Модель производственной границы с пространственными эффектами: процесс исследования

Spatial stochastic frontier model: research process

Dmitry Pavlyuk

TSI TRANSPORT AND TELECOMMUNICATION INSTITUTE

Saratov, 09.11.2017

About myself

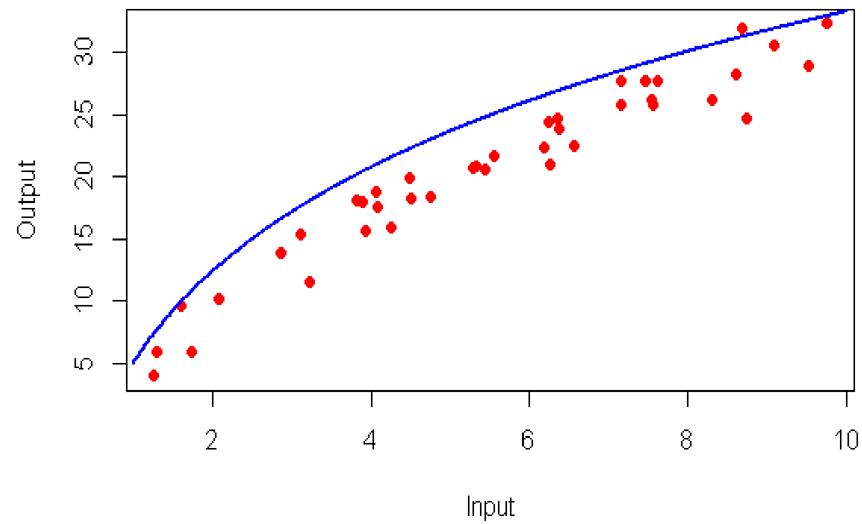
- 2002 – MSc, Information systems (in economics), Saratov State University
- 2005 – Candidate of economic sciences, “Statistical analysis of banks’ efficiency in Russia”, supervisor: prof. Vladimir Balash
- 2015 – Doctor of engineering sciences, “Spatial stochastic frontier analysis of European airports”, Riga, Latvia
- Now – professor and researcher of Transport and Telecommunication Institute, Riga, Latvia

1. Problem statement

Classical Stochastic Frontier model

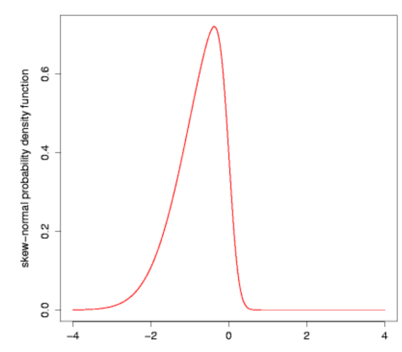
A classical stochastic frontier model (matrix form)

$$Y = X\beta + v - \underline{u}$$



where

- Y is a vector of outputs,
- X is a matrix of inputs,
- v is a vector of random disturbances,
- u is a vector of inefficiency terms,
- β is a vector of unknown coefficients.



Airport benchmarking: origins

- The airport business since 1990-ties:
  from **natural monopolies** to a **competitive market**

⇨

- Increased interest to airport performance and efficiency

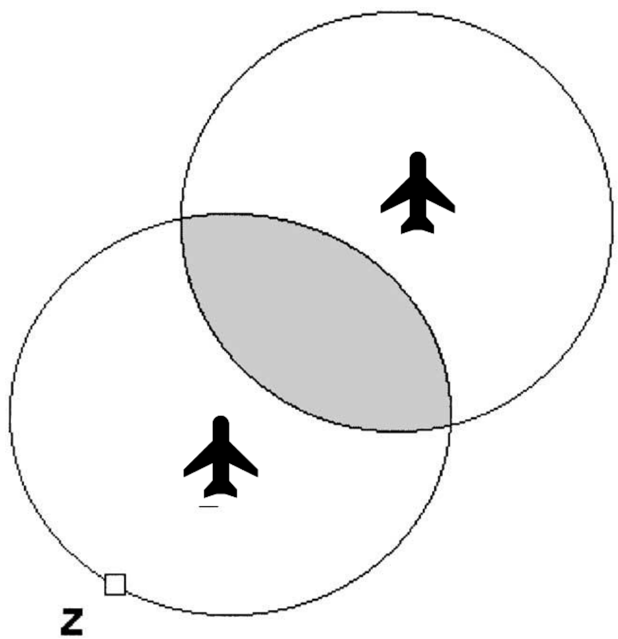
⇨

- ~200 recently published academic and empirical researches, devoted to airport benchmarking.

Airport competition: catchment areas

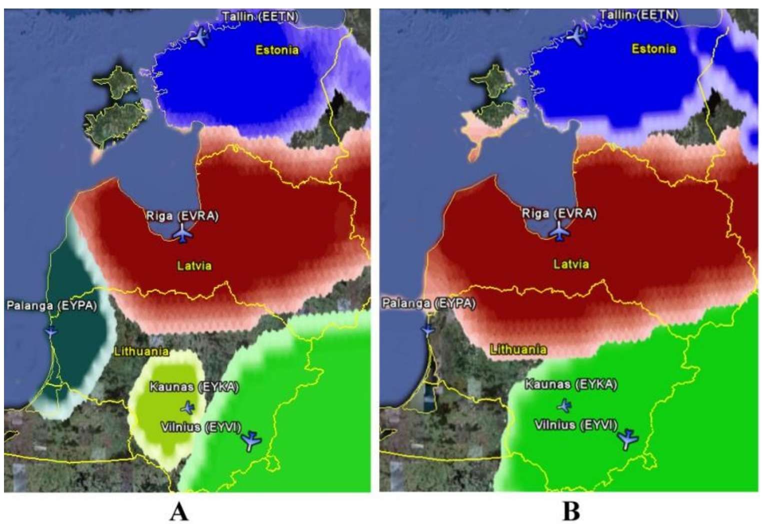
Theory of spatial competition since Hotteling, 1929.

For airports - based on catchment areas (Starkie, 2002):



Airport competition: catchment areas

Pavlyuk, D., 2009. Spatial Competition Pressure as a Factor of European Airports' Efficiency. Transport and Telecommunication 10 (4), 8–17.



Competition pressure estimated on the base of catchment areas intersection and population.

Spatial models

- Local: geographically weighted regression

1998 – Fotheringham et al., Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*. 30 (11): 1905–1927.

- Global: spatial regression models

1988 – Luc Anselin, Spatial econometrics: methods and models, Springer Science & Business Media

Spatial lags

Tobler's Law: "everything is related to everything else, but near things are more related than distant things".

Simplest form of spatial dependency is a linear :

$$y_i \text{ depends on } \sum_{i \neq j} w_{ij} y_j$$

Coefficients w_{ij} can be explained with a distance between objects *i* and *j*, e.g.

$$w_{ij} = \frac{1}{\mathrm{distance}(object_i, object_j)}$$

Spatial regression

Spatial autoregressive model:

$$Y = \rho WY + X\beta + \varepsilon$$

where W – a spatial contiguity matrix (spatial weights).

- ✓ Балаш В.А., Балаш О.С., Харламов А.В. Эконометрический анализ геокодированных данных о ценах на жилую недвижимость, Прикладная эконометрика, 22 (2), 2011

Spatial effects: types

Type 1. **Spatial dependency**:
a relationship between outputs of an airport and outputs of its neighbours

Type 2. **Exogenous spatial effects**:
a relationship between an output of an airport and inputs (resources) of its neighbours

Type 3. **Spatially heterogeneity**: uneven distribution of unobserved influencing factors over the space

Type 4. **Spatially related efficiency**: a relationship between efficiency of neighbour airports

2. Spatial Stochastic frontier model

Spatial Stochastic Frontier model: specification

The proposed SSF model specification:

$$\mathbf{Y} = \underline{\rho_Y \mathbf{W}_Y \mathbf{Y}} + \mathbf{X\beta} + \underline{\mathbf{W}_X \mathbf{X} \beta^{(s)}} + \mathbf{v} - u,$$

$$\mathbf{v} = \underline{\rho_v \mathbf{W}_v \mathbf{v}} + \tilde{\mathbf{v}},$$

$$u = \underline{\rho_u \mathbf{W}_u u} + \tilde{u}.$$

where

- W_Y and ρ_Y are contiguity matrix and coefficient for endogenous spatial effects (spatial dependency),
- W_X and $\beta^{(s)}$ are contiguity matrix and coefficients for exogenous spatial effects,
- W_v and ρ_v are contiguity matrix and coefficients for spatially correlated random disturbances (spatial heterogeneity),
- W_u and ρ_u are contiguity matrix and coefficients for spatially related inefficiency.

Estimation of SSF model parameters: composed error term

- Problem statement
- **SSF model**
- The "spfrontier" package
- Application to European airports

Theorem 1.

Let we have two independent multivariate random variables:

$v = (v_1, v_2, \ldots, v_n)$ with the multivariate normal (MVN) distribution with a zero mean and a covariance matrix Σ_v,

$u = (u_1, u_2, \ldots, u_n)$ with the multivariate truncated normal distribution with a mean μ and a covariance matrix Σ_u and $(0, +\infty)$ truncation interval,

Then an *n*-variate random variable $\varepsilon = v - u$ has the closed skew normal (CSN) distribution:

$$\varepsilon \sim CSN_{n,n}\left(\mu', \Sigma', \Gamma', \nu', \Delta'\right),$$

and the probability density function of ε is:

$$f_\varepsilon(\varepsilon) = \left[\Phi_n\left(0, -\mu, \Sigma_u\right)\right]^{-1} \Phi_n\left(-\Sigma_u\left(\Sigma_v + \Sigma_u\right)^{-1}(\varepsilon + \mu), -\mu, \left(\Sigma_v^{-1} + \Sigma_u^{-1}\right)^{-1}\right) \times$$
$$\times \varphi_n\left(\varepsilon, -\mu, \Sigma_v + \Sigma_u\right),$$

where φ_n and Φ_n are standard MVN p.d.f. and c.d.f.

Estimation of SSF model parameters: MLE

The derived log-likelihood function:

$$\ln L\left(\beta, \beta^{(s)}, \sigma_{\tilde{v}}^2, \sigma_{\tilde{u}}^2, \mu, \rho_Y, \rho_v, \rho_u\right) = -\ln \Phi_n\left(0, -\mu, \Sigma_u\right) +$$
$$+\ln \Phi_n\left(-\Sigma_u\left(\Sigma_v + \Sigma_u\right)^{-1}(e+\mu), -\mu, \left(\Sigma_v^{-1} + \Sigma_u^{-1}\right)^{-1}\right) +$$
$$+\ln \varphi_n\left(e, -\mu, \Sigma_v + \Sigma_u\right),$$

$$e = Y - \rho_Y W_Y Y - X\beta - W_X X\beta^{(s)},$$

$$\Sigma_v = \sigma_{\tilde{v}}^2\left(\left(I_n - \rho_v W_v\right)^{-1}\right)^T \left(I_n - \rho_v W_v\right)^{-1},$$

$$\Sigma_u = \sigma_{\tilde{u}}^2,\left(\left(I_n - \rho_u W_u\right)^{-1}\right)^T \left(I_n - \rho_u W_u\right)^{-1}.$$

Estimation of individual efficiency values

The derived conditional distribution:

$$u|\varepsilon \sim MVTN_{0,+\infty}\left(\mu_{u|\varepsilon}, \Sigma_{u|\varepsilon}\right),$$

where

$$\mu_{u|\varepsilon} = \mu - \Sigma_u\left(\Sigma_v + \Sigma_u\right)^{-1}\left(\varepsilon + \mu\right),$$

$$\Sigma_{u|\varepsilon} = \left(\Sigma_v^{-1} + \Sigma_u^{-1}\right)^{-1}$$

Given the distribution, efficiency values are estimated and their confidence intervals are constructed.

3. SSF model implementation: “spfrontier” package

“spfrontier” package

The derived procedures for the SSF model are implemented as a package for CRAN R software.

Official “spfrontier” package page:

http://cran.r-project.org/web/packages/spfrontier

spfrontier: Spatial Stochastic Frontier models estimation

A set of tools for estimation of various spatial specifications of stochastic frontier models

Version:	0.1.12
Depends:	R (≥ 3.0), moments, ezsim, tmvtnorm, mvtnorm, maxLik
Imports:	methods, parallel, spdep
Published:	2014-12-21
Author:	Dmitry Pavlyuk
Maintainer:	Dmitry Pavlyuk <Dmitry.V.Pavlyuk at gmail.com>
License:	GPL-2 | GPL-3 [expanded from: GPL (≥ 2)]
NeedsCompilation:	no
Materials:	NEWS
CRAN checks:	spfrontier results

Package validation

1. Comparison with classical software (G. Battese, T.Coelli) for non-spatial model specifications

   Results: almost perfect match

2. Monte-Carlo Simulations

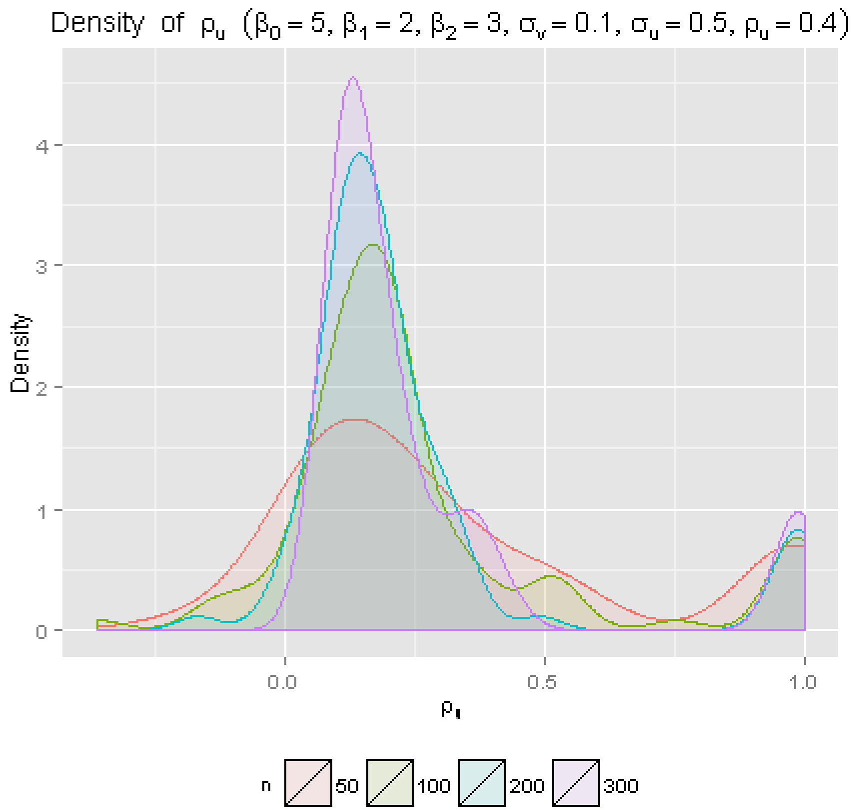
   9 designed experiments:

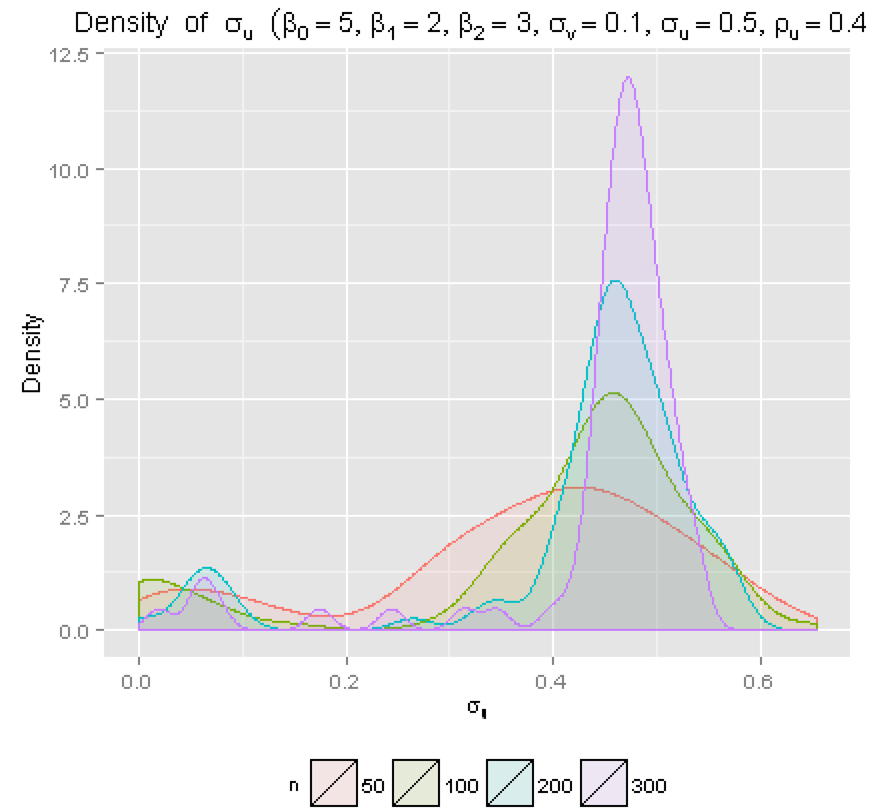
   3 sample sizes (50, 100, 200, and 300 objects) and 100 simulation runs for each.

Problems

- Computational difficulty
  - Solved by AWS cloud
- Discretization error (for more than 1000 observations)
  - Not solved
- Convergence
  - No theoretical convergence of likelihood optimization routine (proved by Azzalini)

Simulation tests: estimates' empirical kernel densities





4. Application of spatial analysis to European airports

Airport benchmarking: approaches

Economic efficiency as usage of available resources (inputs) to maximise the production of goods and services (outputs)

Airport Outputs:

- Passengers, cargo and transport movements (physical approach)
- Profit, total revenue (intermediary approach)

Airport Inputs:

- Infrastructure, employment, location, etc.

Research data sets

- European airports data set, 359 European airports, 2008-2012;
- Spanish airports data set, 38 Spanish airports, 2009-2010;
- UK airports data set, 48 UK airports, 2011-2012;
- Greek airports data set, 42 Greek airports, 2007.

Research model specifications: hierarchy

general to private

Classical Regression (OLS)

$$Y = X\beta + v$$

Spatial Autoregressive Model

$$Y = \rho_Y WY + X\beta + v$$

Spatial Error Model (SEM)

$$Y = X\beta + v$$
$$v = \rho_v Wv + \tilde{v}$$

Stochastic Frontier Model (SF)

$$Y = X\beta + v - u$$

SSF (1,0,0,0) Model

$$Y = \rho_Y WY + X\beta + v - u$$

SSF (0,0,1,0) Model

$$Y = X\beta + v - u,$$
$$v = \rho_v Wv + \tilde{v}$$

SSF (0,0,0,1) Model

$$Y = X\beta + v - u$$
$$u = \rho_u Wu + \tilde{u}$$

SSF (1,0,1,0) Model

$$Y = \rho_Y WY + X\beta + v - u,$$
$$v = \rho_v Wv + \tilde{v}$$

SSF (1,0,0,1) Model

$$Y = \rho_Y WY + X\beta + v - u,$$
$$u = \rho_u Wu + \tilde{u}$$

SSF (0,0,1,1) Model

$$Y = X\beta + v - u,$$
$$v = \rho_v Wv + \tilde{v},$$
$$u = \rho_u Wu + \tilde{u}$$

SSF (1,0,1,1) Model

Application to airports: different environments

Data Set	Significant inefficiency	Significant spatial effects
Data set 1: European airports	Yes	Yes
Data set 2: Spanish airports	No	Yes
Data set 3: UK airports	Yes	Yes
Data set 4: Greek airports	Yes	No

Developed methods work in different environments.

Related Publication

D. Pavlyuk. "Implication of spatial heterogeneity for airports' efficiency estimation", *Research in Transportation Economics*, Vol. 56. 2016, pp. 15 - 24.

ELSEVIER

Research in Transportation Economics

Volume 56, August 2016, Pages 15-24

Implication of spatial heterogeneity for airports' efficiency estimation

Dmitry Pavlyuk

Show more

https://doi.org/10.1016/j.retrec.2016.07.002

Get rights and content

Abstract

This research is devoted to the analysis of a role of spatial heterogeneity in estimation of airports' efficiency. Spatial heterogeneity, caused by uneven distribution of influencing factors over space, is widely acknowledged in the European airport industry, but rarely included in airport benchmarking procedures. We utilise modern methods of spatial econometrics to identify the importance of spatial heterogeneity for estimates of airports' resource elasticity and efficiency values. A set of utilised models includes the spatial error model, spatial stochastic frontier model, geographically weighted regression model, and their modifications.

Conclusions

- Main stages:
  - Spatial stochastic frontier model specification
  - Derivation of the MLE estimator
  - Implementation as “spfrontier’ package for R
  - Application to European airports

- Main problems:
  - Extensive multivariate distributions and matrix algebra
  - Re-inventing the wheel
  - Computational complexity
  - Convergence and publishing

Thank you for your attention!

Questions are very appreciated

Contacts:

email: Dmitry.Pavlyuk@tsi.lv

phone: +37129958338

European airports: model

- Problem statement
- SSF model
- The "spfrontier" package
- **Application to European airports**

Frontier specification:

$$\log(PAX) = \beta_0 + \rho_Y W \log(PAX) + \beta_1 \log(Routes) + \\ + \beta_2 \log(Population100km) + \beta_3 \log(GDPpc)$$

Estimation results:

Model		β_0	β_1	β_2	β_3	σ_v	σ_u	ρ_Y	ρ_v
SSF	Estimate	12.199	0.068	1.091	-0.182	0.557	1.087	-0.001	0.043
(1,0,1,0)	Std. error	1.390	0.045	0.034	0.125	0.053	0.102	0.001	0.000
	Sig.	$< 10^{-16}$	0.035	$< 10^{-16}$	0.262	$< 10^{-16}$	$< 10^{-16}$	0.016	$< 10^{-16}$
	Likelihood								-444.253