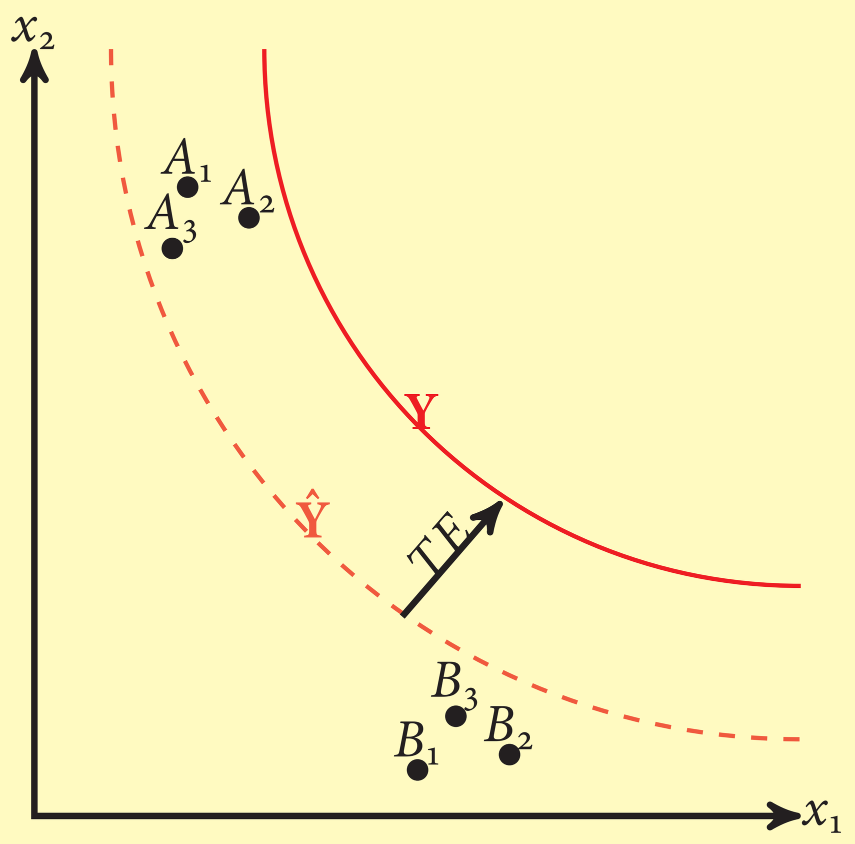
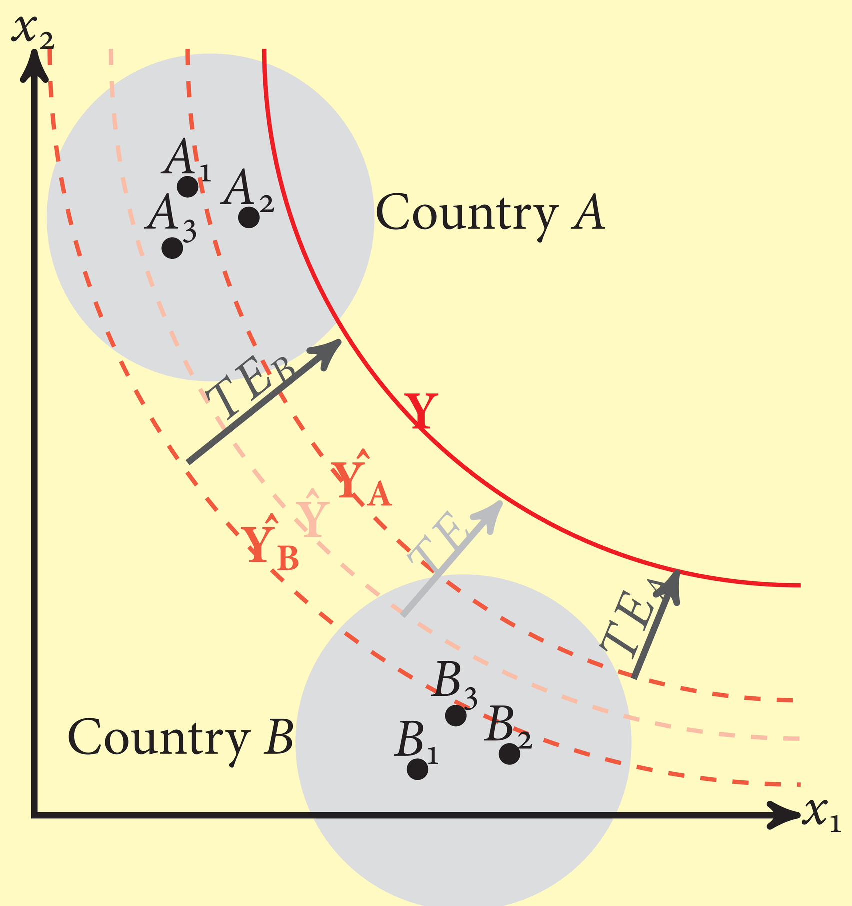
Stochastic frontier models with spatial dependence

Thomas de Graaff tgraaff@feweb.vu.nl VU University Amsterdam & Netherlands Environmental Assessment Agency

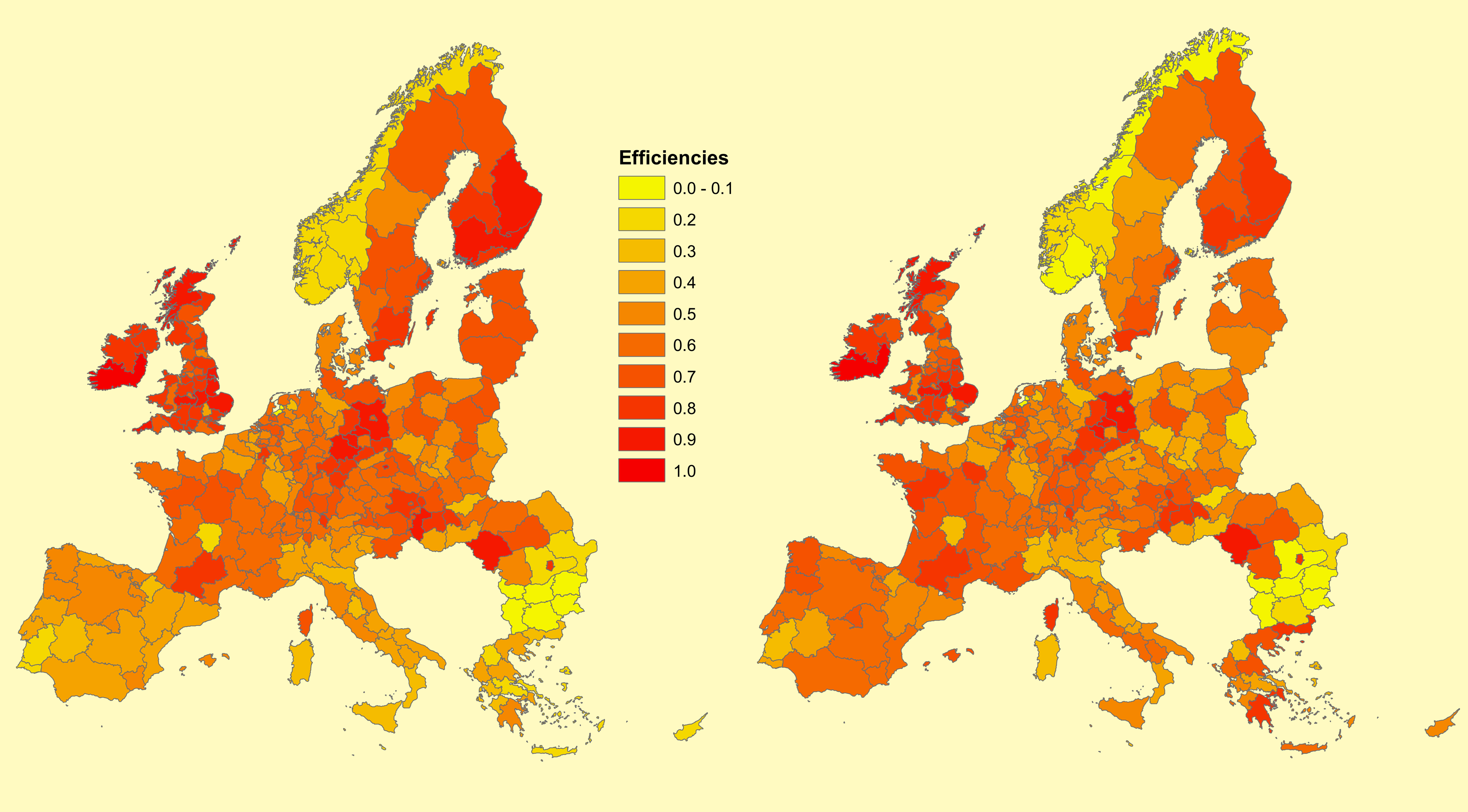
The problem and research aim

Estimation of technical efficiencies may be biased in the presence of spatial dependence or unobserved spatial heterogeneity amongst regions. The aim is therefore to simultaneously model and consistently estimate a model that incorporates both technical inefficiencies and spatial dependence.





Technical inefficiencies in Europe's manufacturing: an application



Left: Standard technical efficiencies *Right:* with additional spatial lag

Stochastic production frontiers

Assume that regional production, y, can be modeled as:

$$y = f(\mathbf{X}; \beta^T) TE,$$

where $\mathbf{X}$ are regional production factors, β the parameters of the production function and TE is the regional specific technical efficiency. By assuming a Cobb-Douglas and that $TE = \exp(-u)$, we get:

$$\ln y = \ln(\mathbf{X})\beta - u + v,$$

Using skew-normal distributions

Let u and v be distributed as:

$$\begin{pmatrix} u \\ v \end{pmatrix} \sim N(0, \Omega^*), \qquad \Omega^* = \begin{pmatrix} 1 & -\delta^T \\ -\delta & \Omega \end{pmatrix}$$

We are interested in $\xi = \ln y - \ln \mathbf{X} = \Pr(v|u < 0)$ (via conditioning): leading to $\ln y \sim \mathrm{SN}(\ln(\mathbf{X})\beta, \Omega, \alpha)$; a multivariate skew-normal distribution with:

$$f(\ln y) = 2\phi(\ln y - \ln(\mathbf{X})\beta; \Omega)\Phi\left(\frac{\alpha}{\omega}(\ln y - \ln(\mathbf{X})\beta\right)$$

- ω is a scale parameter
- α is a measure of skewness
- $\alpha = \left(1 - \delta^T \Omega^{-1}\delta\right)^{-1/2} \Omega^{-1}\delta$

Introducing a spatial lag

Because multivariate skew-normal distributions are closed under affine transformations (similarly to normal distributions), we may write:

$$\mathbf{B}\ln(y) = \ln(\mathbf{X})\beta + \xi \qquad (1)$$

where $\mathbf{B} = (\mathbf{I} - \rho\mathbf{W})$ and with ξ again a multivariate skew normal distribution with $\Omega = \omega(\mathbf{B}'\mathbf{B})^{-1}$). This leads to the following loglikelihood:

$$-\frac{n}{2}\ln(\pi\omega^2) + \ln|\mathbf{B}| - \frac{e'e}{2\omega^2} + \sum \ln 2\Phi\left(\frac{\alpha}{\omega}e\right)$$

where e is the vector of residuals of model (1).

Finding technical inefficiencies

We need to find $TE = \exp(u)$ or $\mathbb{E}(u|\xi)$ given that $u < 0$. Because we can write as well $\xi = \delta|u| + \sqrt{(1-\delta^2)}v$ with $\delta < 0$ (via convolution), where $u \sim N(0,1)$ and $v \sim N(0, \Omega)$, the following general expression holds:

$$u|\xi \sim N^c\left((\mathbf{D}'\Sigma^{-1}\mathbf{D} + \mathbf{I})^{-1}\mathbf{D}'\Sigma^{-1}e, (\mathbf{D}'\Sigma^{-1}\mathbf{D} + \mathbf{I})^{-1}\right)$$

where N^c indicates a normal distribution truncated at 0, $\mathbf{D}$ is a diagonal matrix with δ's on the diagonal and Σ equals $\sqrt{(\mathbf{I} - \mathbf{D}^2)}\Omega$. The expectation can now be readily derived.

Empirical specification

We estimate for the period 1991–2008 a neoclassical growth model of the manufacturing sector across 273 European NUTS-2 regions with the following specification:

$$\ln\frac{y(t)}{y(0)} = \beta_0 + \beta_1 \ln y(0) + \beta_2 \ln s + \beta_3 \ln(n + 0.05) + \xi$$

where s is the savings rate and $y(t)$ the GVA in manufacturing measured at time t, n the manufacturing working population growth rate, ξ is skew-normally distributed and the convergence rate across regions is calculated as: $\hat{\lambda} = -\ln(1 + \hat{\beta}_1)$.

Estimation results

Variable	Growth model	Frontier model	Spatial frontier
Constant	$0.30^{\dagger}$	0.68^{**}	0.07
$\ln(y_0)$	-0.50^{**}	-0.47^{**}	-0.41^{**}
$\ln s$	0.55^{**}	0.51^{**}	0.45^{**}
$\ln(n+g+\delta)$	-0.19^{**}	-0.18^{**}	-0.18^{**}
ω	0.30^{**}	0.44^{**}	0.36^{**}
α		-2.91^{**}	-2.57^{**}
ρ			0.91^{**}
Logl.	-0.21	0.21	0.36
$\hat{\lambda}$	0.053^{**}	0.049^{**}	0.041^{**}

Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%

Conclusions

1. Spatial dependence and stochastic frontiers can be simulteneously and consistently estimated using multivariate skew-normal distribution functions
2. In the presence of spatial dependence, regional technical inefficiency differences can be significantly mitigated

Key references

[1] Abreu, M. *Spatial Determinants of Economic Growth and Technology Diffusion.* Thela Thesis Publishers, Amsterdam, 2005.

[2] Aigner, D. J., Lovell, C. A. K., and Schmidt, P. Formulation and Estimation of Stochastic Production Frontier Models. *Journal of Econometrics 6* (1977), 21–37.

[3] Azzalini, A., and Capitanio, A. Statistical Application of the Multivariate Skew-Normal distribution. *Journal of Royal Statistical Society 61* (1999), 579–602.

[4] Dominguez-Molina, J. A., Gonzalez-Farias, G., and Ramos-Quiroga, R. Skew-normality in stochastic frontier analysis. In *Skew-Elliptical distributions and their applications*, M. G. Genton, Ed. Chapman & Hall/CRC, 2004, ch. 13, pp. 223–242.