On spatial dependence in stochastic frontier models with an application to European regional economic performance

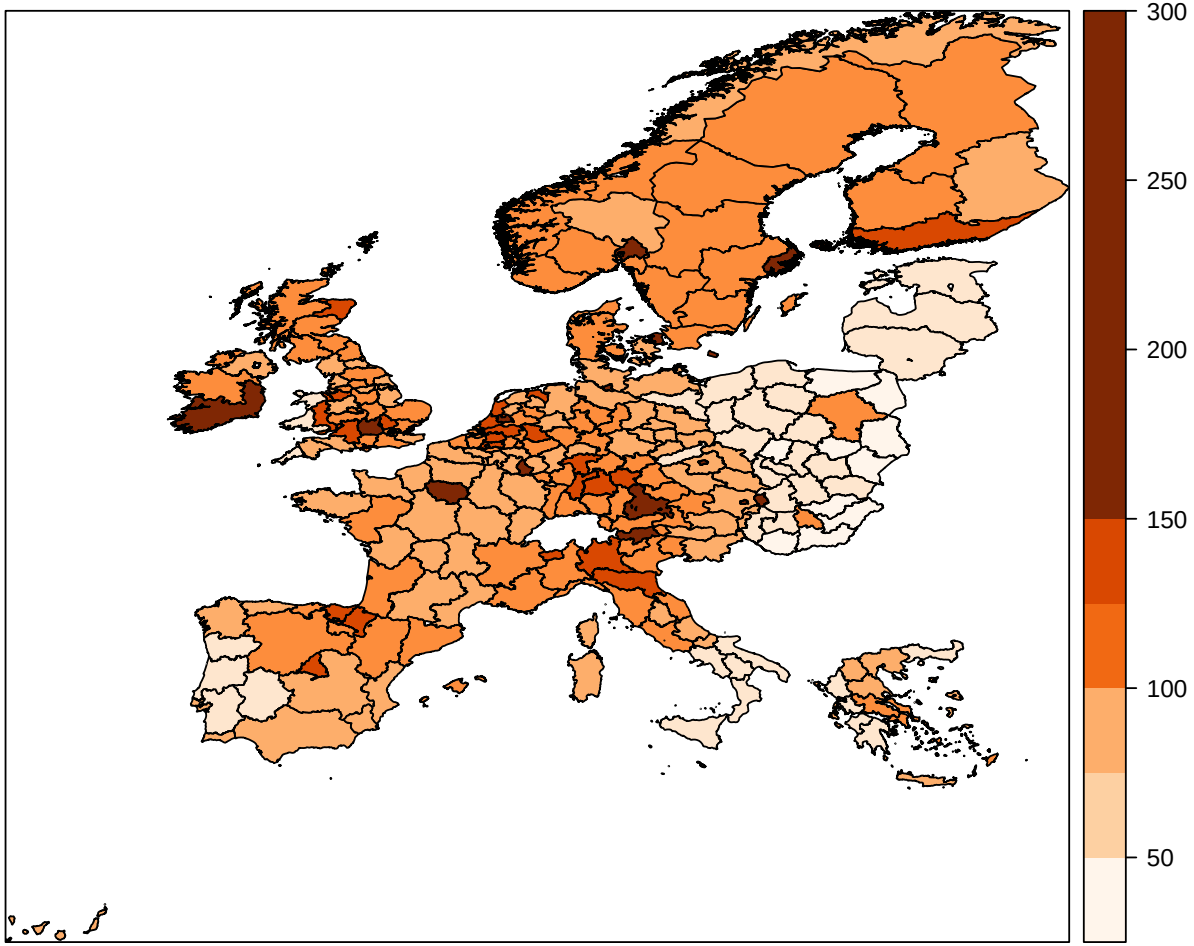
Thomas de Graaff

Department of Spatial Economics, Vrije Universiteit Amsterdam
https://github.com/thdegraaff

GfR-Conference, Marburg, 26 juni 2014

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

GDP per capita across Europe in 2007—EU27 = 100 (source: cambridge econometrics database



Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Background & research question

How to benchmark the economic performance of European regions in the presence of networks

- Data Envelopment Analysis vs Stochastic Frontier Analysis
- Absolute efficiency: the efficiency of the use of input factors vs relative efficiency: accessibility to knowledge/technology
- Relevant for policy
  - Place-based policies versus place-neutral policies (Barca, 2011)
  - Determinants of economic performance might vary over space (Thissen et al., 2013)

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Related literature

- Starting point: Weinstein (Technometrics, 1964)
  - Econometric literature: Aigner et al. (1977); Meeusen & van den Broeck (1977).
  - Statistical literature: Azzalini (1985); Azzalini & Capitanio (1999); Dominguez-Molina et al. (2003)

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Related literature

- Starting point: Weinstein (Technometrics, 1964)
  - Econometric literature: Aigner et al. (1977); Meeusen & van den Broeck (1977).
  - Statistical literature: Azzalini (1985); Azzalini & Capitanio (1999); Dominguez-Molina et al. (2003)
- Regional stochastic frontiers: Puig-Junoy (2001); Alvarez (2007)

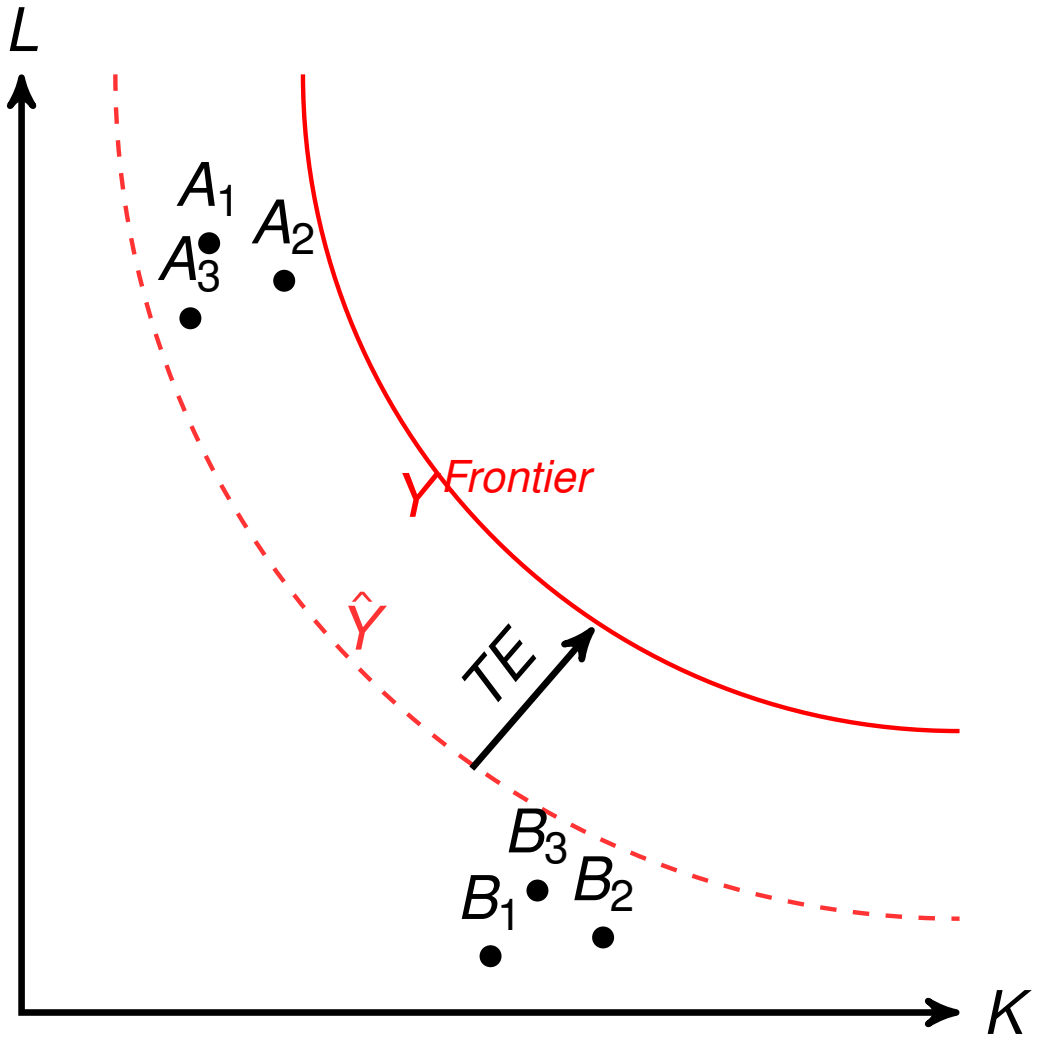
Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Related literature

- Starting point: Weinstein (Technometrics, 1964)
  - Econometric literature: Aigner et al. (1977); Meeusen & van den Broeck (1977).
  - Statistical literature: Azzalini (1985); Azzalini & Capitanio (1999); Dominguez-Molina et al. (2003)
- Regional stochastic frontiers: Puig-Junoy (2001); Alvarez (2007)
- Recent literature concerning spatial stochastic frontiers:
  - Parametric approach: Pavlyuk (2010); Barrios & Lavado (2010)
  - Bayesian approach: Schmidt et al. (2009); Areal et al. (2010)

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Production Isoquant



Stochastic frontiers

In general, regional production, Y, can be modeled as:

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

Assume Cobb-Douglas and $TE = \exp(-U)$, then:

$$\ln Y = \ln A + \beta_1 \ln K + \beta_2 \ln L - U + V,$$

Where usually:

- $U \sim N^+\left(0, \sigma_U^2\right)$
- $V \sim N\left(0, \sigma_V^2\right)$

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

The skew-normal distribution

- Alternative specification, skew normal distribution:

$$Z \sim SN(\alpha) = 2\phi(z)\Phi(\alpha z)$$

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

The skew-normal distribution

- Alternative specification, skew normal distribution:

$$Z \sim SN(\alpha) = 2\phi(z)\Phi(\alpha z)$$

- Which has two types of genesis:
  1. By convolution: $Z = \delta|U| + \sqrt{1-\delta^2}V$
  2. By conditioning: $Z = (V|U < 0)$

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

The skew-normal distribution

- Alternative specification, skew normal distribution:

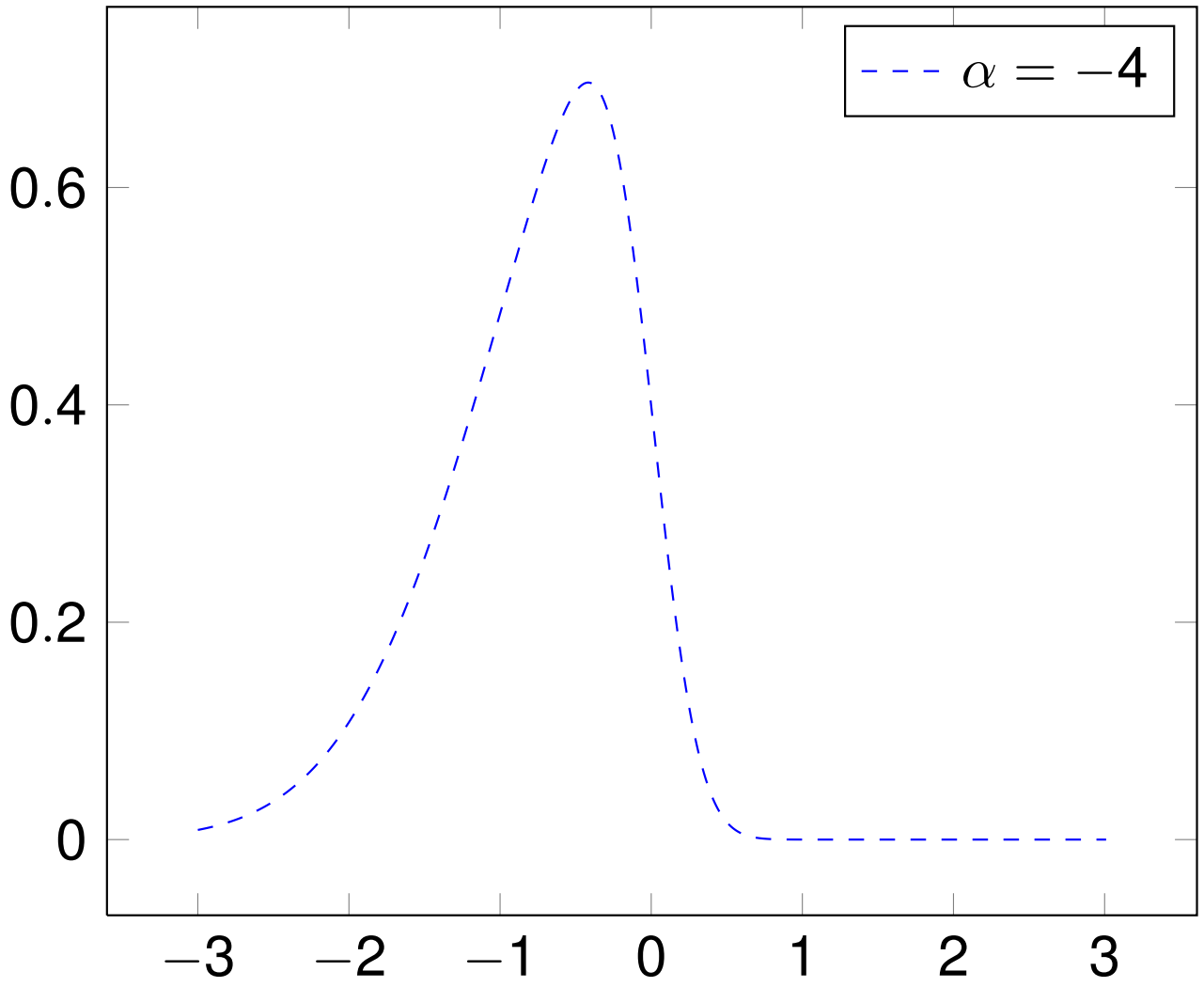
$$Z \sim SN(\alpha) = 2\phi(z)\Phi(\alpha z)$$

- Which has two types of genesis:
  1. By convolution: $Z = \delta|U| + \sqrt{1-\delta^2}V$
  2. By conditioning: $Z = (V|U < 0)$

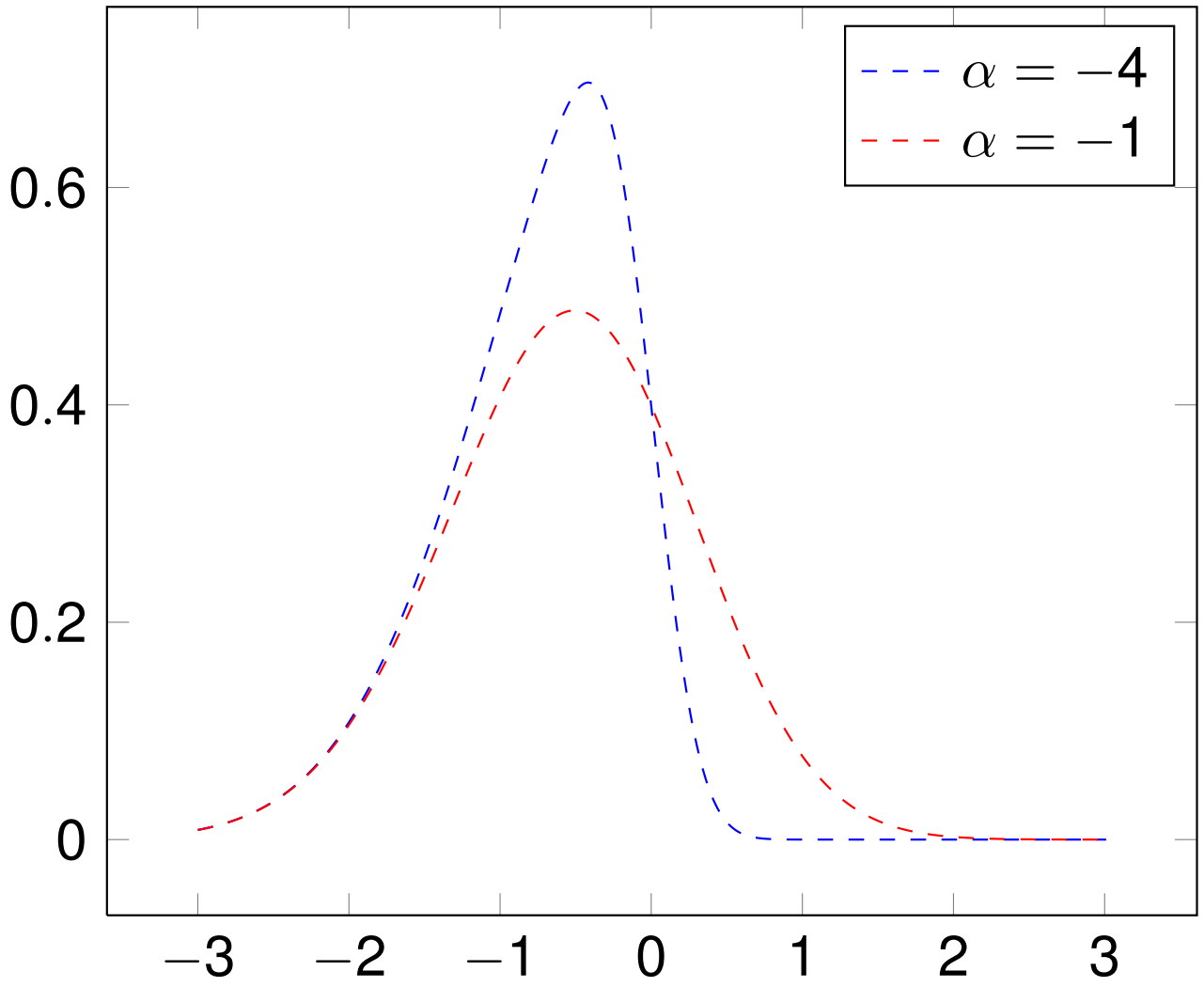
- And is closed under affine transformations:
  - if $Y = \mathbf{X}\beta + \sigma Z$ then $Y \sim SN(\mathbf{X}\beta, \sigma, \alpha)$

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Varying the shape parameter α

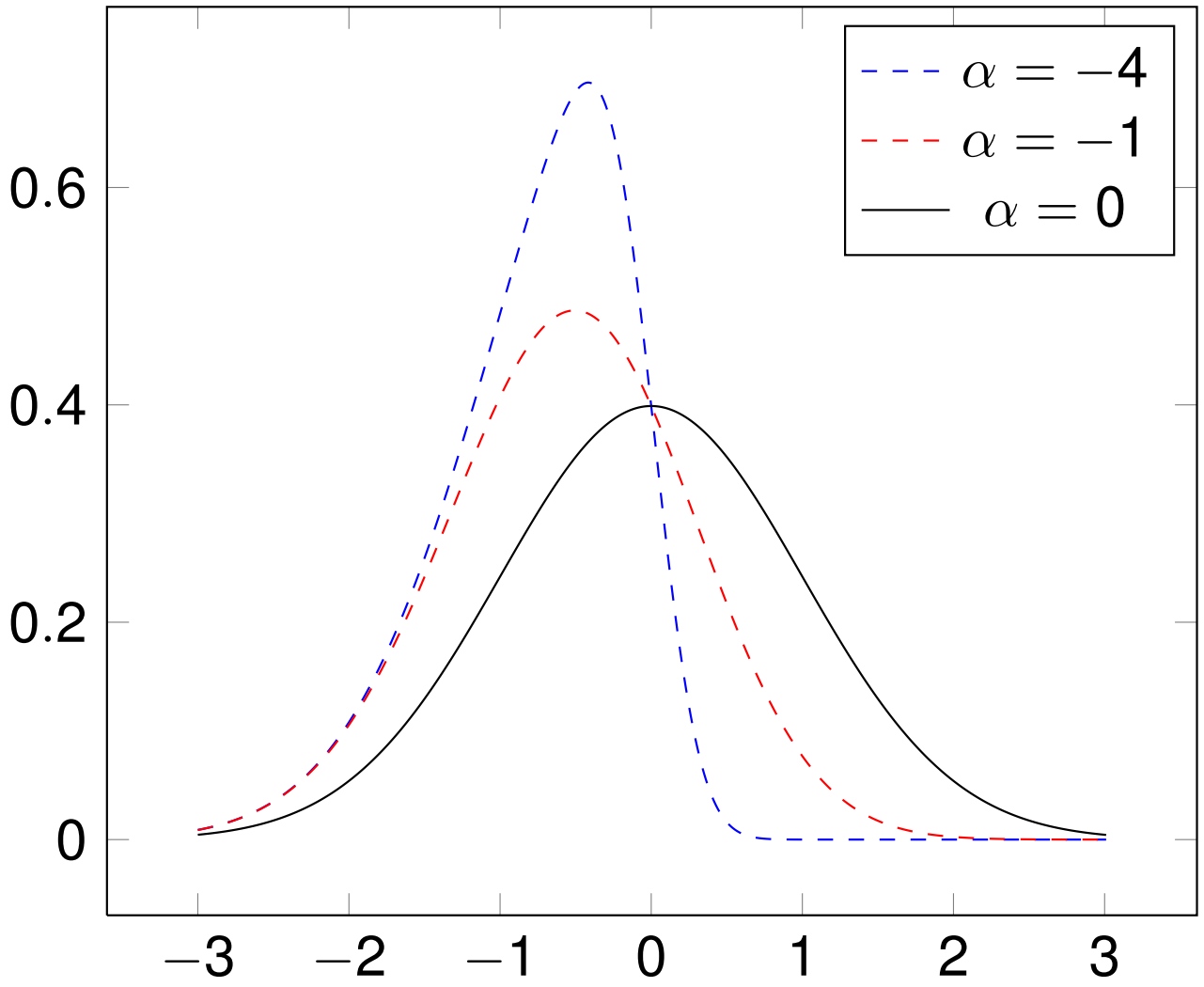


Varying the shape parameter α



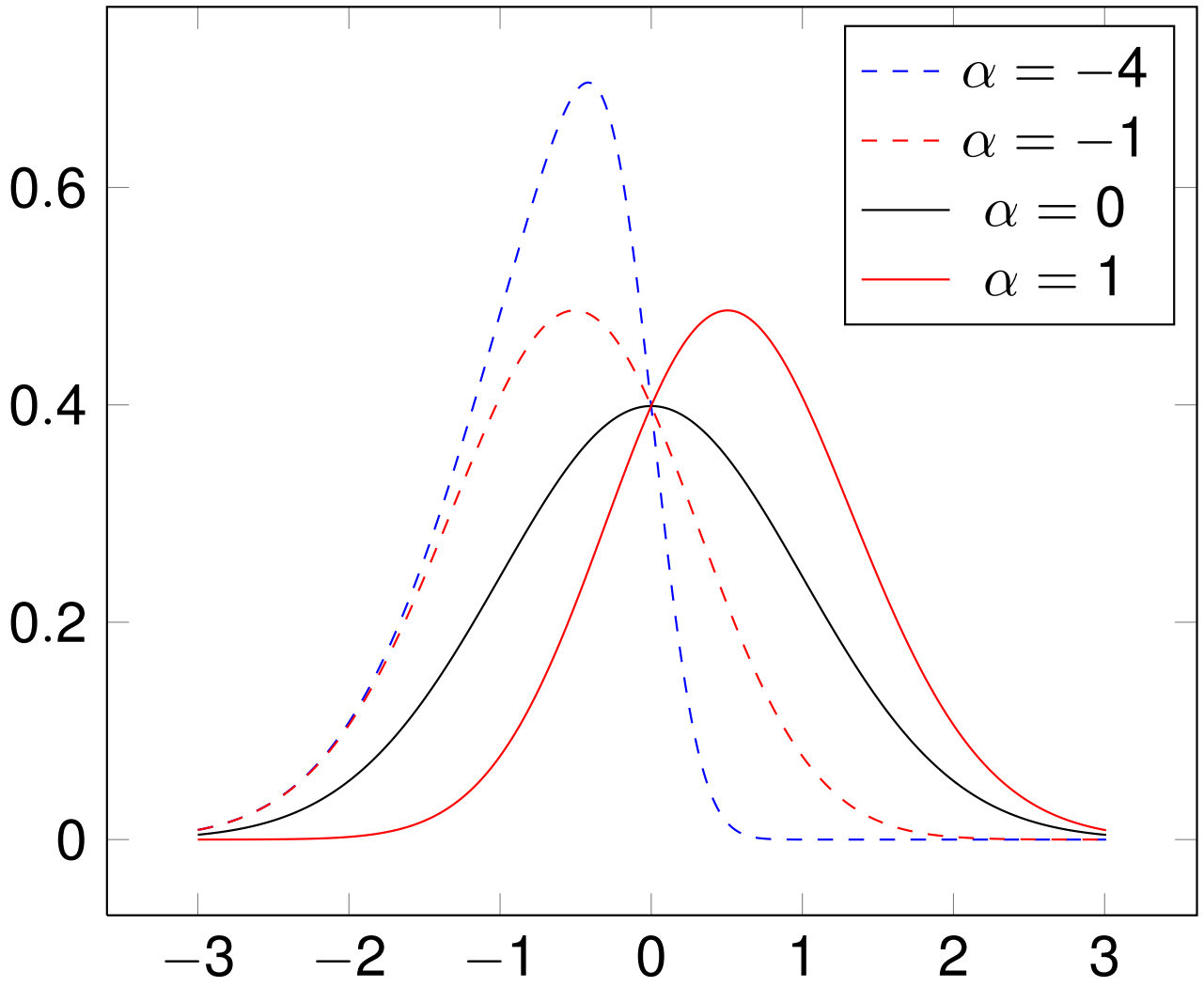
Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Varying the shape parameter α



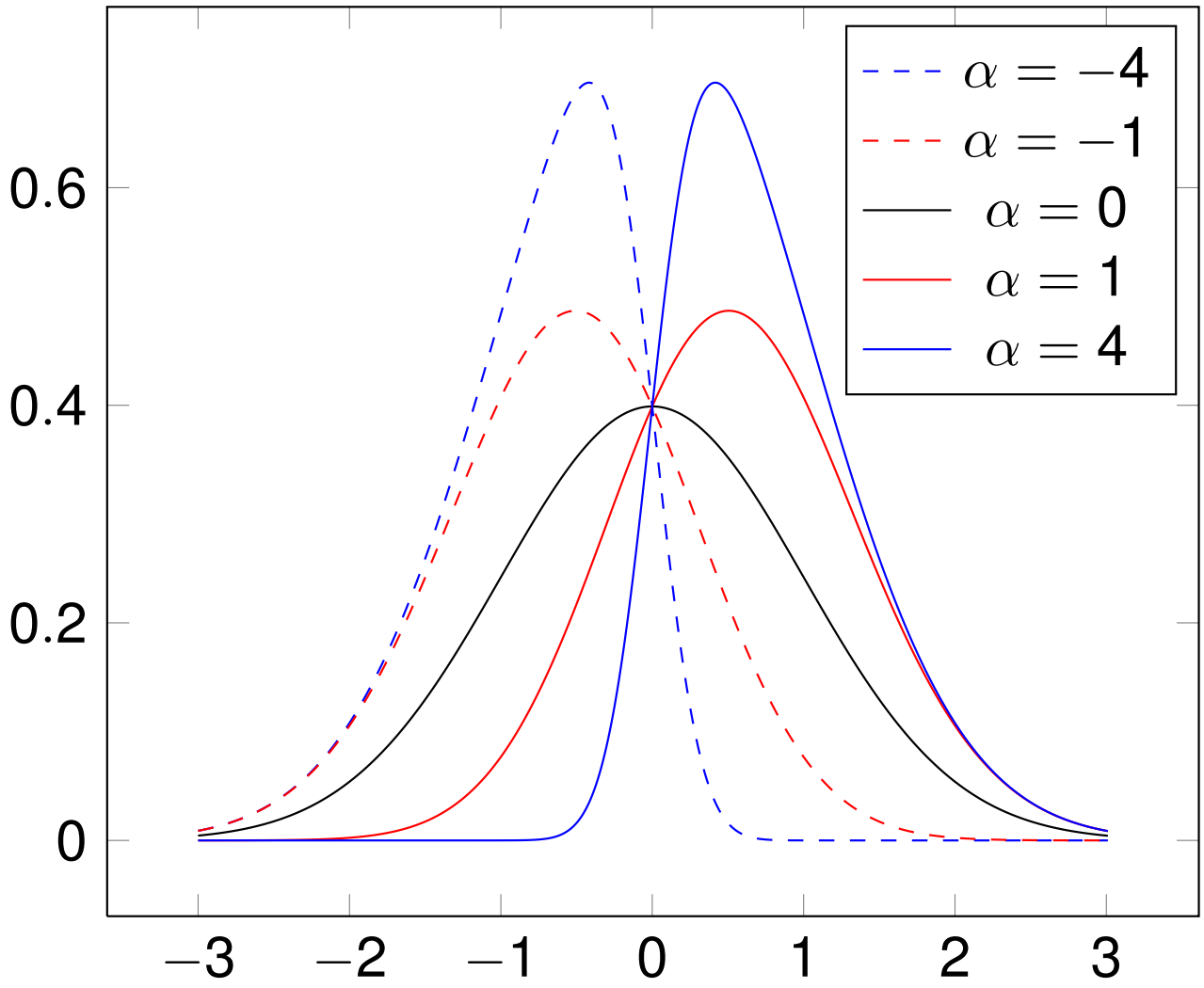
Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Varying the shape parameter α



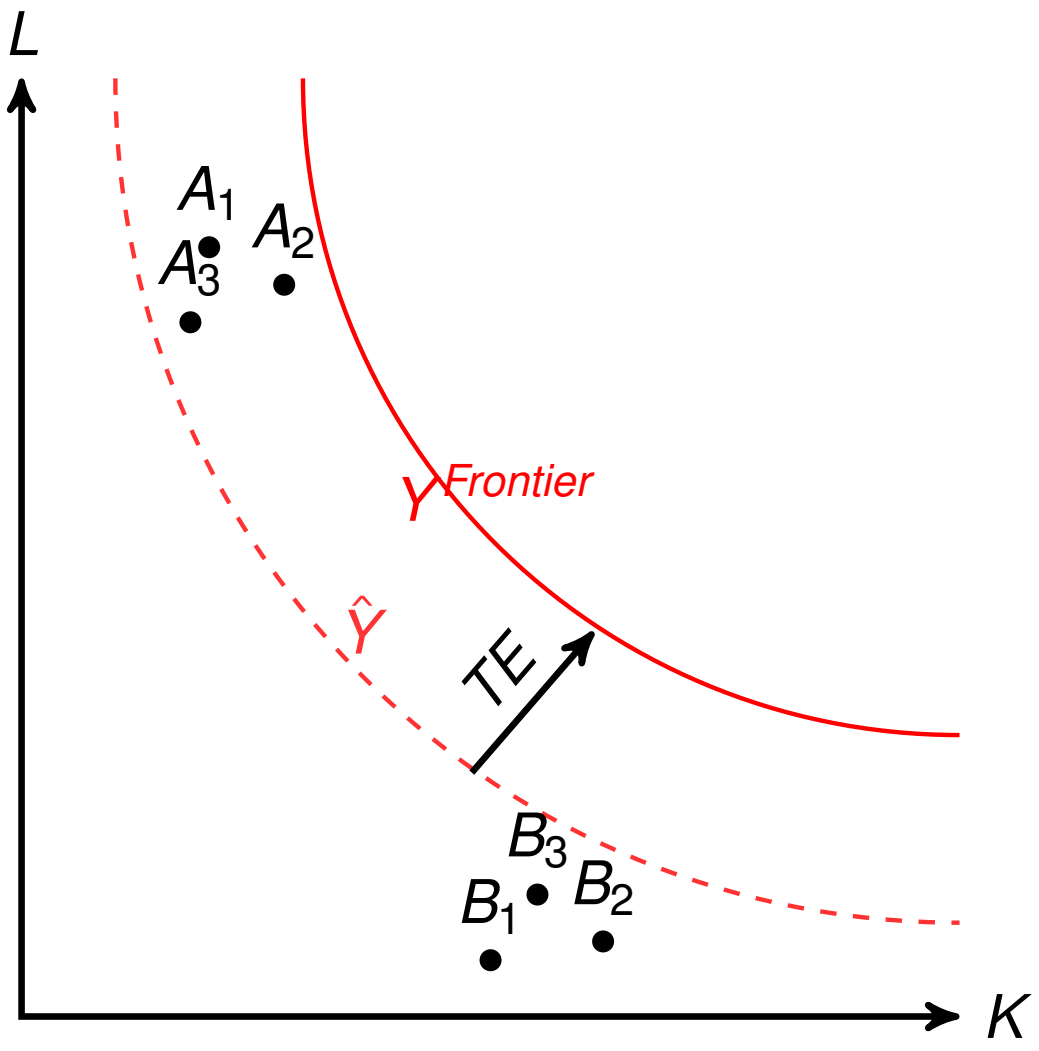
Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Varying the shape parameter α

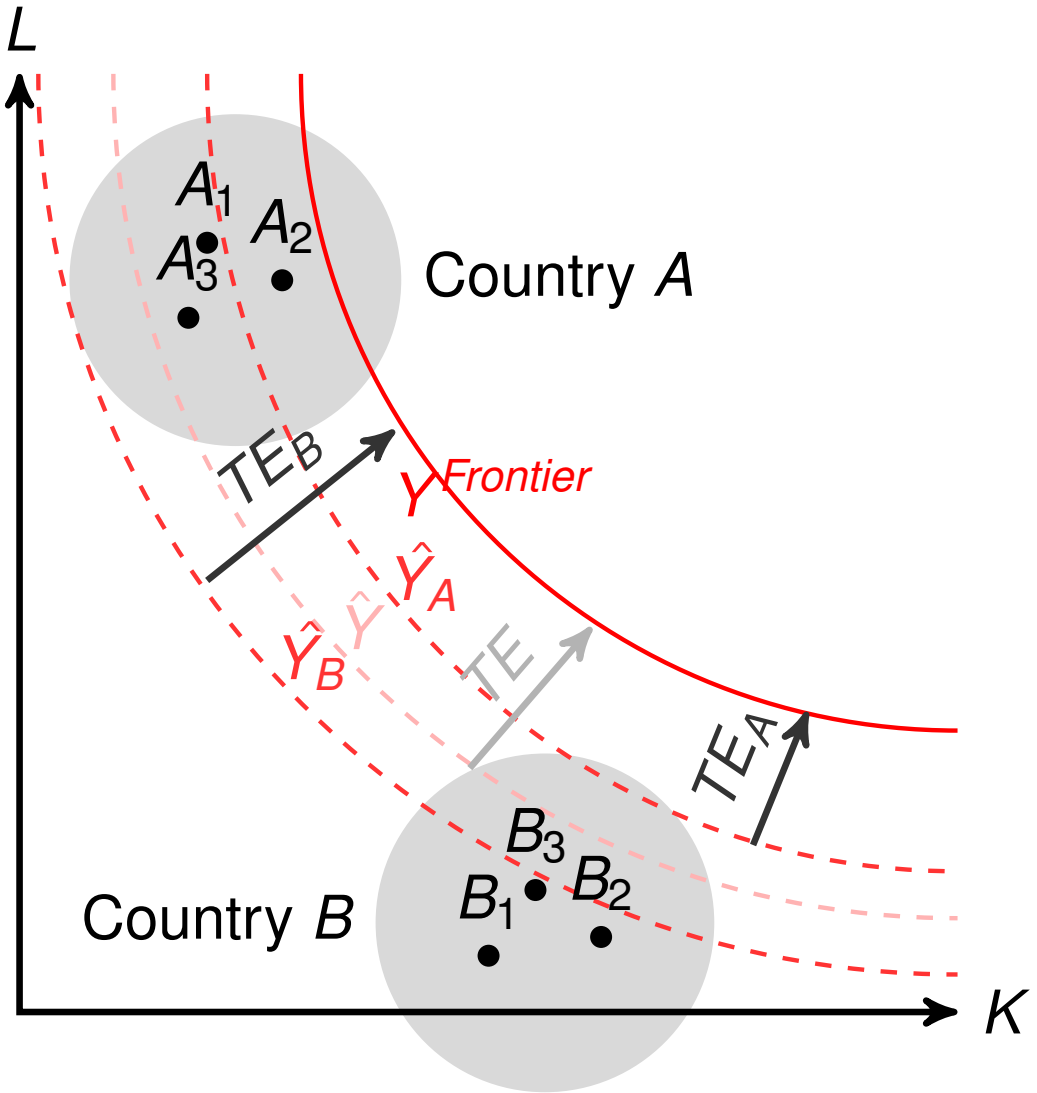


Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Production isoquant revisited (I)



Production isoquant revisited (II)



Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Introducing a spatial lag (SAR)

Because multivariate *skew-normal* distributions are closed under affine transformations, we may write:

$$\mathbf{B}\ln(Y) = \ln A + \beta_1 \ln K + \beta_2 \ln L + Z$$

where $\mathbf{B} = (\mathbf{I} - \rho\mathbf{W})$ and with $Z \sim SN(\mathbf{B}^{-1}\mathbf{X}\beta, \Omega, \alpha)$ a multivariate skew normal distribution with $\Omega = \sigma^2(\mathbf{B}'\mathbf{B})^{-1}$.

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Introducing a spatial lag (SAR)

Because multivariate *skew-normal* distributions are closed under affine transformations, we may write:

$$\mathbf{B}\ln(Y) = \ln A + \beta_1 \ln K + \beta_2 \ln L + Z$$

where $\mathbf{B} = (\mathbf{I} - \rho\mathbf{W})$ and with $Z \sim SN(\mathbf{B}^{-1}\mathbf{X}\beta, \Omega, \alpha)$ a multivariate skew normal distribution with $\Omega = \sigma^2(\mathbf{B}'\mathbf{B})^{-1}$.

Traditional approach

Other genesis:
$\ln Y = \mathbf{B}^{-1}(\ln A + \beta_1 \ln K + \beta_2 \ln L + V) - U$, were:

$$\begin{pmatrix} U \\ V \end{pmatrix} \sim N_{1+k}(0, \Omega^*), \qquad \Omega^* = \begin{pmatrix} 1 & -\delta' \\ -\delta & \Omega \end{pmatrix}$$

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Introducing a spatial error (SEM)

Similar to spatial lag model

$$\ln(Y) = \ln A + \beta_1 \ln K + \beta_2 \ln L + \mathbf{A}^{-1} Z$$

where $\mathbf{A} = (\mathbf{I} - \lambda \mathbf{W})$ and with $Z \sim SN(\mathbf{X}\beta, \Omega, \alpha)$ a multivariate skew normal distribution with $\Omega = \sigma^2(\mathbf{A}'\mathbf{A})^{-1}$.

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Regional production function data

- Cambridge Econometrics Dataset for Employment
- Regional value added and capital stock from regional supply and use framework (Thissen et al. 20013a, 2013b and 2013c)
  - Regional trade database between 256 European Nuts2 regions over 15 sectors (2000–2010)
  - No Perpetual Inventory Method but value added for capital $\rightarrow$ Using country specific interest rates we can then find capital stock ($VA^{capital} = rK$)
- Sample period: mean over 2000–2010
- Spatial weights matrix $\mathbf{W}$: 4 nearest neighbours

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Results for energy and manufacturing sector

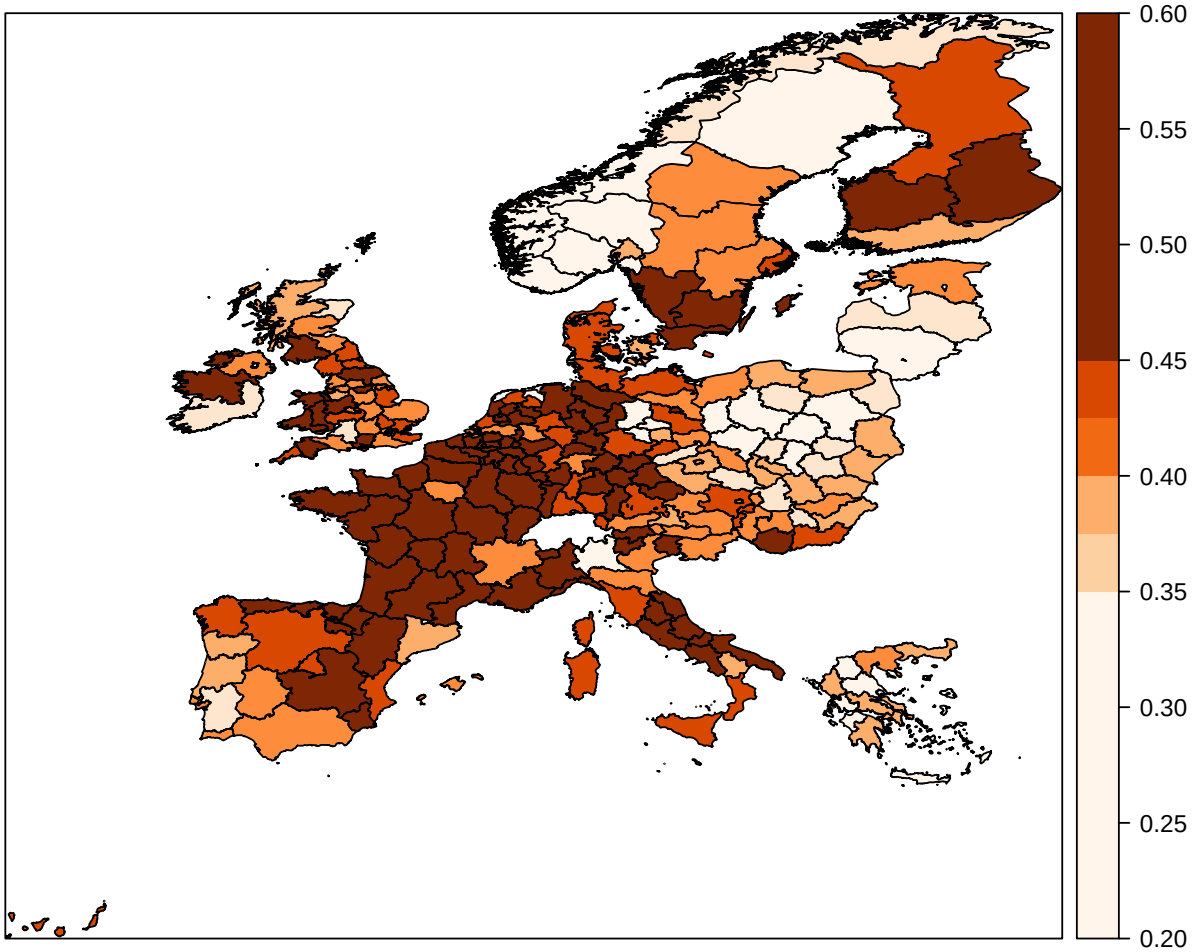
	OLS	SEM	SAR	Frontier	SEM fr.	SAR fr.
Constant	2.40***	3.16***	0.46	2.50***	4.03***	0.49
	(0.16)	(0.14)	(0.26)	(0.18)	(0.25)	(1.22)
ln(Capital)	0.77***	0.58***	0.70***	0.76***	0.58***	0.70***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
ln(Employment)	0.30***	0.44***	0.32***	0.30***	0.44***	0.32***
	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
σ	0.33***	0.22***	0.29***	0.33***	0.29***	0.29*
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.12)
λ		0.79***			0.78***	
		(0.09)			(0.09)	
ρ			0.25***			0.25***
			(0.03)			(0.03)
α				−1.75***	−1.41***	−0.13
				(0.53)	(0.37)	(5.09)
Log Lik.	−80.10	92.28	41.45	11.03	94.86	41.45

Results all sectors for frontier model with spatial error

	Agric.	E & M	Constr.	Distri.	Serv.	NM. Serv
Constant	2.43***	4.03***	2.11***	1.24***	2.05***	1.74***
	(0.71)	(0.25)	(0.22)	(0.17)	(0.29)	(0.19)
ln(Capital)	0.68***	0.58***	0.61***	0.69***	0.71***	0.61***
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
ln(Employment)	0.21***	0.44***	0.39***	0.30***	0.29***	0.39***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)
σ	0.26***	0.29***	0.28***	0.24***	0.17***	0.25***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)
α	−0.03	−1.41***	−1.19***	−1.54***	−0.02	−1.68***
	(1.44)	(0.37)	(0.29)	(0.34)	(0.75)	(0.32)
λ	0.57***	0.78***	0.76***	0.78***	0.59***	0.81***
	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)
Log Lik.	62.92	94.86	94.19	149.78	164.88	143.52

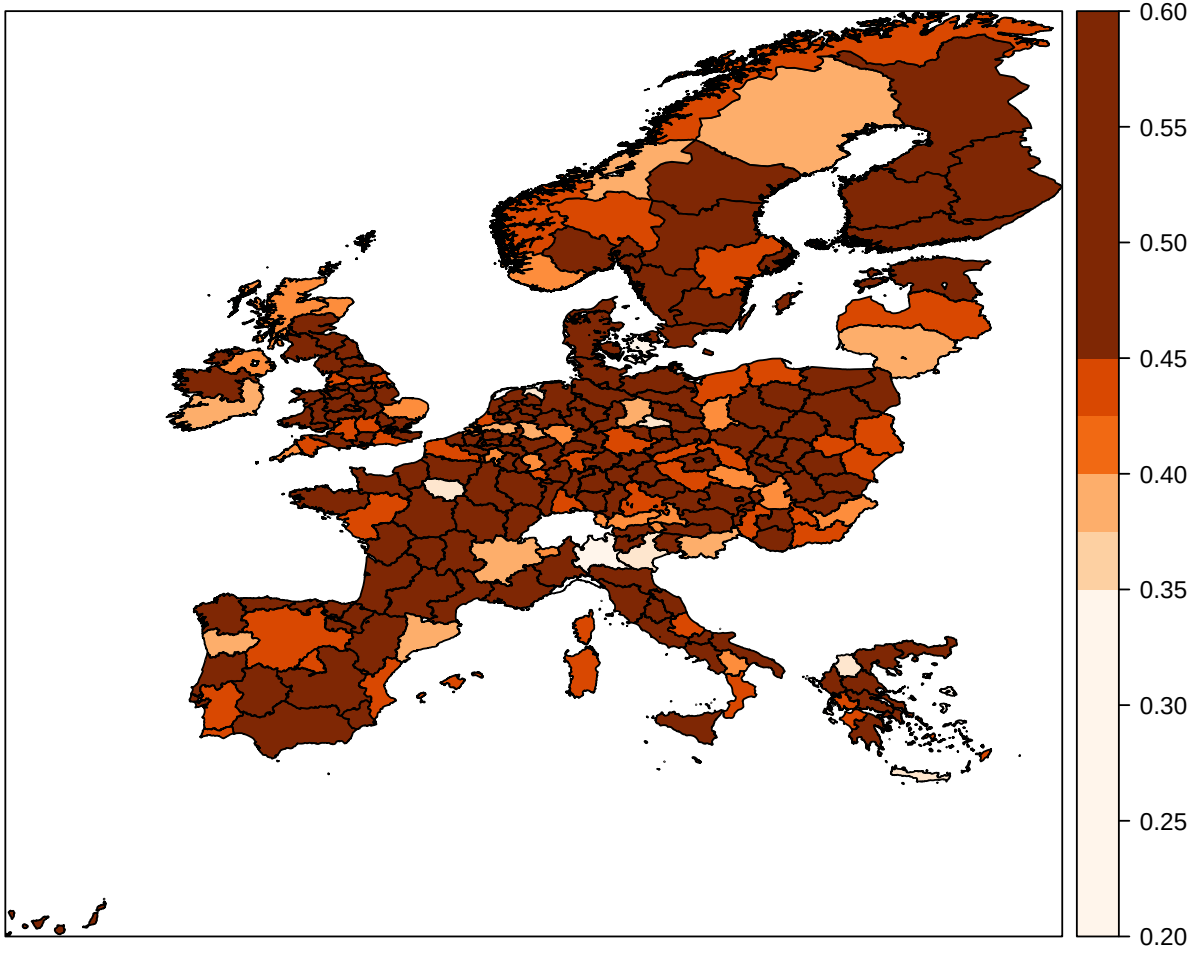
Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Regional efficiency manufacturing



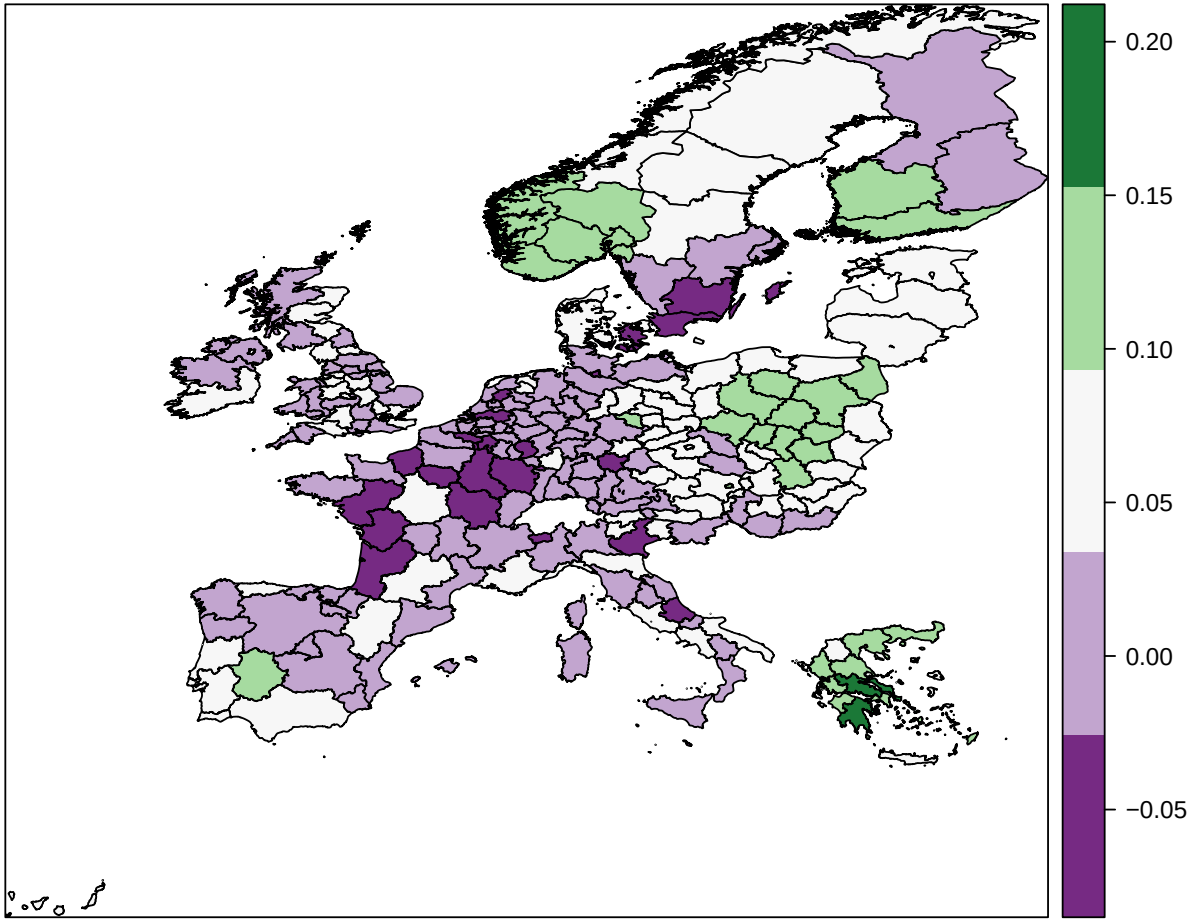
Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Regional efficiency manufacturing controlling for distance



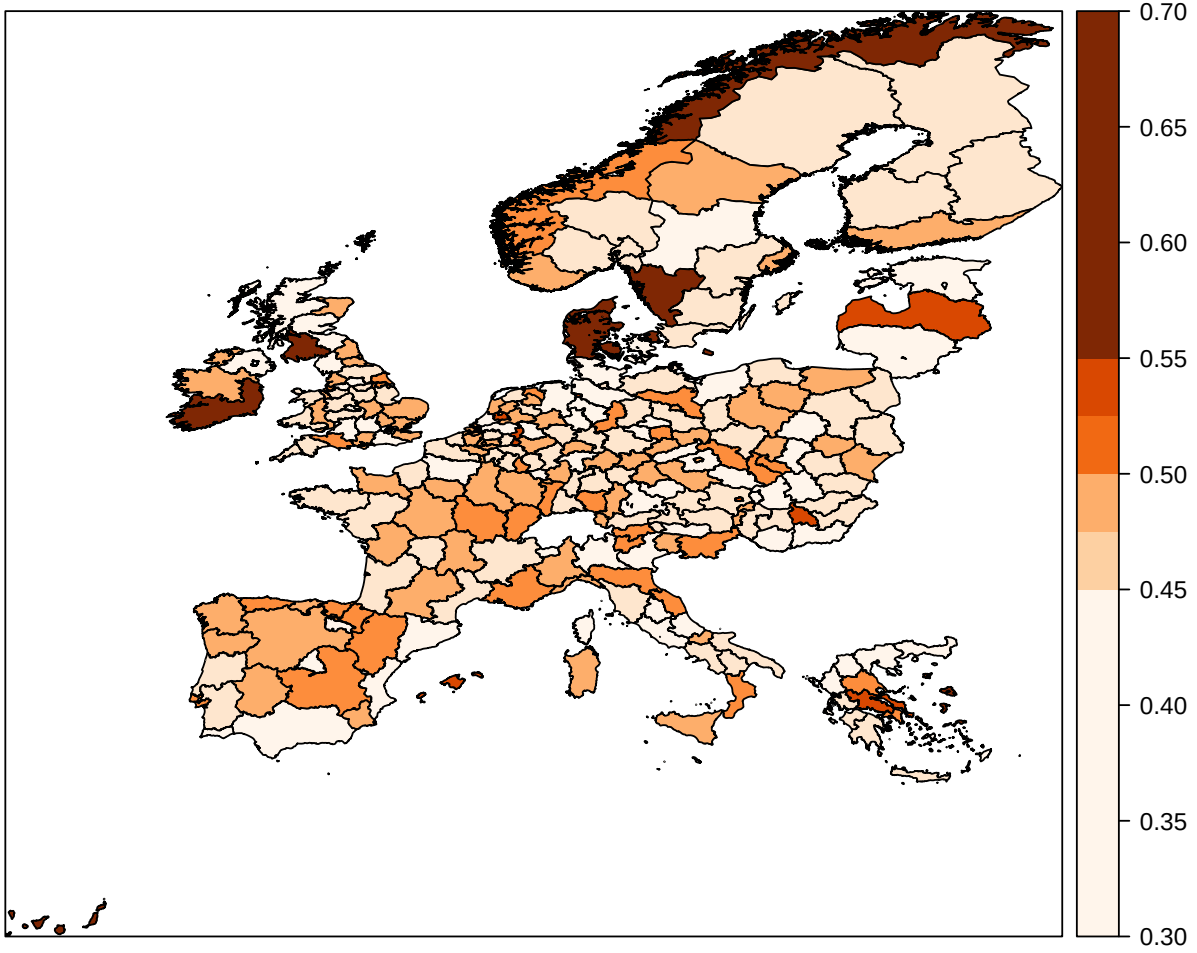
Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Difference frontier vs. spatial error frontier for E & M



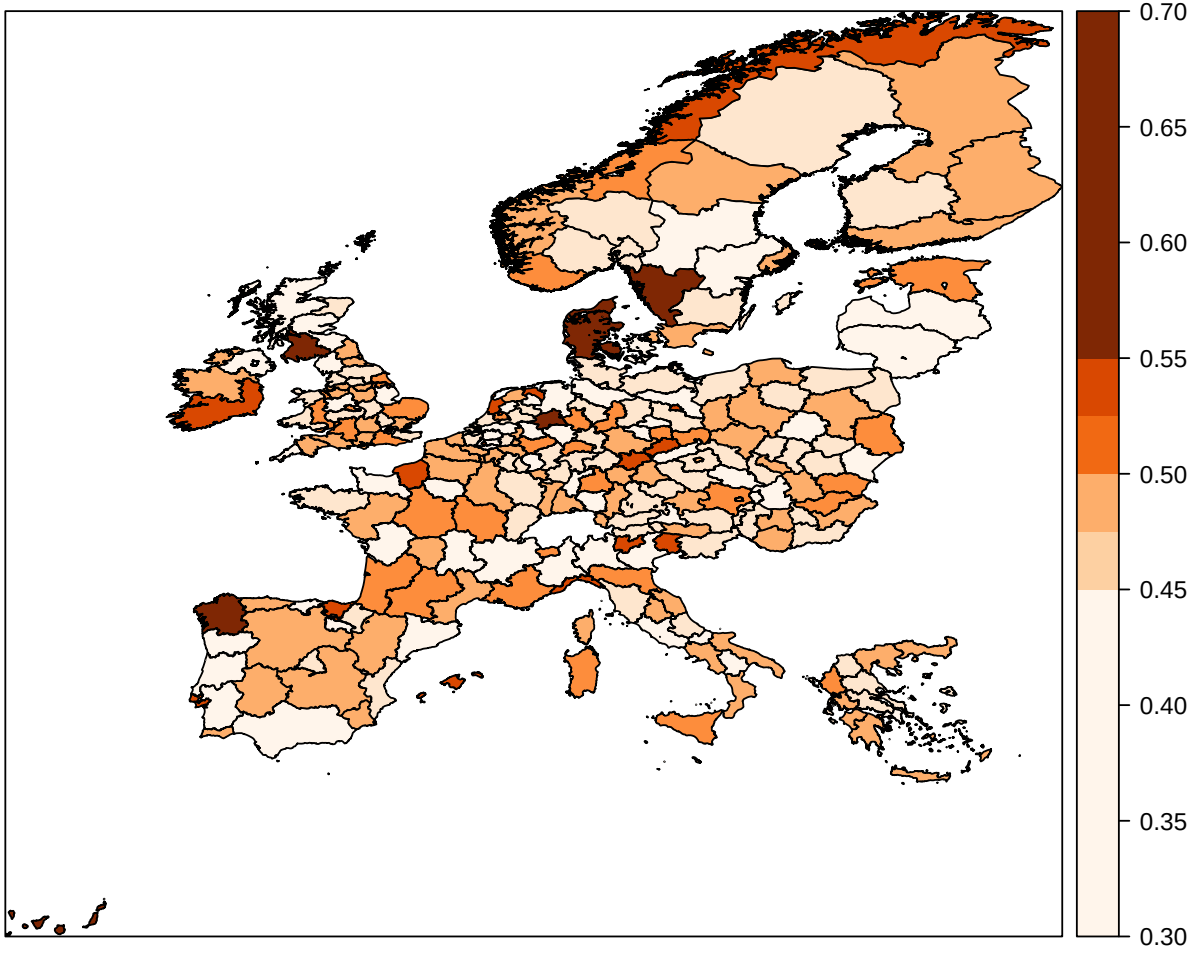
Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Regional efficiency construction controlling for distance



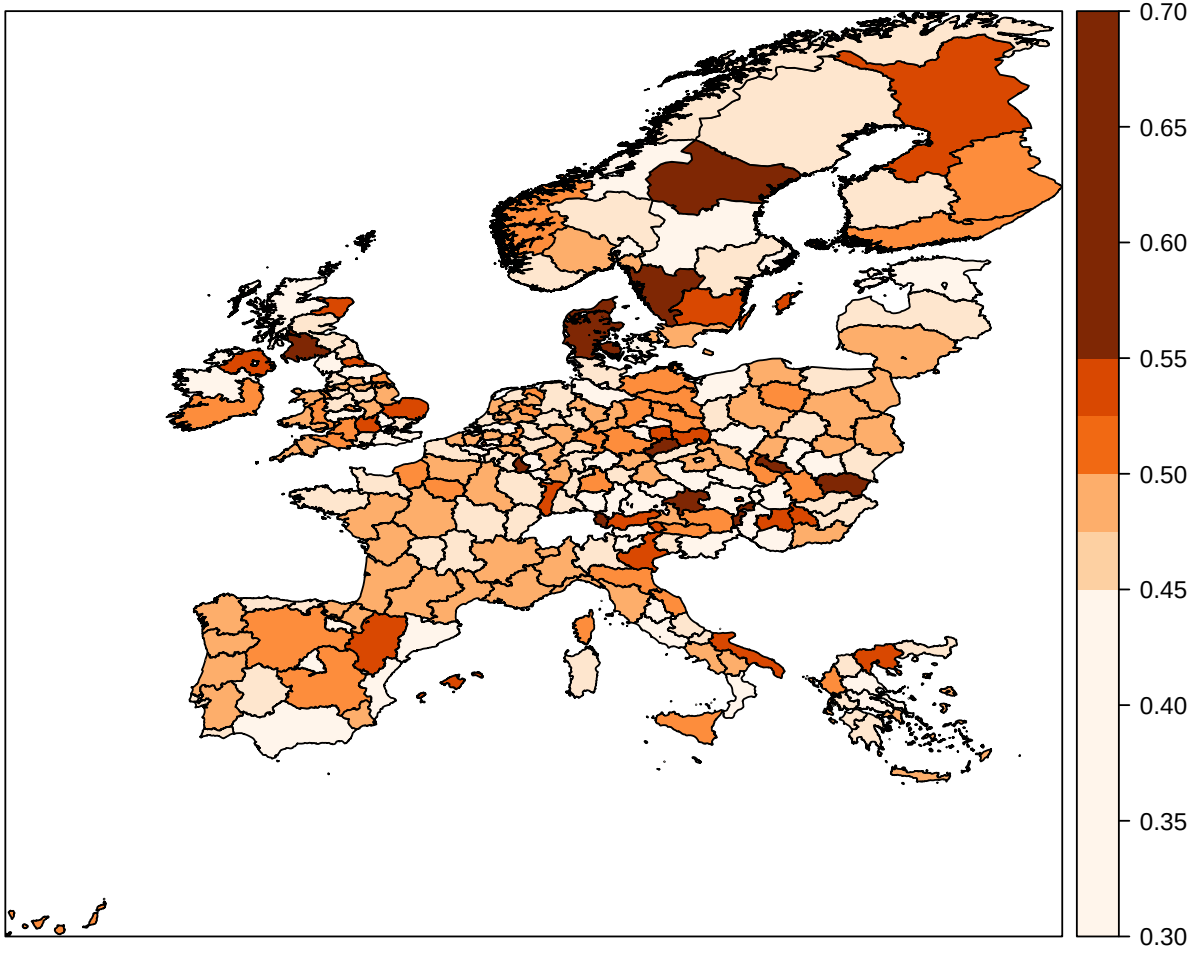
Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Regional efficiency Distribution controlling for distance



Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Regional efficiency non market services controlling for distance



Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Conclusion

1. Spatial dependence and stochastic frontiers can be simultaneously and consistently estimated
   - relevant for regions and policy
   - skew-normal distribution is not necessary but much more elegant, more straightforward and less messy
2. In the presence of spatial dependence, regional technical inefficiency differences are significantly mitigated
3. Spatial dependence seems to dominate efficiency mechanism, but . . .

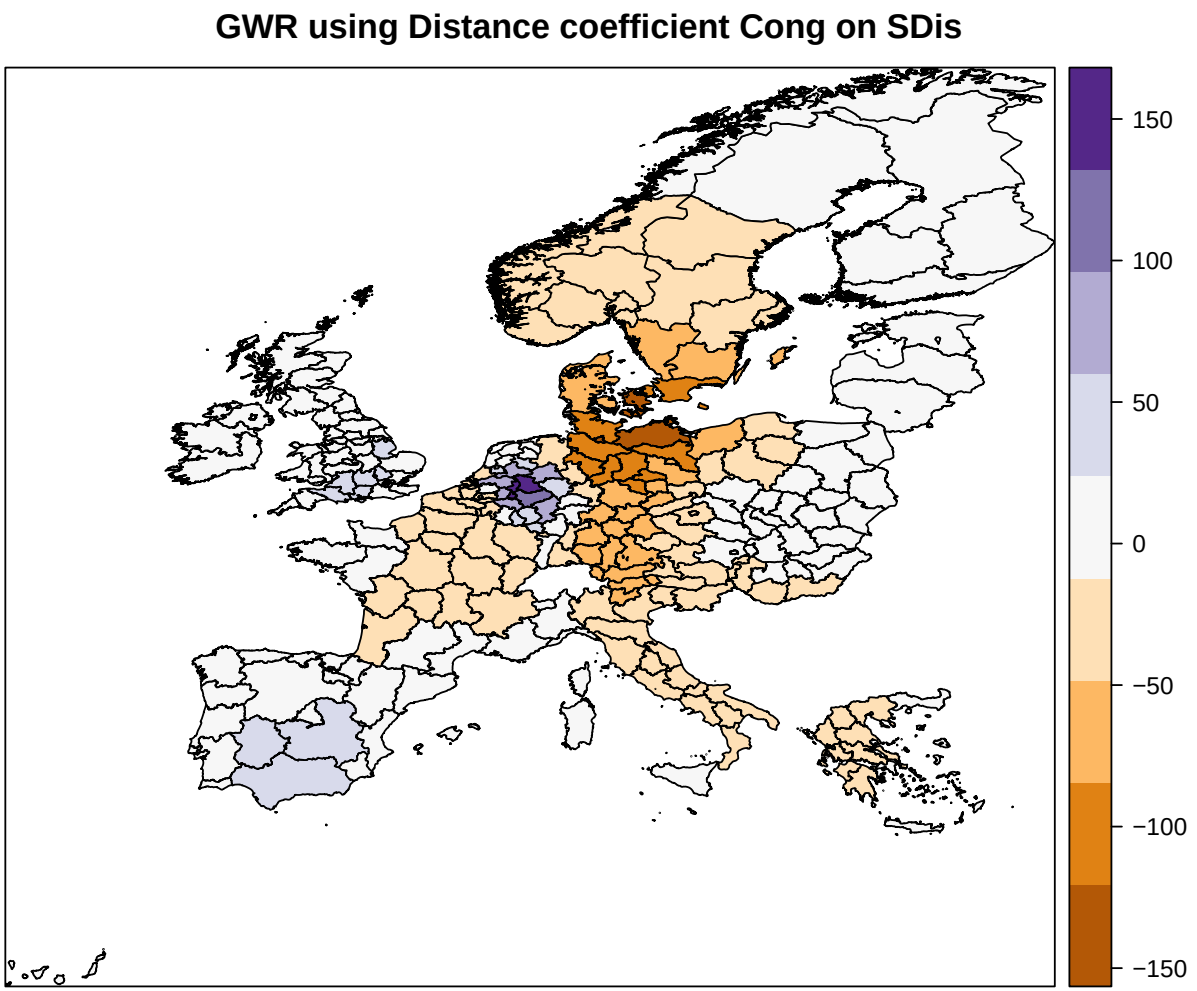
Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Further research

1. Using panel structure to control for regional fixed effects (first differencing)
2. Using different weight matrices (FDI, co-patenting and structural trade)
3. Regional production functions do not always behave as production functions, perhaps due to:
   - omitted variables; country dummies
   - missing human capital
4. Policy makers might be more interested in what determines (congestion, accessibility, private R&D) regional efficiency ...

Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion

Example: spatially varying impact of congestion on efficiency distribution



Why bother?
Background
Literature
Benchmarking regions
Frontiers in production
Stochastic frontiers
Skew-normal alternative
Benchmarking in the presence of networks
Main concept
Spatial frontiers
European regional efficiency
Data
Results
Regional efficiency
In conclusion