Proceedings of the 12th International Conference "Reliability and Statistics in Transportation and Communication" (RelStat'12), 17–20 October 2012, Riga, Latvia, p. 11–19. ISBN 978-9984-818-49-8
Transport and Telecommunication Institute, Lomonosova 1, LV-1019, Riga, Latvia

MAXIMUM LIKELIHOOD ESTIMATOR FOR SPATIAL STOCHASTIC FRONTIER MODELS

Dmitry Pavlyuk

Transport and Telecommunication Institute
Lomonosova 1, Riga, LV-1019, Latvia
Ph.: +371 29958338. E-mail: Dmitry.Pavlyuk@tsi.lv

This research is devoted to analysis of efficiency estimation in presence of spatial relationships and spatial heterogeneity in data. We presented a general specification of the spatial stochastic frontier model, which includes spatial lags, spatial autoregressive disturbances and spatial autoregressive inefficiencies. Maximum likelihood estimators are derived for two special cases of the spatial stochastic frontier. Small-sample properties of these estimators and comparison with a standard non-spatial estimator were implemented using a set of Monte Carlo experiments. Finally, we tested our estimators on a real-world data set of European airports and discovered significant spatial components in data.

Keywords: spatial stochastic frontier, maximum likelihood, efficiency, heterogeneity

Introduction

During the last three decades the theoretical and empirical interest to spatial aspects of the economy is significantly increased [1]. New economic geography postulates an important role of spatial interactions between economic units. Economic activity is heterogeneously distributed across space and its spatial structure is considerably determined by forces of spatial interaction like spatial competition or spatial conglomeration of businesses. Thus spatial relationships become an essential component of contemporary economic models, and spatial econometrics, which incorporates spatial effects into classical statistical approaches, is a thriving branch of modern econometrics [2].

Estimation of economic efficiency is an extensive research area, where econometric models are intensively used. Stochastic frontier model [3], one of the most popular approaches to empirical research of efficiency, is widely used. Nevertheless a standard form of the stochastic frontier model doesn't take spatial effects into account and a theory of its spatial modifications is weakly developed.

Influence of spatial interaction on units' efficiency is even more important factor for transport enterprises like airports, sea ports, and coach terminals. Increasing people mobility and enhancing road network not only increase transport terminals accessibility, but make them 'closer' to their neighbour competitors and intensify spatial effects. Spatial competition among terminals and heterogeneity of a spatial structure exert growing influence on their efficiency, and should be included into benchmarking approach [4], [5].

A specification of a spatial stochastic frontier model per se is insufficiently researched subject. In this research we consider a general form of the spatial stochastic frontier model with all types of spatial components – spatial lags, spatial autoregressive disturbances and spatial autoregressive inefficiencies.

Different econometric techniques (maximum likelihood, two-step least squares, general method of moments) can be adapted estimation of this model. In this article we develop a maximum likelihood estimator for different forms of the stochastic frontier model. Derived maximum likelihood estimator formulas are tested on the base of simulated data to discover its properties on limited samples.

Airport benchmarking attracted significant scientific and practical attention after industry liberalisation in the nineties [4], [6]. Spatial interactions between European airports and spatial heterogeneity of the industry structure are widely acknowledged and should be included into airport benchmarking techniques. In this research the spatial stochastic frontier model is applied to a data set of European airports and significant spatial effects are discovered.

Specifications of a Spatial Stochastic Frontier Model

A classical stochastic frontier model is usually presented in a matrix form as [3]:

$$\begin{aligned} &y = X\beta + \varepsilon, \\ &\varepsilon = v - u, \\ &u \geq 0, \end{aligned} \tag{1}$$

where
y is an (n x 1) vector of a dependent variable, output (n is a size of the sample);
X is an (n x k+1) matrix of explanatory variables, inputs (k is a number of explanatory variables);
β is a (1 x k+1) vector of unknown coefficients (model parameters);
ε is an (n x 1) vector of composite error terms;
v is an (n x 1) vector of independent identically distributed (i.i.d.) error terms;
u is an (n x 1) vector of inefficiency terms with non-negative values.

The classical stochastic frontier model doesn't include any spatial dependencies and assumes that all objects in a sample are independent. This assumption is too restrictive in some practical cases. Spatial effects can be presented almost in all components of the classical model:
- spatial influence of neighbours' output values on a given unit's output (spatial lags);
- spatial influence of neighbours' input values on a given unit's output;
- spatial relationship between neighbour unit's error terms (spatial heterogeneity);
- spatial correlation between efficiency of neighbour units.

We define a general spatial stochastic frontier model, including all these effects into the classical stochastic frontier model specification:

$$\begin{aligned} &y = \rho W_1 y + X\beta + \gamma W_2 X + \varepsilon, \\ &\varepsilon = v - u, \\ &v = \xi W_3 v + \tilde{v}, \\ &u = \eta W_4 u + \tilde{u}, u \geq 0, \end{aligned} \tag{2}$$

where
$W_1 y$ is a spatial lag vector of output values (with a coefficient ρ);
$W_2 X$ is a spatial inputs-output lag vector (with a coefficient γ);
$W_3 v$ are spatial errors (with a coefficient ξ);
$W_4 u$ are spatial inefficiency lags (with a coefficient η).

Matrices W_1, W_2, W_3, and W_4 represent levels of spatial dependency between units (spatial weights) and can be different for every spatial component. All spatial weights matrices have zero values on the main diagonal (to prevent self-dependency). Construction of these matrices is usually research-specific and can be based on geographical distances, travel times, etc.

Estimation of the general spatial stochastic frontier model's parameters is a complicated task, which is related with identification problems, computation performance issues and requires a significant volume of data. In this research we consider two special cases of the general spatial stochastic frontier model.

Let apply following constraints on the general spatial stochastic frontier model:

$$\gamma = 0, \tag{3}$$

$$\xi = 0, \tag{4}$$

$$\eta = 0. \tag{5}$$

Under these constraints the model includes only spatial lags for unit's outputs, all other spatial effects are excluded:

$$\begin{aligned} &y = \rho W_1 y + X\beta + \varepsilon, \\ &\varepsilon = v - u, \\ &v = \tilde{v}, \\ &u = \tilde{u}, u \geq 0. \end{aligned} \tag{6}$$

Following LeSage [7] notation for naming of spatial models, this model is a mixed first-order spatial autoregressive-regressive stochastic frontier model. We will refer this model as a spatial autoregressive stochastic frontier (SARSF) model.

Another model we consider in this research includes spatial relationship of a symmetric error term v as well as spatial lags. Suppressing the constraint (4), we obtain the following model specification:

$$\begin{aligned} &y = \rho W_1 y + X\beta + \varepsilon, \\ &\varepsilon = v - u, \\ &v = \xi W_2 v + \tilde{v}, \\ &u = \tilde{u}, u \geq 0. \end{aligned} \tag{7}$$

We will refer this model as a mixed first-order spatial autoregressive-regressive stochastic frontier model with spatial autoregressive disturbances (SARARSF).

Maximum likelihood estimator for spatial stochastic frontier models

A wide range of statistical methods is used for estimation of spatial model parameters. The most popular are maximum likelihood estimator [7], [8], two-step least squares [9], and generalised method of moments [10], [11]. In this research we derive maximum likelihood estimators for SARSF and SARARSF models, test them with Monte Carlo experiments and apply to a real dataset of European airports.

The maximum likelihood estimator requires an assumption about distributions of error and inefficiency components. The distribution of the symmetric error term v is usually set to normal, and the distribution of the non-negative inefficiency term u is selected from half-normal [12], truncated normal [13], or gamma [14]. We consider the simplest normal-half-normal type of the composite error term ε:

$$\begin{aligned} &v = \tilde{v} \sim N(0, \sigma_{\tilde{v}}^2 I), \\ &u = \tilde{u} \sim N^+(0, \sigma_{\tilde{u}}^2 I). \end{aligned} \tag{8}$$

The probability density function for this case is well known:

$$f(\varepsilon) = \frac{2}{\sigma}\varphi\left(\frac{\varepsilon}{\sigma}\right)\Phi\left(\frac{-\varepsilon\lambda}{\sigma}\right), \tag{9}$$

where

$$\sigma = \sqrt{\sigma_{\tilde{u}}^2 + \sigma_{\tilde{v}}^2}, \lambda = \frac{\sigma_{\tilde{u}}}{\sigma_{\tilde{v}}},$$

φ and Φ are standard normal probability density and cumulative distribution functions accordingly.

Maximum Likelihood Estimator for the SARSF Model

Derivation of the maximum likelihood estimator formula for the SARSF model (6) is quite straightforward. According to the model specification, the composite error terms vector can be expressed as:

$$\varepsilon = (I - \rho W_1)y - X\beta.$$

Using the multivariate change of variables formula and the Jacobian matrix, we can produce the probability density function for y:

$$f(y) = f(\varepsilon)\det\left(\frac{\partial\varepsilon}{\partial y}\right) = \frac{2}{\sigma}\varphi\left(\frac{\varepsilon}{\sigma}\right)\Phi\left(\frac{-\varepsilon\lambda}{\sigma}\right)\det(I - \rho W_1). \tag{10}$$

Now the log-likelihood function can be easily obtained:

$$LogL(\rho, \beta, \sigma, y, X, W_1) = C - n\ln(\sigma) - \frac{\varepsilon^T\varepsilon}{2\sigma^2} + \sum_{i=1}^{n}\ln\left(\Phi\left(\frac{-\varepsilon\lambda}{\sigma}\right)\right) + \ln(\det(I - \rho W_1)). \tag{11}$$

Maximum Likelihood Estimator for the SARARSF Model

Derivation of the maximum likelihood estimator formula for the SARARSF model (7) requires some technical efforts. We follow the procedure, described in Kumbhakar and Lovell [3] to produce the probability density function and likelihood function.

Initial model specification includes the distribution of the error term in an implicit form:

$$v = \xi W_2 v + \tilde{v},$$

$$\tilde{v} \sim N(0, \sigma_{\tilde{v}}^2 I).$$

Straightforward transformations give us:

$$v \sim N(0, \Sigma), \text{where } \Sigma = \sigma_{\tilde{v}}^2 \tilde{\Sigma} = \sigma_{\tilde{v}}^2 \left(\left(I - \xi W_2 \right)^{-1} \right)^T \left(I - \xi W_2 \right)^{-1}. \quad (12)$$

So the error term has a multivariate normal distribution with a covariance matrix Σ and its respective probability density function is given as:

$$f(v) = \left(\frac{1}{\sqrt{2\pi}} \right)^n \cdot \left(\det(\Sigma) \right)^{-\frac{1}{2}} \cdot \exp\left\{ -\frac{1}{2} v^T \Sigma^{-1} v \right\}. \quad (13)$$

The half-normal probability density function is given as:

$$f(u) = \begin{cases} \left(\frac{2}{\sigma_{\tilde{u}} \sqrt{2\pi}} \right)^n \cdot \exp\left\{ -\frac{u^T u}{2\sigma_{\tilde{u}}^2} \right\}, u \geq 0 \\ 0, u < 0 \end{cases}. \quad (14)$$

We omit the derivation details of the composite error term's probability density. Finally the probability density function of the composite error term ε is obtained as:

$$f(\varepsilon) = 2^n \left(\Phi(\mathbf{0}, \mu\varepsilon, \Omega) \right) \varphi(\varepsilon, \mathbf{0}, \Theta), \quad (15)$$

where $\varphi(x, mean, covariance)$ and $\Phi(x, mean, covariance)$ are multivariate probability density and cumulative distribution functions with a mean vector *mean* and a covariance matrix *covariance*,

$$\Theta = \sigma_{\tilde{u}}^2 I + \sigma_{\tilde{v}}^2 \tilde{\Sigma},$$

$$\Omega = \sigma_{\tilde{u}}^2 \sigma_{\tilde{v}}^2 \tilde{\Sigma} \Theta^{-1},$$

$$\mu = -\sigma_{\tilde{v}}^{-2} \Omega \tilde{\Sigma}^{-1}.$$

This finding conforms to the results presented in [15].
Using the density function (16) and following the same logic as in (10), we obtained the respective log-likelihood function for the SARARSF model:

$$LogL(y \mid \beta, \sigma_{\tilde{v}}, \sigma_{\tilde{u}}, \rho) = C + \ln\left(\Phi(\mathbf{0}, \mu\varepsilon, \Omega) \right) + \ln\left(\varphi(\varepsilon, \mathbf{0}, \Theta) \right) + \ln\left(\det\left(I - \rho W_1 \right) \right), \quad (16)$$

Maximisation of the log-likelihood function with respect to its parameters is a separate computational problem.

Monte Carlo experiments

In order to discover additional information about properties of the derived estimators on limited samples, we organised two Monte Carlo studies, completely based on artificial data. The first study is devoted to comparison of the classical stochastic frontier and the SARSF model in presence of spatial lags. The second study is used to compare SARSF and SARARSF model estimators in presence of both spatial lags and spatial errors.

Both studies are based on a series of 100 experiments and a generated sample of 100 units. The sample size was chosen for practical purposes – application to a set of international airports in Europe, which have approximately this size.

Spatial settings were modelled as a "rook" spatial weights matrix for spatial lags (W_1) and as a "queen" spatial weights matrix for spatial errors (W_2).

Study 1. Stochastic frontier without spatial components vs. the SARSF model

The first study was repeated 7 times for different values of ρ: 0, ± 0.1, ± 0.5, ± 0.9. Higher values of ρ correspond to stronger spatial effects.

The results of the first study are presented in the Table 1.

Table 1. Simulation results for the classical stochastic frontier (SF) and the spatial autoregressive stochastic frontier (SARSF) estimators.*

	β_0	β_1	β_2	β_3	σ_v	σ_u	ρ
Real values	*10*	*1*	*-2*	*3*	*20*	*100*	
							0
SF	1.769 (1.3E-04)	1.002 (7.0E-01)	-2.001 (7.6E-01)	3.001 (8.9E-01)	21.937 (2.4E-01)	91.544 (1.1E-03)	-
SARSF	2.338 (5.0E-04)	0.996 (4.8E-01)	-2.004 (4.7E-01)	2.994 (2.7E-01)	19.676 (8.5E-01)	89.465 (5.5E-05)	0.004 (1.1E-01)
							0.1
SF	-10.774 (9.5E-13)	1.001 (9.1E-01)	-2.008 (1.1E-01)	3.013 (1.2E-02)	35.581 (1.2E-13)	82.302 (1.1E-07)	-
SARSF	-1.321 (8.3E-06)	1.006 (2.5E-01)	-2.002 (7.3E-01)	2.996 (5.4E-01)	23.039 (9.0E-02)	86.259 (2.2E-05)	0.103 (1.7E-01)
							-0.1
SF	4.527 (3.2E-02)	0.998 (6.6E-01)	-2.005 (3.1E-01)	3.011 (3.1E-02)	32.342 (3.1E-10)	86.649 (5.1E-05)	-
SARSF	-2.079 (3.0E-07)	0.995 (3.4E-01)	-2.005 (3.6E-01)	3.001 (8.8E-01)	23.242 (8.8E-02)	86.277 (2.7E-06)	-0.098 (5.1E-01)
							0.5
SF	-92.485 (3.0E-34)	1.054 (8.2E-05)	-2.138 (9.2E-15)	3.225 (8.1E-26)	142.612 (4.0E-71)	40.556 (2.9E-39)	-
SARSF	-11.746 (6.0E-09)	0.999 (8.5E-01)	-1.990 (3.5E-02)	3.001 (8.0E-01)	28.926 (2.0E-04)	71.537 (1.9E-09)	0.499 (5.2E-01)
							-0.5
SF	4.313 (2.4E-01)	1.065 (1.3E-05)	-2.181 (2.6E-20)	3.269 (4.9E-32)	146.612 (3.1E-71)	46.783 (3.5E-32)	-
SARSF	1.561 (9.4E-05)	0.999 (9.0E-01)	-1.996 (4.4E-01)	3.002 (7.7E-01)	21.000 (5.2E-01)	90.037 (1.5E-04)	-0.498 (3.3E-01)
							0.9
SF	-636.538 (1.5E-30)	1.442 (1.0E-11)	-2.936 (8.8E-26)	4.347 (3.6E-38)	545.887 (5.3E-57)	148.401 (5.8E-04)	-
SARSF	1.269 (6.9E-03)	0.999 (8.8E-01)	-1.996 (5.2E-01)	2.988 (3.6E-02)	20.292 (8.9E-01)	88.623 (9.7E-03)	0.892 (1.5E-02)
							-0.9
SF	91.309 (2.3E-10)	1.668 (1.1E-17)	-3.221 (6.6E-31)	4.727 (4.5E-41)	686.998 (2.3E-66)	156.142 (6.6E-06)	-
SARSF	11.080 (5.6E-01)	0.995 (3.1E-01)	-2.005 (3.6E-01)	3.007 (2.2E-01)	10.839 (1.4E-12)	108.119 (2.3E-02)	-0.898 (1.8E-02)

* Values in brackets are p-values for Student's test of a hypothesis "an average estimate equals to a corresponding real value"

The classical stochastic frontier without spatial components estimator became biased in presence of spatial lags (as expected), and the bias becomes significant for higher values of ρ. Sensitivity of the model's β parameters to changes of spatial settings is relatively lower – almost no bias for $\rho = 0.1$ (although the bias should be proved theoretically). Estimates of the standard deviations σ_v and σ_u are more susceptible to spatial effects. As these parameters are critical to estimation of inefficiency levels, this finding is important for understanding of unrecorded spatial effects on estimated efficiency values. According to simulation results, a standard deviation of the error term v is overestimated, and a standard deviation of the inefficiency term u is underestimated. This leads to a false conclusion about lower inefficiency values of data units in presence of spatial effects.

The SARSF estimator identifies spatial lags and other data generation process parameters correctly.

Study 2. The SARSF model vs. the SARARSF model

The second study was repeated 3 times for different values of parameters ρ (spatial lags) and ξ (spatial errors).

The results of this study are presented in the Table 2.

Table 2. Simulation results for the spatial autoregressive stochastic frontier (SARSF) and the spatial autoregressive stochastic frontier with spatial autoregressive disturbances (SARARSF) estimators.*

	β_1	β_2	β_3	ρ	ξ
Real values	*1*	*-2*	*3*		
				0.5	*0*
SARSF	1.000 (9.8E-01)	-2.007 (1.4E-01)	2.992 (1.3E-01)	0.500 (8.2E-01)	-
SARARSF	0.995 (4.5E-01)	-1.992 (1.7E-01)	2.988 (7.0E-02)	0.488 (6.9E-04)	0.091 (2.6E-02)
				0	*0.5*
SARSF	1.021 (3.4E-01)	-1.993 (6.5E-01)	2.989 (4.7E-01)	0.040 (1.1E-02)	-
SARARSF	0.998 (9.2E-01)	-2.014 (2.0E-01)	3.008 (5.0E-01)	0.012 (3.2E-02)	0.561 (1.2E-01)
				0.5	*0.5*
SARSF	0.993 (2.3E-01)	-1.993 (2.0E-01)	2.992 (1.4E-01)	0.518 (8.2E-07)	-
SARARSF	0.998 (7.4E-01)	-2.000 (9.3E-01)	2.995 (4.1E-01)	0.501 (7.8E-01)	0.469 (4.5E-02)

* Values in brackets are p-values for the Student test of a hypothesis "an average estimate equals to a corresponding real value"

Application to Airports

We applied both SARSF and SARARSF specifications of the spatial stochastic frontier model to a data set of European airports.

Significant demand for airports benchmarking attracted academic researchers to this problem. There are more than a hundred research papers, published during last two decades, and devoted to airports efficiency estimation. The most significant reports are the Global Airport Performance Benchmarking Reports produced by Air Transport Research Society [16], the Airport Performance Indicators and Review of Airport Charges reports by Jacobs Consulting, the Airport Service Quality programme by Airports Council International. Some local authorities which control the airport sector also provide their own benchmarking reports, e.g. Civil Aviation Authority (UK) [17], and others. Many related researches are also executed within the bounds of the German Airport Performance research project, a joint study between three German universities.

Despite intensified research of efficiency in the airport industry and obvious geographical issues, spatial effects are rarely included into consideration [6].

There are many different approaches to understanding of airport business, a set of used resources and a output of airport activity [18]. In this research we used a number of transferred passengers of a main result of airport activity and infrastructure units (gateways, check-ins) as airport resources. The production function used in this research is estimated in the form:

$$\begin{aligned} &\ln(Passengers) = \beta_0 + \rho W \ln(Passengers) + \beta_1 \ln(Runways) + \beta_3 \ln(CheckIns) + v - u, \\ &v = \xi W v + \tilde{v}, \tilde{v} \sim N(0, \sigma_{\tilde{v}}^2 I), \\ &u \sim N^+(0, \sigma_{\tilde{u}}^2 I), \end{aligned} \tag{17}$$

where

Passengers is a total number of passengers carried by an airport (both departure and arrival);
Runways and *CheckIns* are numbers of airport's runways and check-ins respectively.

Other airport infrastructure characteristics are excluded due to a multicollinearity problem.

A matrix of spatial weights W was constructed on the base of Euclidean distances. Realising all shortcomings of this approach, we think that general influence of spatial effects will be estimated correctly.

Spatial relationships and spatial heterogeneity in airport industry

Presence of both spatial relationships and spatial heterogeneity in the airport industry is widely acknowledged in literature (see [6] for a related literature review). Spatial relationships are usually presented in a form of spatial competition. Theory of spatial competition is well developed, but rarely

applied to airports. Scientific and government interest to spatial competition among airports was raised after industry liberalisation in earlier nineties. Taking spatial interactions between airports into account is acknowledged to be critically important for estimation of airports efficiency levels. A wide range of instruments like overlapping catchment areas, a network connectivity index and other were suggested, but the spatial econometric models can be highlighted as the most theoretically supported methodology.

Spatial heterogeneity is a very import drawback in airport efficiency research. Region-specific settings can significantly affect airports activity, but their inclusion into a model is not straightforward. There are some sources of regional heterogeneity of airport activity:

- Climate. Airport activity can be significantly affected by a climate. For example, snow-belt airports have to spend additional efforts on runway service, which reduce their production in relation with airports located in a region with softer weather conditions. Mapping this difference out the model will lead to incorrect (underestimated) values of snow-belt airports' efficiency.
- Economics. Economic situation in European countries is very heterogeneous. Significantly different income per capita and price levels define different demand to air transport.
- Region attractiveness. Regions also are not equal in relation to their demand for air transport. Business activities, required air flights, tourism attractiveness significantly vary across Europe.

A standard approach to include these factors into the model is based on a set of region-specific dummy variables, and looks weak in case of a complex spatial structure. Spatial effects are usually not limited with country borders, so using of an administrative division in this case is not well-grounded. Nevertheless, spatial structure should be included into efficiency estimation to prevent a bias in efficiency estimates.

Data

The study data set includes characteristics of 122 European airports in 2009. The characteristics include:

- A number of passengers carried (direct transit passengers are excluded). This indicator is used as the main output of airport's activity.
- Airport infrastructure – a number of check-in facilities, gates, runways, and parking spaces are used as input resources of airports' activity.

Estimation results

Six different model specifications were tested for the data set:

- classical regression (OLS estimates);
- classical stochastic frontier (SF estimates);
- classical spatial lag model (SAR estimates);
- classical spatial error model (SEM estimates);
- spatial autoregressive stochastic frontier (SARSF estimates);
- spatial autoregressive stochastic frontier with spatial autoregressive disturbances (SARARSF estimates).

The results are presented in the Table 3.

Table 3. Estimation results of different model specifications*

	OLS	SF	SAR	SEM	SARSF	SARARSF
β_0	12.073 (<2E-16)	12.498 (<2E-16)	6.980 (6.9E-02)	12.092 (<2E-16)	12.50 (<2E-16)	12.120 (<2E-16)
β_1	0.658 (6.3E-06)	0.669 (7.3E-07)	0.649 (1.3E-06)	0.658 (7.7E-07)	0.668 (9.4E-07)	0.653 (5.7E-07)
β_2	0.813 (<2E-16)	0.785 (<2E-16)	0.803 (<2E-16)	0.809 (<2E-16)	0.786 (<2E-16)	0.796 (<2E-16)
σ_v	0.514	0.549	0.501	0.501	0.444	0.494
σ_u	-	0.248	-	-	0.400	0.005
ρ	-	-	0.331 (2.1E-01)	-	0.0001 (9.4E-01)	0.0002 (9.0E-01)
ξ	-	-	-	0.480 (2.6E-01)	-	0.132 (4.7E-08)

* Values in brackets are significance

Standard SF estimates show significant inefficiency in data (inefficiency variance $\sigma_u = 0.248$, which is comparable with the error term variance σ_u). SARSF model supports this conclusion, showing insignificant spatial lags and significant inefficiency. But taking spatial heterogeneity into account (SARARSF model) produces alternative results. The model demonstrates significant spatial heterogeneity in data (which is expected for airport industry), and this spatial component supplants inefficiency – it becomes insignificant. This result is expected due to incompleteness of the production function components (only infrastructure metrics are included), but from our opinion the general findings are very important for analysis. Generally, imperfection of the production frontier can produce incorrect conclusions about inefficiency term, when including spatial components into the model can partly improve the situation.

Estimates of input elasticises (coefficient β_1 for a number of runways and coefficient β_2 for a number of check-ins) are quite stable for all models, which is quite reasonable in absence of significant spatial lags.

Conclusions

We presented a general specification of the spatial stochastic frontier model which includes spatial lags, spatial autoregressive disturbances and spatial autoregressive inefficiencies. These spatial components describe present spatial relationships and spatial structure of study units and critically important for econometric models, specifically for efficiency estimation.

Maximum likelihood estimators are derived for two specific cases of the spatial stochastic frontier – a spatial autoregressive stochastic frontier and spatial autoregressive stochastic frontier with spatial autoregressive disturbances. Small sample properties of the derived estimates are investigated with a set of Monte Carlo experiments. We discovered that for a sample of size 100 both estimators provide stable correct results for unknown parameters of a production function. Further small-sample studies are necessary for more definite conclusion about the estimators. Additional research is necessary to assess the influence of spatial components on efficiency estimates and investigate properties of the presented maximum likelihood estimators.

We tested our estimators on a real-world data set of European airports. We discovered significant spatial components in data, which proves our initial hypothesis about presence of spatial relationships and spatial heterogeneity in airport industry.

References

1. Krugman, P. R., Fujita, M., Venables, A. (2001). *The spatial economy: cities, regions and international trade*. Cambridge, Mass.: MIT Press.
2. Anselin, L. (2010). Thirty years of spatial econometrics. *Papers in Regional Science*, 89(1), 3–25.
3. Kumbhakar, S. C., Lovell, C. A. K. (2003). *Stochastic frontier analysis*. Cambridge: Cambridge Univ. Press.
4. Adler, N., Liebert, V. (2011). Competition and regulation (when lacking the former) outrank ownership form in generating airport efficiency. In the GAP workshop "Benchmarking of Airports," November 25–26, 2010 (p. 27). Berlin, Germany: The German Airport Performance (GAP).
5. Scotti, D. (2011). *Measuring Airports' Technical Efficiency: Evidence from Italy*. PhD thesis, University of Bergamo, Italy.
6. Pavlyuk, D. (2012). Airport Benchmarking and Spatial Competition: A Critical Review. *Transport and Telecommunication*, 13(2), 123–137.
7. LeSage, J. P. (1999). *Spatial econometrics*. West Virginia University: Regional Research Institute.
8. Anselin, L. (1988). *Spatial econometrics: methods and models*. Dordrecht: Kluwer Academic Publishing.
9. Kelejian, H. H., Prucha, I. R. (1998). A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17(1), 99–121.
10. Druska, V., Horrace, W. C. (2004). Generalized moments estimation for spatial panel data: Indonesian rice farming. *American Journal of Agricultural Economics*, 86(1), 185–198.
11. Lee, L. (2007). GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *Journal of Econometrics*, 137(2), 489–514.
12. Aigner, D., Lovell, C., Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of econometrics*, 6(1), 21–37.

13. Stevenson, R. (1980). Likelihood Functions for Generalized Stochastic Frontier Functions. *Journal of Econometrics*, 13, 57–66.
14. Greene, W. H. (1990). A gamma-distributed stochastic frontier model. *Journal of econometrics*, 46(1–2), 141–163.
15. Lin, J., Long, Z., Lin, K. (2010). Spatial Panel Stochastic Frontier Model and Technical Efficiency Estimation. *Journal of Business Economics*, 223(5), 71–78.
16. Oum, T. H., Yu, C., Choo, Y. (2011). *ATRS Global Airport Performance Benchmarking Project.* Vancouver, Canada: University of British Columbia.
17. Civil Aviation Authority. (2000). *The Use of Benchmarking in the Airport Reviews.* London, UK: Civil Aviation Authority.
18. Doganis, R. (1992). *The airport business*. London: Routledge.