A Regional Perspective on Industrial Production in Ireland over 1979-2009:

A Spatial Stochastic Frontier Approach with Unknown W

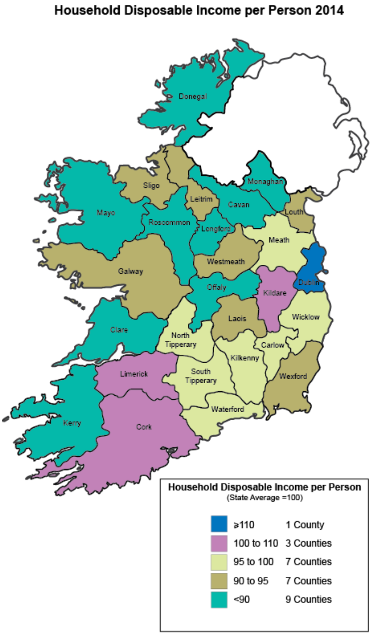
Achim Ahrens *(ESRI, Dublin)* &
Edgar Morgenroth *(Dublin City University)*

58th ERSA Congress

Cork, August 28-31, 2018

Background & Motivation

Ireland is characterised by regional disparities: the Dublin region and city of Cork stand in stark contrast to the rest of the country.

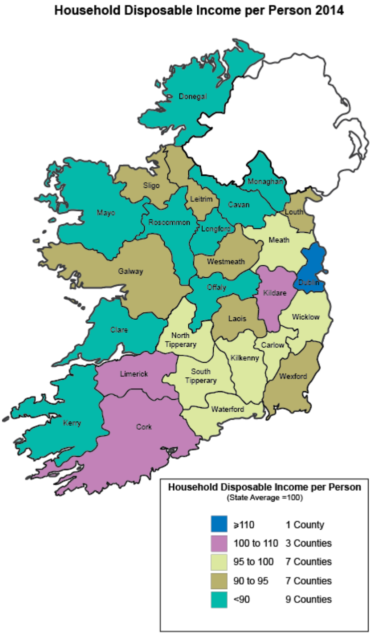


Background & Motivation

Ireland is characterised by regional disparities: the Dublin region and city of Cork stand in stark contrast to the rest of the country.

What drives economic differences across counties?

We try to contribute to the debate by focusing on the **Industrial Sector**, using a county-level data from the *Census of Industrial Production* covering 1979-2009.



Research question

Decompose growth & level of industrial output:

- technological change: *shifts in the production frontier*,
- technological catch-up: *movements toward the frontier*,
- labour & capital accumulation: *movement along the frontier*.

Identify regions which are characterised by **technical inefficiency** in the industrial sector.

Research question

Decompose growth & level of industrial output:

- technological change: *shifts in the production frontier*,
- technological catch-up: *movements toward the frontier*,
- labour & capital accumulation: *movement along the frontier*.

Identify regions which are characterised by **technical inefficiency** in the industrial sector.

Open question: Why do some countries exhibit lower output for given inputs? What are the drivers of technical inefficiency?

Do we see **convergence across counties** in terms of output, labour productivity and technical efficiency?

Methodology

Two econometric challenges:

(1) decompose technical inefficiency from random noise

(2) address spatial or cross-section dependence

Methodology

Two econometric challenges:

(1) decompose technical inefficiency from random noise

(2) address spatial or cross-section dependence

Challenge (1) is commonly addressed by Data Envelopment Analysis (DEA) or Stochastic Frontier Analysis (SFA). Both widely applied in firm-level or macro-economic studies.

Methodology

Two econometric challenges:

(1) decompose technical inefficiency from random noise

(2) address spatial or cross-section dependence

Challenge (1) is commonly addressed by Data Envelopment Analysis (DEA) or Stochastic Frontier Analysis (SFA). Both widely applied in firm-level or macro-economic studies.

Challenge (2) is increasingly acknowledged in panel studies. In regional studies we expect strong interactions between counties, e.g., due to cross-border commuting, migration.

Econometric Model

The economic model:

$$Y_{it} = A_{it} F(L_{it}, K_{it})$$

where A_{it} is **total factor productivity (TFP)**.

Standard interpretation: TFP refers to the (unobservable) portion of output not explained by inputs used in production.

TFP is known to account for the largest proportion of income/output differences (up to 90%; e.g. Easterly & Levine, 2001).

Econometric Model

We assume

$$A_{it} = \bar{A}_t TE_{it}.$$

That is, TFP consists of two components:

- $\bar{A}_t$ represents **technology available** to all counties at time t,
- $TE_{it} \in (0, 1]$ is a **technical efficiency score** with $TE_{it} = e^{-u_{it}}$ where $u_{it} \geq 0$ is a one-sided error term.

This allows us to decompose TFP into

- technical change (changes in $\bar{A}_t$),
- and changes in TE_{it} (movements toward the frontier).

Econometric Model

The econometric model:

$$\ln Y_{it} = \beta_1 \ln L_{it} + \beta_2 \ln K_{it} + \bar{a}_t - u_{it} + v_{it}$$

where $\ln \bar{A}_t = \bar{a}_t$, $u_{it} \geq 0$ is a measure of technical *in*efficiency and v_{it} is the idiosyncratic error term.

Econometric Model

The econometric model:

$$\ln Y_{it} = \beta_1 \ln L_{it} + \beta_2 \ln K_{it} + \bar{a}_t - u_{it} + v_{it}$$

where $\ln \bar{A}_t = \bar{a}_t$, $u_{it} \geq 0$ is a measure of technical *in*efficiency and v_{it} is the idiosyncratic error term.

- $u_{it} = 0$ implies efficiency; i.e., county i is at the technological frontier at time t.
- $u_{it} > 0$ implies inefficiency, and the size of u_{it} indicates the degree of inefficiency.

Econometric Model

The econometric model:

$$\ln Y_{it} = \beta_1 \ln L_{it} + \beta_2 \ln K_{it} + \bar{a}_t - u_{it} + v_{it}$$

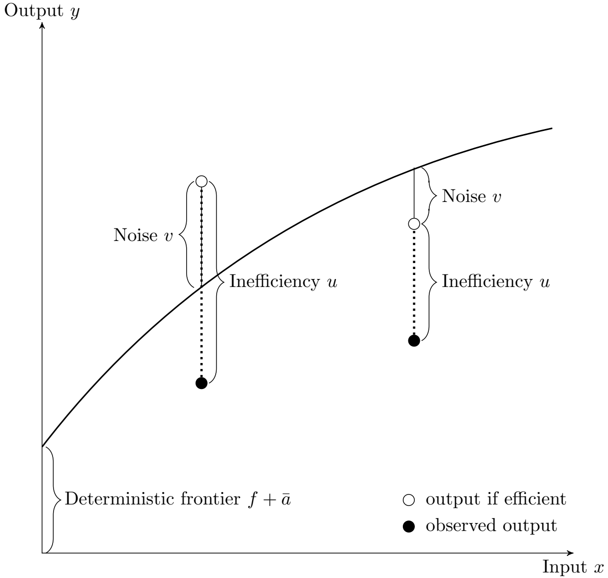
where $\ln \bar{A}_t = \bar{a}_t$, $u_{it} \geq 0$ is a measure of technical *in*efficiency and v_{it} is the idiosyncratic error term.

- $u_{it} = 0$ implies efficiency; i.e., county i is at the technological frontier at time t.
- $u_{it} > 0$ implies inefficiency, and the size of u_{it} indicates the degree of inefficiency.

$f + \bar{a}$ is the **deterministic frontier**. It represents the level of production that is technically feasible for given inputs (L, K).

$f + \bar{a} + v$ is the **stochastic frontier**. It allows for idiosyncratic shocks that affect the frontier.

Efficiency Analysis



Efficiency Analysis

How can we disentangle inefficiency & noise?

Stochastic Frontier Analysis (SFA).

Consider the cross-section model:

$$y_i = \alpha + \boldsymbol{x}_i' \boldsymbol{\beta} - u_i + v_i.$$

Traditionally SFA requires distributional assumptions.

Aigner, Lovell, and Schmidt (1977) assume:

- v_i is normally distributed (thus symmetric)
- u_i follows a one-sided distribution (half-normal or exponential)

Based on these distributional assumptions, it is possible to disentangle **idiosyncratic noise** from **technical inefficiency**.

Efficiency Analysis

How can we disentangle inefficiency & noise?

Stochastic Frontier Analysis (SFA).

Consider the panel (fixed effects) model,

$$y_{it} = \boldsymbol{x}_{it}'\boldsymbol{\beta} - \alpha_i + v_{it}.$$

The frontier is $\hat{\alpha}^\star = \max_i(\hat{\alpha}_i)$ and technical inefficiency is $\hat{u}_i = \hat{\alpha}^\star - \hat{\alpha}_i$.

Advantage: no distributional assumptions required.

Efficiency Analysis

How can we disentangle inefficiency & noise?

Stochastic Frontier Analysis (SFA).

Consider the panel (fixed effects) model,

$$y_{it} = \boldsymbol{x}_{it}'\boldsymbol{\beta} - \alpha_i + v_{it}.$$

The frontier is $\hat{\alpha}^\star = \max_i(\hat{\alpha}_i)$ and technical inefficiency is $\hat{u}_i = \hat{\alpha}^\star - \hat{\alpha}_i$.

Advantage: no distributional assumptions required.

But:

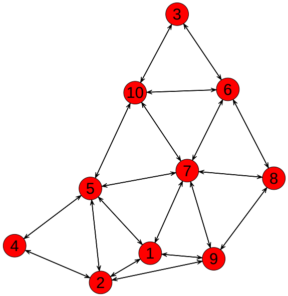
- incidental parameter problem as $N \to \infty$,
- we cannot distinguish time-invariant unobserved heterogeneity from technical inefficiency,
- assumption of time-invariant technical inefficiency.

Cross-section Dependence

There is another issue with SFA: we have to account for cross-section or spatial dependence.

We expected interaction effects between counties:

- cross-border commuting,
- knowledge spill-overs,
- national migration.



Cross-section Dependence

Two types of models for cross-section dependence.

Spatial econometric models use a deterministic spatial weights matrix to accommodate interaction effects between units.

Cross-section Dependence

Two types of models for cross-section dependence.

Spatial econometric models use a deterministic spatial weights matrix to accommodate interaction effects between units.

For example, the *spatial autoregressive model* is

$$y_{it} = \lambda \sum_{i=1}^{N} w_{ij} y_j + \boldsymbol{x}'_{it} \beta + v_{it}.$$

w_{ij} are the spatial weights which determine the strength of interaction between county i and j. The model allows county i's output to directly affect county j's output.

Cross-section Dependence

Two types of models for cross-section dependence.

Spatial econometric models use a deterministic spatial weights matrix to accommodate interaction effects between units.

For example, the *spatial autoregressive model* is

$$y_{it} = \lambda \sum_{i=1}^{N} w_{ij} y_j + \boldsymbol{x}'_{it} \beta + v_{it}.$$

w_{ij} are the spatial weights which determine the strength of interaction between county i and j. The model allows county i's output to directly affect county j's output.

Recent work has accommodated spatial effects into SFA, e.g., Tsionas and Michaelides (2016) and Glass, Kenjegalieva, and Sickles (2016).

Cross-section Dependence

There is an alternative, more general approach to cross-section dependence.

Unobserved factor models.

$$y_{it} = \boldsymbol{x}'_{it}\beta + \boldsymbol{f}'_t\gamma_i + v_{it}$$

Factor models are more general than spatial econometric models. They encompass **strong and weak** forms of cross-section dependence.

A **strong factor** could be due to global shock affecting all counties with heterogeneous intensity (e.g., Brexit, exchange rate changes, financial crisis). Its effect doesn't vanish as $N \to \infty$.

Cross-section Dependence

There is an alternative, more general approach to cross-section dependence.

Unobserved factor models.

$$y_{it} = \boldsymbol{x}'_{it}\beta + \boldsymbol{f}'_t\gamma_i + v_{it}$$

Factor models are more general than spatial econometric models. They encompass **strong and weak** forms of cross-section dependence.

A **strong factor** could be due to global shock affecting all counties with heterogeneous intensity (e.g., Brexit, exchange rate changes, financial crisis). Its effect doesn't vanish as $N \to \infty$.

Why factor models? Unobserved factors may be correlated to $\boldsymbol{x}_{it}$ and y_{it}, thus inducing a bias.

Cross-section Dependence

Unobserved factor model: $y_{it} = \boldsymbol{x}'_{it}\beta + \boldsymbol{f}'_t\gamma_i + v_{it}$

Stochastic frontier model: $y_{it} = \boldsymbol{x}'_{it}\beta - u_{it} + v_{it}$

It's easy to see that

$$\boldsymbol{f}'_t\gamma_i = -u_{it}.$$

If we estimate the factor structure, we also have an estimate of technical inefficiency!

This approach follows Mastromarco, Selenga & Shin (2013; 2016).

Cross-section Dependence

We assume

$$\boldsymbol{f}'_t\gamma_i = -u_{it} = \alpha_i + \gamma_1 t + \gamma_1 f_{2,t} + \ldots + \gamma_m f_{m,t}$$

Thus, inefficiency is determined by:

- Unobserved time-invariant heterogeneity, α_i;
- Time trend, $\gamma_1 t$;
- Additional **unobserved factors**.

Cross-section Dependence

We assume

$$\boldsymbol{f}'_t\gamma_i = -u_{it} = \alpha_i + \gamma_1 t + \gamma_1 f_{2,t} + \ldots + \gamma_m f_{m,t}$$

Thus, inefficiency is determined by:

- Unobserved time-invariant heterogeneity, α_i;
- Time trend, $\gamma_1 t$;
- Additional **unobserved factors**.

Note that we don't need to estimate each unobserved factor. We are only interested in estimating the compound term $\boldsymbol{f}'_t\boldsymbol{\gamma}_i$.

Cross-section Dependence

We assume

$$\boldsymbol{f}'_t\boldsymbol{\gamma}_i = -u_{it} = \alpha_i + \gamma_1 t + \gamma_1 f_{2,t} + \ldots + \gamma_m f_{m,t}$$

Thus, inefficiency is determined by:

- Unobserved time-invariant heterogeneity, α_i;
- Time trend, $\gamma_1 t$;
- Additional **unobserved factors**.

Note that we don't need to estimate each unobserved factor. We are only interested in estimating the compound term $\boldsymbol{f}'_t\boldsymbol{\gamma}_i$.

In contrast to the standard SFA approach, we don't need distributional assumptions about u and v, or about the specific form of the factor structure.

Cross-section Dependence

(1) Common Correlated Effects (CCE) estimator due to Pesaran (2006) is obtained by the augmented regression model:

$$y_{it} = \boldsymbol{x}'_{\boldsymbol{it}}\beta + \alpha_i + \boldsymbol{\gamma}_{\boldsymbol{i}}^{\star\prime}\boldsymbol{f}_{\boldsymbol{t}}^{\star} + \epsilon_{it},$$

where $\boldsymbol{f}_{\boldsymbol{t}}^{\star}$ is a vector of cross-sectional averages, i.e., $\boldsymbol{f}_{\boldsymbol{t}}^{\star\prime} = (\bar{y}_t, \bar{\boldsymbol{x}}'_{\boldsymbol{t}})$ with $\bar{y}_t = N^{-1}\sum_i y_{it}$ and $\bar{\boldsymbol{x}}'_{\boldsymbol{t}}$ defined accordingly.

The cross-sectional averages approximate the unobserved factor structure asymptotically as $N \to \infty$.

Cross-section Dependence

(1) Common Correlated Effects (CCE) estimator due to Pesaran (2006) is obtained by the augmented regression model:

$$y_{it} = \boldsymbol{x}_{\boldsymbol{it}}'\beta + \alpha_i + \boldsymbol{\gamma}_{\boldsymbol{i}}^{\star\prime}\boldsymbol{f}_{\boldsymbol{t}}^{\star} + \epsilon_{it},$$

where $\boldsymbol{f}_{\boldsymbol{t}}^{\star}$ is a vector of cross-sectional averages, i.e., $\boldsymbol{f}_{\boldsymbol{t}}^{\star\prime} = (\bar{y}_t, \bar{\boldsymbol{x}}_{\boldsymbol{t}}')$ with $\bar{y}_t = N^{-1}\sum_i y_{it}$ and $\bar{\boldsymbol{x}}_{\boldsymbol{t}}'$ defined accordingly.

The cross-sectional averages approximate the unobserved factor structure asymptotically as $N \to \infty$.

The estimator is consistent in a wide range of settings:

- non-stationary factors (Kapetanios, Pesaran, and Yamagata, 2011),
- weak & strong forms of CD (Pesaran and Tosetti, 2011),
- and an infinite number of non-strong factors (Chudik, Pesaran, and Tosetti, 2011).

Cross-section Dependence

(2) Interactive Fixed Effects (Bai, 2009)

Our model in matrix form is

$$\boldsymbol{Y_i} = \boldsymbol{X_i}\beta + \boldsymbol{F}\gamma_{\boldsymbol{i}} + \boldsymbol{v_i}.$$

Cross-section Dependence

(2) Interactive Fixed Effects (Bai, 2009)

Our model in matrix form is

$$\boldsymbol{Y_i} = \boldsymbol{X_i}\beta + \boldsymbol{F}\gamma_{\boldsymbol{i}} + \boldsymbol{v_i}.$$

For a given $\boldsymbol{F}$, the problem reduces to a simple least square problem and we can obtain $\hat{\beta}(\boldsymbol{F})$ using OLS.

For a given β, one can obtain $\hat{\boldsymbol{F}}$ using standard factor analysis.

Cross-section Dependence

(2) Interactive Fixed Effects (Bai, 2009)

Our model in matrix form is

$$\boldsymbol{Y_i} = \boldsymbol{X_i}\beta + \boldsymbol{F}\gamma_{\boldsymbol{i}} + \boldsymbol{v_i}.$$

For a given $\boldsymbol{F}$, the problem reduces to a simple least square problem and we can obtain $\hat{\beta}(\boldsymbol{F})$ using OLS.

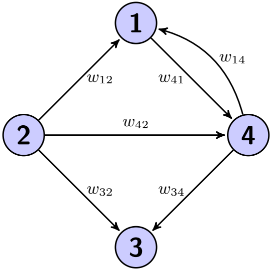
For a given $\boldsymbol{\beta}$, one can obtain $\hat{\boldsymbol{F}}$ using standard factor analysis.

Bai (2009) show that an iterative estimation procedure can be used to compute the interactive fixed effects estimator.

However, the number of factors m is assumed to be known. We adopt the framework proposed by Bada and Kneip (2014) who suggest to integrate m into the optimization problem through penalized least squares.

Spatial analysis

The standard approach in spatial econometrics is to apply pre-specified weights matrices, e.g. binary contiguity or inverse distance matrix.



Based on recent econometric results by Kock and Callot (2015) and Medeiros and Mendes (2016), we apply the adaptive LASSO to estimate the spatial weights matrix.

Spatial analysis

In the second part of the analysis, we analyse the de-factored residuals for spatial dependence.

Spatial analysis

In the second part of the analysis, we analyse the de-factored residuals for spatial dependence.

We follow the two-step approach of Holly, Pesaran, and Yamagata (2010) and Bailey, Holly, and Pesaran (2016) who suggest to first control for strong dependence, and then analyse the de-factored residuals for weak/spatial dependence.

Spatial analysis

In the second part of the analysis, we analyse the de-factored residuals for spatial dependence.

We follow the two-step approach of Holly, Pesaran, and Yamagata (2010) and Bailey, Holly, and Pesaran (2016) who suggest to first control for strong dependence, and then analyse the de-factored residuals for weak/spatial dependence.

Instead of using a pre-specified weights matrix, we employ the adaptive LASSO (Zou, 2006) for each county:

$$v_{it} = \sum_{i=1}^{N} w_{ij} v_{i,t-1} + \varepsilon_{it},$$

where we replace v_{it} by the de-factored residuals.

Econometric Results

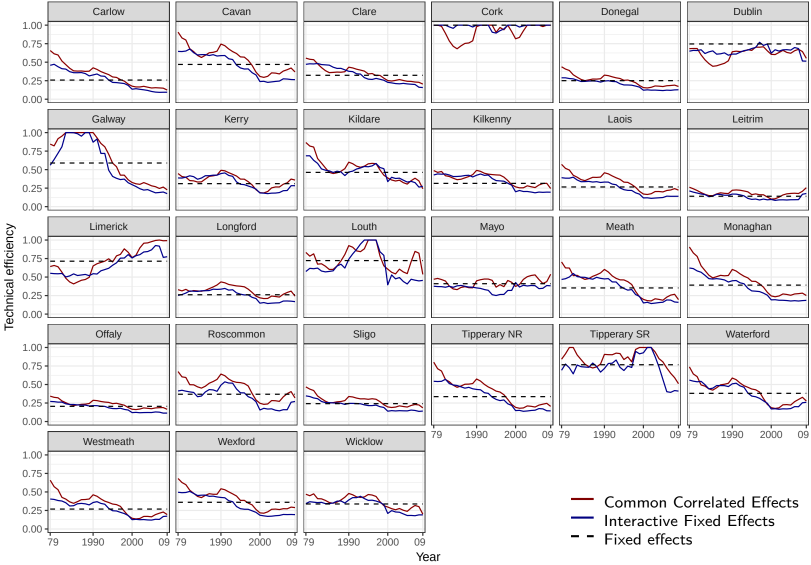
Production function

	Dependent variable: $\ln Y_{it}$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln K_{it}$	0.394^{***}	0.339^{***}	0.346^{***}	0.241^{**}	0.189^{**}	0.170^{**}
	(0.034)	(0.031)	(0.121)	(0.118)	(0.075)	(0.071)
$\ln L_{it}$	0.359^{***}		0.521^{***}		0.582^{***}	
	(0.059)		(0.122)		(0.047)	
$\ln L_{it}$ (dom.)		-0.088^{*}		0.155^{*}		0.301^{***}
		(0.051)		(0.091)		(0.048)
$\ln L_{it}$ (for.)		0.339^{***}		0.312^{***}		0.241^{***}
		(0.030)		(0.094)		(0.026)
Trend	0.039^{***}	0.045^{***}				
	(0.003)	(0.002)				
Model	FE	FE	CCEP	CCEP	IFE	IFE
Trend	common	common	heterogenous	heterogenous	–	–
CD test	19.98	13.37	-3.17	-3.56	-3.63	-3.63
Observations	837	837	837	837	837	837

We distinguish between employment in foreign-owned and domestic companies to account for FDI and the effects of MNEs.

Econometric Results

Technical efficiency score



Econometric Results

Technical efficiency score

- **Cork:** concentration of pharmaceutical firm (e.g. Pfizer, Novartis) & chemical industry. — Role of transfer pricing?
- **Limerick** (manufacture of medical/electrical devices) is the only country that shows some evidence of catch-up relative to Cork.

Econometric Results

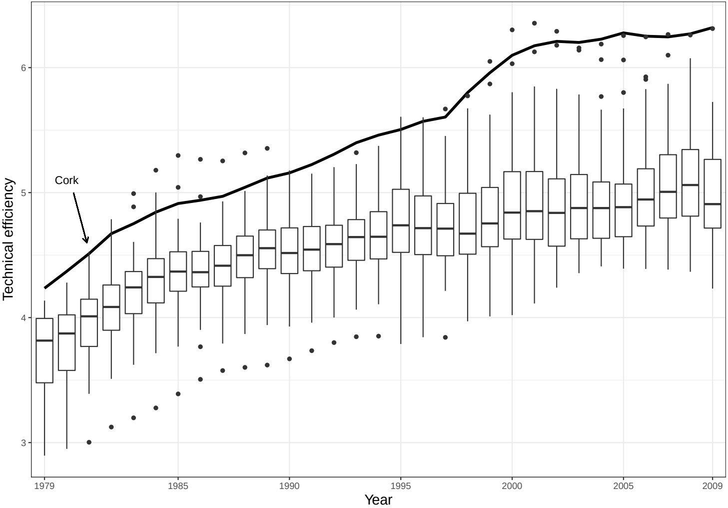
Efficiency ranking: average TE score by county

County	FE	CCE	IFE
Cork	1.000 (1)	0.914 (1)	1.000 (1)
Tipperary SR	0.767 (2)	0.850 (2)	0.767 (2)
Louth	0.722 (4)	0.742 (3)	0.722 (4)
Limerick	0.714 (5)	0.726 (4)	0.714 (5)
Galway	0.589 (6)	0.646 (5)	0.589 (6)
Dublin	0.746 (3)	0.612 (6)	0.746 (3)
Cavan	0.469 (7)	0.546 (7)	0.469 (7)
Kildare	0.463 (8)	0.499 (8)	0.463 (8)
Monaghan	0.391 (10)	0.455 (9)	0.391 (10)
Roscommon	0.369 (12)	0.451 (10)	0.369 (12)
Mayo	0.410 (9)	0.430 (11)	0.410 (9)
Waterford	0.381 (11)	0.407 (12)	0.381 (11)
Meath	0.353 (14)	0.406 (13)	0.353 (14)
Tipperary NR	0.336 (16)	0.405 (14)	0.336 (16)
Wexford	0.360 (13)	0.401 (15)	0.360 (13)
Kilkenny	0.314 (18)	0.372 (16)	0.314 (18)
Wicklow	0.338 (15)	0.370 (17)	0.338 (15)
Kerry	0.311 (19)	0.350 (18)	0.311 (19)
Clare	0.323 (17)	0.348 (19)	0.323 (17)
Laois	0.266 (21)	0.329 (20)	0.266 (21)
Westmeath	0.269 (20)	0.325 (21)	0.269 (20)
Carlow	0.258 (23)	0.320 (22)	0.258 (23)
Longford	0.261 (22)	0.315 (23)	0.261 (22)
Sligo	0.241 (25)	0.278 (24)	0.241 (25)
Donegal	0.250 (24)	0.254 (25)	0.250 (24)
Offaly	0.205 (26)	0.231 (26)	0.205 (26)
Leitrim	0.140 (27)	0.181 (27)	0.140 (27)

Note: Rank shown in parentheses.

Econometric Results

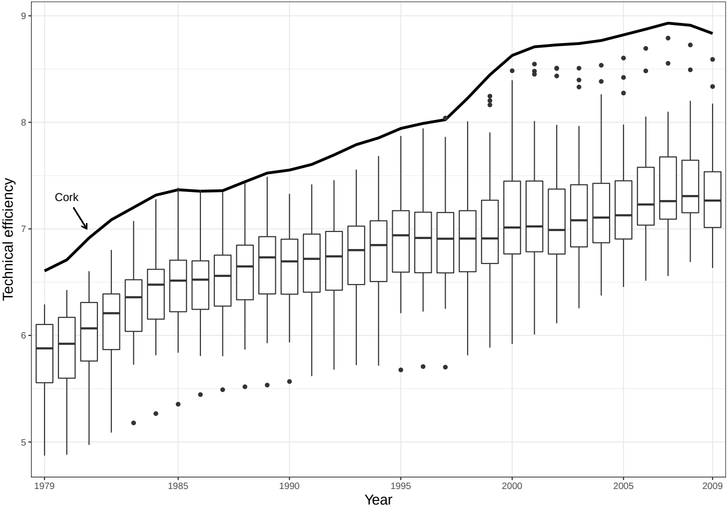
Production frontier



The frontier (represented by Cork) increases over time due to *technological progress*. — Evidence of *technological divergence* between the frontier and the median county.

Econometric Results

Production frontier



The frontier (represented by Cork) increases over time due to *technological progress*. — Evidence of *technological divergence* between the frontier and the median county.

Growth decomposition

Average annual growth over 1979-2009 by county and component (in %).

County	Output	Output per worker	Labour	Capital	T.E.	Unexplained
Cavan	6.49	5.57	0.87	6.51	-2.96	-0.29
Donegal	5.61	8.15	-2.35	7.58	-3.06	0.33
Leitrim	6.73	7.88	-1.06	1.83	-0.08	-0.42
Louth	7.52	9.84	-2.11	8.13	-1.46	0.18
Monaghan	5.36	5.20	0.15	7.55	-4.19	-0.04
Sligo	6.37	6.66	-0.27	6.56	-3.05	0.27
Laois	4.99	5.09	-0.09	5.95	-3.08	-0.89
Longford	8.89	7.47	1.33	3.22	-1.02	0.81
Offaly	7.47	6.00	1.39	5.15	-2.40	0.23
Westmeath	7.55	7.69	-0.13	11.70	-3.98	0.63
Mayo	9.72	10.52	-0.73	8.18	0.45	-0.46
Roscommon	8.91	8.60	0.28	7.69	-2.47	1.39
Dublin	7.21	10.25	-2.75	7.95	-0.71	-0.47
Kildare	8.39	6.74	1.54	14.67	-4.16	-0.17
Meath	5.95	5.68	0.26	8.38	-4.19	0.19
Wicklow	8.30	8.81	-0.47	9.74	-2.80	0.90
Clare	4.89	5.59	-0.66	5.77	-3.27	-0.44
Tipperary NR	4.34	4.68	-0.33	7.08	-4.45	-0.34
Carlow	2.94	4.23	-1.23	7.86	-5.41	-0.46
Kilkenny	4.47	5.77	-1.23	2.49	-2.27	-0.49
Tipperary SR	8.88	7.38	1.40	10.48	-1.65	-0.94
Wexford	6.20	5.70	0.47	5.67	-2.82	-0.22
Kerry	6.81	7.98	-1.09	2.67	-0.73	0.03
Cork	10.09	10.30	-0.18	9.17	0.00	-0.27
Galway	7.47	6.05	1.34	10.23	-4.29	0.58
Limerick	12.61	11.86	0.67	10.30	1.46	-0.26
Waterford	6.49	7.51	-0.94	7.29	-3.21	0.67

Level decomposition

Variance decomposition (in %):

	Inputs			TFP
	Total	*Labour*	*Kapital*	
CCE (base)	58.80	28.02	30.78	41.20
IFE (base)	48.19	26.78	21.41	51.81
CCE (extended)	47.93	31.14	16.80	52.07
IFE (extended)	44.05	29.00	15.06	55.95

Convergence

The decomposition allows us to test for convergence by component.

	Dependent variable: Annual growth rate over 1979-2009				
	Y	Y/L	K	L	TE
Log of 1979 level	0.410	-1.847^{*}	1.135^{**}	1.321^{***}	-0.708
	(0.363)	(0.983)	(0.578)	(0.251)	(0.886)
Contant	2.048	13.964^{***}	3.306	-4.986^{**}	0.115
	(4.424)	(3.492)	(6.594)	(2.151)	(3.196)
Observations	27	27	27	27	27
R^2	0.031	0.091	0.016	0.137	0.018

There is limited evidence of convergence in labour productivity, and no evidence for convergence in technical efficiency.

But the results indicate divergence in output and capital!

Spatial analysis

The full model is given by

$$y_{it} = \boldsymbol{x}_{it}'\boldsymbol{\beta} - u_{it} + v_{it}$$
$$v_{it} = \sum_{i=1}^{N} w_{ij} v_{i,t-1} + \varepsilon_{it}$$

We employ the adaptive Lasso due to Zou (2006) to estimate w_{ij}.

The adaptive Lasso allows for estimation and selection of spatial lags under the assumption of sparsity, even if N is large relative to T.

Spatial analysis

The adaptive Lasso minimizes (for each i)

$$\frac{1}{T}\sum_{t=1}^{T}(v_{it}-\sum_{j=1}^{N}w_{ij}v_{i,t-1})^2+\frac{\lambda}{T}\sum_{j\neq i}|w_{ij}|/|\hat{w}_{ij}^0|.$$

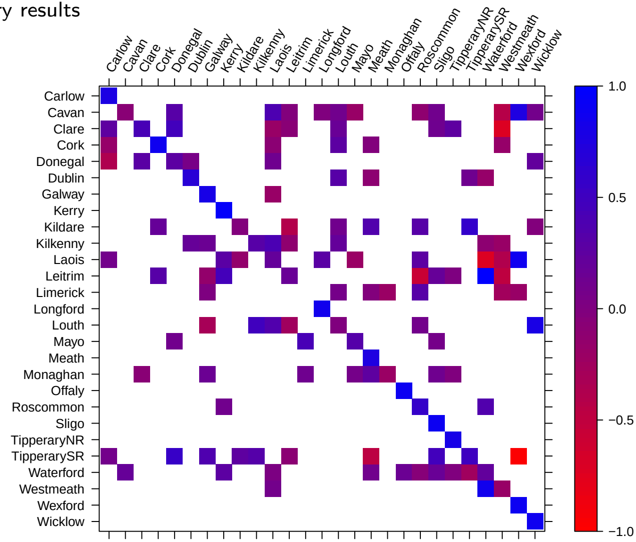
where $\hat{w}_{ij}^0$ is a first-stage estimator, e.g., standard Lasso or Ridge.

Recent econometric results show that the adaptive Lasso is **model selection consistent** when applied to time-series data with non-Gaussian & heteroskedastic errors (Kock and Callot, 2015; Medeiros and Mendes, 2016).

We select λ using the Extended BIC (Chen & Chen, 2008, 2012) which is more approriate in high-dimensional settings than AIC/BIC.

Spatial analysis

Preliminary results



Using Stata packages **lassovar** (unpublished) and **lassopack** (Ahrens, Schaffer, Hansen, 2018).

Discussion

Based on Mastromarco, Selenga & Shin (2013; 2016), we use CCE & IFE to estimate the time-varying efficiency.

We identify Cork as the technological leader in the industrial sector of Ireland. Results are consistent across three estimation methods (FE, CCE, IFE).

There is no evidence for convergence in terms of technical inefficiency. Graphical analysis even suggests technological divergence across counties relative to the frontier (represented by Cork).

With regard to labour & capital accumulation, we find statistical evidence for **divergence**.

Labour & capital divergence might evidence of clustering (technology clusters, largely driven by foreign companies).

Open question: What drives (lack of) technological catch-up? — Leading candidates: infrastructure, human capital, cluster effects.

Discussion

Following Holly, Pesaran, and Yamagata (2010), we apply a two-step approach where we analyse the de-factored residuals for spatial dependence.

Rather than relying on *ad-hoc* specifications, we employ the adaptive LASSO to estimate the spatial weights matrix.

The framework allows to gain insights into drivers of spatial spill-overs.

References I

- Aigner, Dennis, C.A.Knox Lovell, and Peter Schmidt (1977). “Formulation and estimation of stochastic frontier production function models”. In: *Journal of Econometrics* 6.1, pp. 21–37. URL: http://www.sciencedirect.com/science/article/pii/0304407677900525.
- Bada, Oualid and Alois Kneip (2014). “Parameter cascading for panel models with unknown number of unobserved factors: An application to the credit spread puzzle”. In: *Computational Statistics & Data Analysis* 76.Supplement C, pp. 95–115. URL: http://www.sciencedirect.com/science/article/pii/S0167947313004453.
- Bai, Jushan (2009). “Panel Data Models With Interactive Fixed Effects”. In: *Econometrica* 77.4, pp. 1229–1279. URL: http://dx.doi.org/10.3982/ECTA6135.

References II

Bailey, Natalia, Sean Holly, and M Hashem Pesaran (2016). "A Two Stage Approach to Spatiotemporal Analysis with Strong and Weak Cross-Sectional Dependence". In: *Journal of Applied Econometrics* 31.1, pp. 248–280. URL: http://dx.doi.org/10.1002/jae.2468.

Chudik, Alexander, M Hashem Pesaran, and Elisa Tosetti (2011). "Weak and strong cross-section dependence and estimation of large panels". In: *The Econometrics Journal* 14.1, pp. C45–C90. URL: http://dx.doi.org/10.1111/j.1368-423X.2010.00330.x.

Glass, Anthony J., Karligash Kenjegalieva, and Robin C. Sickles (2016). "A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers". In: *Journal of Econometrics* 190.2, pp. 289–300. URL: http://linkinghub.elsevier.com/retrieve/pii/S0304407615001803.

References III

Holly, Sean, M Hashem Pesaran, and Takashi Yamagata (2010). "A spatio-temporal model of house prices in the USA". In: *Journal of Econometrics* 158.1, pp. 160–173. URL: http://www.sciencedirect.com/science/article/pii/S0304407610000837.

Kapetanios, G, M Hashem Pesaran, and Takashi Yamagata (2011). "Panels with non-stationary multifactor error structures". In: *Journal of Econometrics* 160.2, pp. 326–348. URL: http://www.sciencedirect.com/science/article/pii/S0304407610002022.

Kock, Anders Bredahl and Laurent Callot (2015). "Oracle inequalities for high dimensional vector autoregressions". In: *Journal of Econometrics* 186.2, pp. 325–344. URL: http://www.sciencedirect.com/science/article/pii/S0304407615000378.

References IV

Mastromarco, Camilla, Laura Serlenga, and Yongcheol Shin (2013). "Globalisation and technological convergence in the EU". In: *Journal of Productivity Analysis* 40.1, pp. 15–29. URL: https://doi.org/10.1007/s11123-012-0308-9.

– (2016). "Modelling Technical Efficiency in Cross Sectionally Dependent Stochastic Frontier Panels". In: *Journal of Applied Econometrics* 31.1, pp. 281–297. URL: http://dx.doi.org/10.1002/jae.2439.

Medeiros, Marcelo C and Eduardo F Mendes (2016). "L1-regularization of high-dimensional time-series models with non-Gaussian and heteroskedastic errors". In: *Journal of Econometrics* 191.1, pp. 255–271. URL: http://www.sciencedirect.com/science/article/pii/S0304407615002638.

References V

Pesaran, M Hashem (2006). “Estimation and inference in large heterogeneous panels with a multifactor error structure”. In: *Econometrica* 74.4, pp. 967–1012.

Pesaran, M Hashem and Elisa Tosetti (2011). “Large panels with common factors and spatial correlation”. In: *Journal of Econometrics* 161, pp. 182–202.

Tsionas, Efthymios G and Panayotis G Michaelides (2016). “A Spatial Stochastic Frontier Model with Spillovers: Evidence for Italian Regions”. In: *Scottish Journal of Political Economy* 63.3, pp. 243–257. URL: http://dx.doi.org/10.1111/sjpe.12081.

Zou, Hui (2006). “The Adaptive Lasso and Its Oracle Properties”. In: *Journal of the American Statistical Association* 101.476, pp. 1418–1429.