital, I adopt the supply and use tables as used previously in Thissen et al. (2016) and explained in detail in Thissen and Diodato (2013a, b). This allows us to deal with one of the prevailing data problems in this literature: the calculation of the capital stock. Typically, this is done with a perpetual inventory method. However, this could be problematic, since shocks in the capital stocks (e.g. by deaths or migrations of a firms) do not manifest themselves in the short run. Because there is information on regional value added of capital (V^K) across regions, sectors and years (so $V_{r,s,t}^K = r_{r,s,t} K_{r,s,t}$), I can circumvent this problem by using data on sector-specific interest rates for capital and thus calculate the capital stock per region, year, and sector ($K_{r,s,t}$).¹³

To avoid idiosyncratic shocks, the data used are the mean over the period 2000–2010, and the economic sectors that they comprise are: agriculture, energy and manufacturing, construction, distribution market services, and non-market services. The countries included in the estimation can be seen in Fig. 1 and are basically all EU25 countries except for Romania and Bulgaria. The total number of NUTS-2 regions in the dataset is 256. Its geographical distribution is shown in Fig. 1.

To define the spatial weight matrix W , I use a k -nearest neighbour algorithm with $k = 4$, where the k -nearest neighbours get a weight of 1. The weights of all other neighbours are set at 0. Finally, I row-standardize W .

These data allow us to estimate the following Cobb–Douglas function:

$$\ln(Y_r) = \beta_0 + \ln(L_r)\beta_1 + \ln(K_r)\beta_2 + \epsilon_r, \quad (17)$$

where Y is gross value added, L is the number of workers multiplied by the average hours worked per week, K is the amount of capital, r is the region, and ϵ is an error term that can be distributed normally or skew normally.

The next subsection provides the results for various sectors and specifications of the production function of (17).

4.2 Results

Table 3 gives the results for various econometrics specifications for the energy and manufacturing sector. I start with the OLS estimation. The factor rewards (or output elasticities) for capital and labour are not conform theory (typically, labour should get an elasticity of around 0.7 and capital of around 0.3), although not

¹³ Unfortunately but not surprisingly, there is only country-specific interest rates instead of region specific ones.

Table 3 Estimation results for energy and manufacturing

	OLS	Sp. error	Sp. lag	Frontier	Fr. error	Fr. lag
Constant	2.40*** (0.16)	3.16*** (0.15)	0.46 (0.26)	2.50*** (0.19)	4.03*** (0.25)	0.51 (1.08)
ln(Capital)	0.77*** (0.03)	0.58*** (0.03)	0.70*** (0.03)	0.76*** (0.04)	0.58*** (0.03)	0.70*** (0.03)
ln(Employment)	0.30*** (0.03)	0.44*** (0.03)	0.32*** (0.03)	0.30*** (0.04)	0.44*** (0.03)	0.32*** (0.03)
$\hat{\sigma}$	0.33*** (0.01)	0.22*** (0.01)	0.29*** (0.01)	0.33*** (0.02)	0.29*** (0.02)	0.29* (0.14)
$\hat{\lambda}$		0.79*** (0.09)			0.78*** (0.09)	
$\hat{\rho}$			0.25*** (0.03)			0.12*** (0.01)
$\hat{\alpha}$				- 1.75*** (0.52)	- 1.41*** (0.37)	- 0.17 (4.41)
Log Lik.	- 80.10	92.28	41.45	11.03	94.86	41.45

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

significantly different from constant returns to scale. A frontier analysis does not alter those strange results, although the likelihood improves significantly. Finally, allowing for spatial dependence (whether that would be a SEM or a SAR frontier model) does not change the estimates of the factor rewards, considerably. However, it is clear that a SEM frontier model performs best in terms of log likelihood. Moreover, a λ of almost 0.8 indicates significant spatial dependence in the error terms, which should be reflected in the estimations of the regional technical efficiencies.

Figure 4 shows the technical efficiencies across European regions for the energy and manufacturing sector as generated by a models with an error structure as given by Eq. 8. Technical efficiencies are measured between 0.2 and 0.6, and clearly, they are spatially correlated, with relatively high technical efficiencies in the centre of Europe (as in France, Belgium, the Netherlands, and Germany) and relatively low technical efficiencies in Poland, Portugal, Greece, and the northern part of the UK.

Figure 5 shows the difference between the efficiencies which are generated by a SEM frontier model and the non-spatial technical efficiencies. Clearly, the introduction of spatial dependence ensures that regions in the centre become less efficient and regions in the periphery become relatively more efficient. (The distribution of technical efficiencies over regions becomes more homogeneous.) In other words, regions in the periphery produce technically inefficient but less so when taking their geographical location into account. Thus, where regions are located matters just as their economic structure.

As the frontier model with spatial error structure perform best in Table 3, I extend the analysis and apply this model to other sectors. The results are shown in Table 4. Clearly, and already indicated by the simulation exercise, the other sectors do now

Fig. 4 Regional technical efficiencies for the energy and manufacturing sector

show evidence of a frontier model structure as the α parameters are not only all close to zero but also have very large standard errors.¹⁴

5 In conclusion

The main aim of this paper is to introduce spatial dependence in stochastic production frontier analysis. I do so by using a skew-normal distribution function approach, which I argue is (1) straightforward to use, (2) able to separate

¹⁴ This is most likely due to the amount of observations. Spatial dependence structures, different starting values and centred parameter transformations as suggested by Azzalini (2005), all yield similar results.

Fig. 5 Difference in regional technical efficiencies for the energy and manufacturing sector between spatially lagged errors and non-spatial model

spatial dependence and technical efficiencies, and (3) produce consistent estimates. These results can be interpreted using the discussion on relative and absolute geographic location. The size of endowments and thus maximum attainable production are caused by a mixture of absolute geographic location and historical path dependence. Similarly for regional efficiency, as it can be argued that they are mainly caused by institutions and social structures. However, spatial dependence measures the location within the network and could thus be a measure for the relative location. Central regions just performed better because they have better access to production inputs and technology. When comparing regions' performance, it would be fairer to control for the region's location within the network.

Table 4 Estimation results for all sectors for frontier model with spatial error

	Agric.	E and M	Constr.	Distri.	Serv.	NM Serv.
Constant	2.43*** (0.57)	4.03*** (0.25)	2.88 (2.79)	2.02* (1.01)	2.05*** (0.55)	2.70*** (0.78)
ln(Capital)	0.68*** (0.03)	0.58*** (0.03)	0.59*** (0.04)	0.66*** (0.03)	0.71*** (0.03)	0.58*** (0.03)
ln(Employment)	0.21*** (0.03)	0.44*** (0.03)	0.40*** (0.04)	0.33*** (0.03)	0.29*** (0.03)	0.43*** (0.04)
$\hat{\sigma}$	0.26*** (0.01)	0.29*** (0.03)	0.22*** (0.01)	0.18*** (0.01)	0.17*** (0.01)	0.18*** (0.01)
$\hat{\alpha}$	- 0.04 (1.15)	- 1.41*** (0.37)	- 0.00 (3.90)	- 0.01 (1.56)	- 0.01 (1.57)	- 0.02 (1.03)
$\hat{\lambda}$	0.57*** (0.08)	0.78*** (0.09)	0.75*** (0.08)	0.78*** (0.08)	0.59*** (0.08)	0.81*** (0.08)
Log Lik.	62.92	94.86	91.66	144.50	164.88	136.06

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Obviously, there is more to this because technical efficiencies itself may be spatially dependent instead of the error structure in total. (For instance, there are spillovers in the adoption of new technology that improve the efficiency or there are specific institutions, such as former guilds or unions, that prohibit the adoption of new technologies spatially concentrated.) In any case, when looking at the efficiency of regions, taking into account spatial dependence—whether in the inefficiency part or not—strongly affects the estimates of technical efficiencies in the energy and manufacturing sector.

Unfortunately, the parameter which governs technical efficiencies (in this case the skew-normal parameter α or the δ parameter in the traditional literature) is volatile when the parameter itself is small or with small amounts of observations (which is typically the case in spatial econometrics applications)—whether in a spatial setting or not. The simulation exercise shows that this does not affect the other parameters but that one should be careful in drawing strong conclusions when applying (spatial) frontier analyses with a small number of observations, such when analysing European regional performance.

For our empirical application, when looking at the energy and manufacturing sector in European regions, taking spatial dependence into account controls more or less for the core-periphery pattern in Europe. Thus, regions in the periphery do not produce that inefficiently only because of their economic structure, but partly as well because of their location and the related diminished access to knowledge and information. Obviously, the estimations are restrictive regarding the data and specification I use. Ideally, one would like to model larger regional datasets, to test the alternative skew-normal approach to spatial stochastic frontier models. A viable avenue for further research would be to use regional panel data instead of cross-sectional data.

Acknowledgements First of all, I would like to dedicate this paper to the memory of my late colleague and mentor Raymond Florax who started working on spatial stochastic frontier models already 10 years ago. This paper was prepared for the 17th workshop on Spatial Econometrics and Statistics in Dijon 2018. I would like to thank Ferdinand Paraguas, Paul Elhorst and Henri de Groot for useful comments on earlier versions of this paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Aigner DJ, Lovell CAK, Schmidt P (1977) Formulation and estimation of stochastic production frontier models. *J Econ* 6(1):21–37
- Alvarez A (2007) Decomposing regional productivity growth using an aggregate production frontier. *Ann Reg Sci* 41(2):431–441
- Anselin L (1988) *Spatial econometrics: methods and models*, vol 4. Springer, Berlin
- Anselin L, Florax RJ (1995) Small sample properties of tests for spatial dependence in regression models: some further results. In: Anselin L, Florax RJGM (eds) *New directions in spatial econometrics*. Springer, Berlin, pp 21–74
- Areal FJ, Balcombe K, Tiffin R (2010) Integrating spatial dependence into stochastic frontier analysis. *Aust J Agric Resour Econ* 56:521–541
- Arellano-Valle RB, Azzalini A (2006) On the unification of families of skew-normal distributions. *Scand J Stat* 33(3):561–574
- Arellano-Valle RB, Azzalini A (2008) The centred parametrization for the multivariate skew-normal distribution. *J Multivar Anal* 99(7):1362–1382
- Azzalini A (1985) A class of distributions which includes the normal ones. *Scand J Stat* 12(2):171–178
- Azzalini A (2005) The skew-normal distribution and related multivariate families. *Scand J Stat* 32:159–188
- Azzalini A, Capitanio A (1999) Statistical application of the multivariate skew-normal distribution. *J R Stat Soc* 61(3):579–602
- Azzalini A, DallaValle A (1996) The multivariate skew-normal distribution. *Biometrika* 83(4):715–726
- Barrios EB, Ladado RF (2010) *Spatial stochastic frontier models*. Technical report, Philippine Institute for Development Studies, Discussion paper series no. 2010-08
- Basile R, Capello R, Caragliu A (2012) Technological interdependence and regional growth in Europe: proximity and synergy in knowledge spillovers. *Pap Reg Sci* 91(4):697–722
- Brock GJ (1999) Exploring a regional technical efficiency frontier in the former USSR. *Econ Plan* 32(1):23–44
- Cambridge Econometrics (2015) *European regional data*. Technical report, database information can be retrieved from Cambridge Econometrics. <https://www.camecon.com/european-regional-data/>. Accessed 20 Aug 2017
- Chen YY, Schmidt P, Wang HJ (2014) Consistent estimation of the fixed effects stochastic frontier model. *J Econom* 181(2):65–76
- Diamond JM (1998) *Guns, germs and steel: a short history of everybody for the last 13,000 years*. Random House, New York
- Dominguez-Molina JA, González-Farías G, Ramos-Quiroga R (2003) Skew-normality in stochastic frontier analysis. Technical report, Comunicacion tecnica no. I-03-18
- Dominguez-Molina JA, González-Farías G, Ramos-Quiroga R, Gupta AK (2007) A matrix variate closed skew-normal distribution with applications to stochastic frontier analysis. *Commun Stat Theory Methods* 36(9):1691–1703
- Driffield N, Munday M (2001) Foreign manufacturing, regional agglomeration and technical efficiency in UK industries: a stochastic production frontier approach. *Reg Stud* 35(5):391–399
- Favero CA, Papi L (1995) Technical efficiency and scale efficiency in the Italian banking sector: a non-parametric approach. *Appl Econ* 27(4):385–395

- Fujita M, Krugman P, Venables AJ (1999) *The spatial economy: cities, regions, and economic trade*. MIT Press, Cambridge
- Fusco E, Vidoli F (2013) Spatial stochastic frontier models: controlling spatial global and local heterogeneity. *Int Rev Appl Econ* 27(5):679–694
- Glass A, Kenjegalieva K, Paez-Farrell J (2013) Productivity growth decomposition using a spatial autoregressive frontier model. *Econ Lett* 119(3):291–295
- Glass AJ, Kenjegalieva K, Sickles RC (2016) A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers. *J Econom* 190(2):289–300
- Jiang L, Folmer H, Ji M, Tang J (2017) Energy efficiency in the chinese provinces: a fixed effects stochastic frontier spatial Durbin error panel analysis. *Ann Reg Sci* 58(2):301–319
- Kinfu Y, Sawhney M (2015) Inefficiency, heterogeneity and spillover effects in maternal care in India: a spatial stochastic frontier analysis. *BMC Health Serv Res* 15(1):118
- Kumbhakar SC, Knox Lovell CA (2000) *Stochastic frontier analysis*. Cambridge University Press, Cambridge
- Kumbhakar SC, Tsionas EG (2006) Estimation of stochastic frontier production functions with input-oriented technical efficiency. *J Econom* 133(1):71–96
- Landes DS (1998) *The wealth and poverty of nations: why some are so rich and some so poor*. W.W. Norton and Co., New York
- LeSage JP, Pace RK (2009) *Introduction to spatial econometrics*. Chapman & Hall, Boca Raton
- Meeusen W, van den Broek J (1977) Efficiency estimation from Cobb–Douglas production functions with composed error. *Int Econ Rev* 18(2):435–444
- Otsuka A (2017) Regional determinants of total factor productivity in Japan: stochastic frontier analysis. *Ann Reg Sci* 58(3):579–596
- Pavlyuk D (2010) Regional tourism competition in the Baltic states: a spatial stochastic frontier approach. MPRA Paper 25052, University Library of Munich, Germany
- Puig-Junoy J (2001) Technical inefficiency and public capital in US states: a stochastic frontier approach. *J Reg Sci* 41(1):75–96
- Puig-Junoy J, Pinilla J (2008) Why are some Spanish regions so much more efficient than others? *Environ Plan C Gov Policy* 26(6):1129–1142
- Resti A (1997) Evaluating the cost-efficiency of the Italian banking system: what can be learned from the joint application of parametric and non-parametric techniques. *J Bank Finance* 21(2):221–250
- Rodríguez-Pose A, Crescenzi R (2008) Research and development, spillovers, innovation systems, and the genesis of regional growth in Europe. *Reg Stud* 42(1):51–67
- Schmidt AM, Moreira ARB, Helfand SM, Fonseca TCO (2009) Spatial stochastic frontier models: accounting for unobserved local determinants of inefficiency. *J Prod Anal* 31:101–112
- Thissen M, Diodato D (2013a) Trade between European NUTS2 regions from 2000 to 2010. Technical report, The PBL Netherlands Environmental Assessment Agency, The Hague
- Thissen M, Diodato D (2013b) Trade between European NUTS2 regions in 2000. Technical report, The PBL Netherlands Environmental Assessment Agency, The Hague
- Thissen M, de Graaff T, van Oort F (2016) Competitive network positions in trade and structural economic growth: a geographically weighted regression analysis for European regions. *Pap Reg Sci* 95(1):159–180
- Tsionas EG, Michaelides PG (2016) A spatial stochastic frontier model with spillovers: evidence for Italian regions. *Scott J Polit Econ* 63(3):243–257
- Vidoli F, Cardillo C, Fusco E, Canello J (2016) Spatial nonstationarity in the stochastic frontier model: an application to the Italian wine industry. *Reg Sci Urb Econ* 61:153–164
- Wang HJ, Ho CW (2010) Estimating fixed-effect panel stochastic frontier models by model transformation. *J Econom* 157(2):286–296
- Wang WS, Schmidt P (2009) On the distribution of estimated technical efficiency in stochastic frontier models. *J Econom* 148(1):36–45
- Weinstein MA (1964) The sum of values from a normal and a truncated normal distribution. *Technometrics* 6(1):104–105