



# Composite indicators. Methodology and Applications to R&D indicators

*by Andrea Saltelli  
EC, JRC*

*andrea.saltelli@jrc.it*

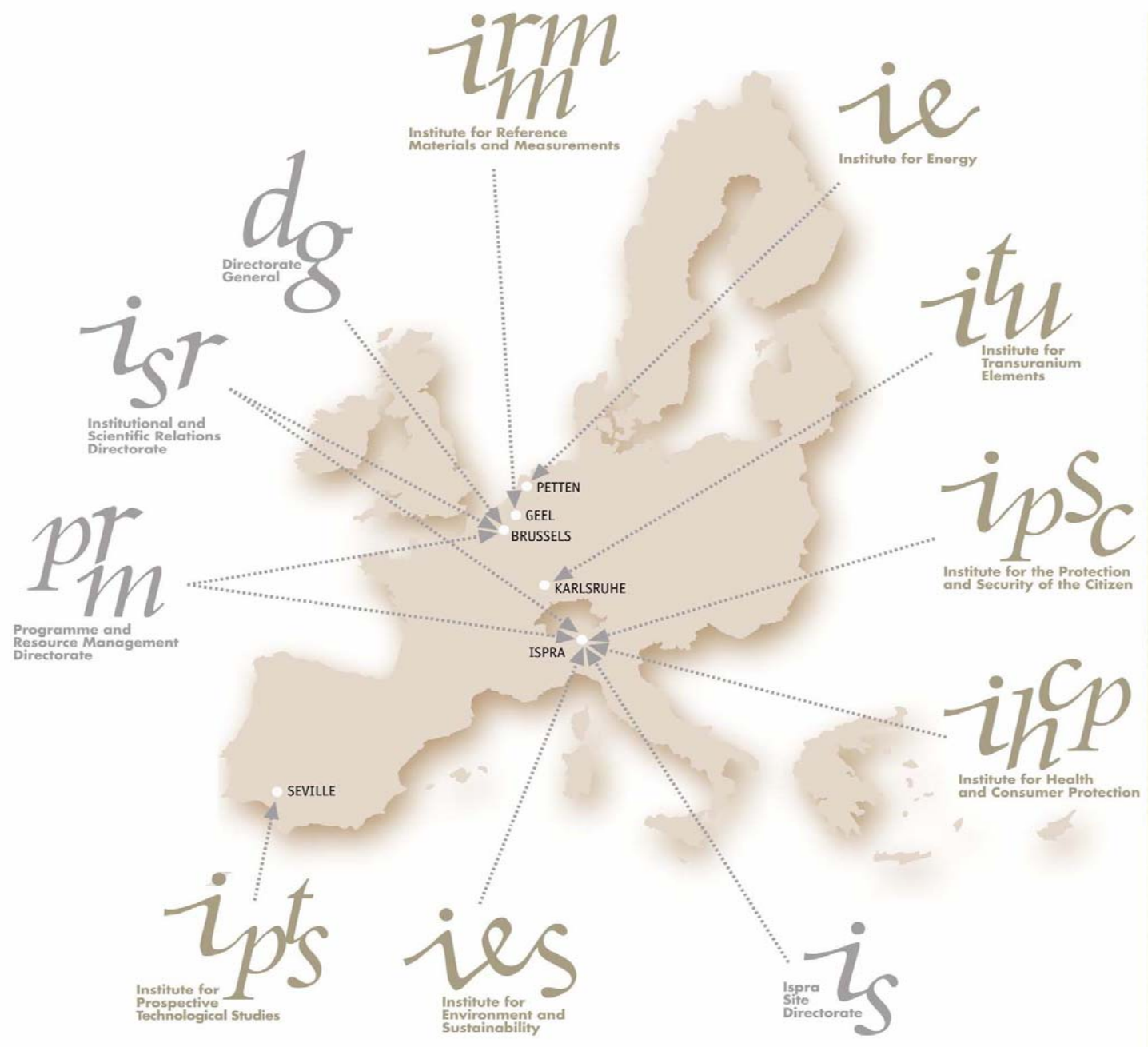
Workshop on Entrepreneurship  
Indicators:

R&D and Technology

Copenhagen, 30 and 31 October 2008



# Joint Research Centre





# Whereabouts







Borgomanero

Verbania

Lesa

Besozzo

Meina

Gavirate

Angera

Arona

Sesto Calende

Varese

Malnate

Varallo Pombia

© 2006 Europa Technologies Jambardo

Image © 2006 TerraMetrics

© 2006 Google™





# Methodology from: Joint OECD-JRC **handbook.**

- **5 years** of preparation,
- **2 rounds** of consultation with **OECD high level statistical committee,**
- finally endorsed March 2008 with one abstention



Handbook on Constructing Composite Indicators METHODOLOGY AND USER GUIDE





- Composite indicators' fortune



How many of them are around?  
Searching "composite indicators"



	Scholar Google
October 2005	992
June 2006	1,440
May 2007	1,900
October 2008	3,030



Use: an example, from the Economist web edition, October 15

**Economist.com**

**The Economist's house-price indicators**  
% change

	Latest	Q2 2007	1997-2008
	on a year earlier		or latest
Hong Kong	23.9	7.6	-23
Singapore	20.1	21	na
Australia	8.2	10.1	173
Belgium	7.5	9.5	147
China	7.0	6.3	na
Sweden	4.9	9.9	149
Italy	4.3	5.6	107
France	4.2	6.8	149
South Africa	3.2	15.5	398
Canada	3.1	7.9	70
Switzerland	2.7	1.4	21
Netherlands	2.3	4.5	105
Spain	2.0	5.8	194
Japan	-0.7	-0.7	-32
United States (OFHEO)	-1.7	3.4	91
Germany	-1.7	-0.8	na
Denmark	-1.9	4.9	123
New Zealand	-4.5	13.7	117
Ireland	-9.4	2.9	210
Britain	-10.5	10.5	179
United States (S&P Case-Shiller national index)	-15.4	-3.4	86
United States (S&P Case-Shiller ten-city index)	-17.0	-3.3	125

Sources: ABSA; ESRI; Hypoport; Japan Real Estate Institute; Nationwide; Nomisma; NWM; OFHEO; Quotable Value; Stadim; Swiss National Bank; Standard & Poor's; government offices

Markets and data

All markets and data

Daily chart

Weekly indicators

World markets

Currencies

Rankings

Big Mac index

**Business travel ranking**

(The most and least attractive destinations for business travellers)

**Distance-learning MBAs**

(A ranking of the top distance-learning MBA programmes)

**Democracy index**

(The world's most and least democratic countries)

**Operational risk ratings**

(Rating the risk of doing business in 150 markets)

**E-readiness**

(The best countries for conducting electronic business)

**Global house-price indices**

(Tracking house-price inflation around the world)







Another example, about Mauritius,  
Economist October 16

**Economist.com**

THE **1.3m people** of Mauritius love to prove famous people wrong. On independence from Britain in 1968, pundits such as a Nobel prize-winning economist, James Meade, and a novelist, V.S. Naipaul, did not give much of a chance to this tiny, isolated Indian Ocean island 1,800km (1,100 miles) off the coast of east Africa. Its people depended on a sugar economy and enjoyed a GDP per person of only \$200. Yet the island now boasts a **GDP** per person of **\$7,000**, and very few of its people live in absolute poverty. It once again **ranks first** in the latest annual **Mo Ibrahim index**, which measures governance in Africa. And it bagged **24th** spot in the World Bank's global ranking for **ease of doing business**—the only African country in the top 30, ahead of countries such as Germany and France. How does it pull it off?



Ideally, a composite indicator should be based on:

- Solid theoretical framework,
- Underlying data of good quality,
- Tested methodology.

When these conditions are met, it can be used for advocacy and analysis.

Saltelli, A., 2007, Composite indicators between analysis and advocacy, *Social Indicators Research*, **81** , 65-77.



- The steps to build CIs





A good technical preparation for a CI can make it more **robust** (to uncertainties in data, weights,...) more **resilient** (remain relevant over time), more **defensible** (in dialogue with stakeholders...)



## Step 1. Developing a solid theoretical framework

*What is badly defined is likely to be badly measured ...*

The challenges are:

- To integrate a broad set of (probably conflicting) points of view while keeping within a manageable construct. **A model.**
  - ➔ a storyteller
  - ➔ a representation/narrative meaningful to a community of peers

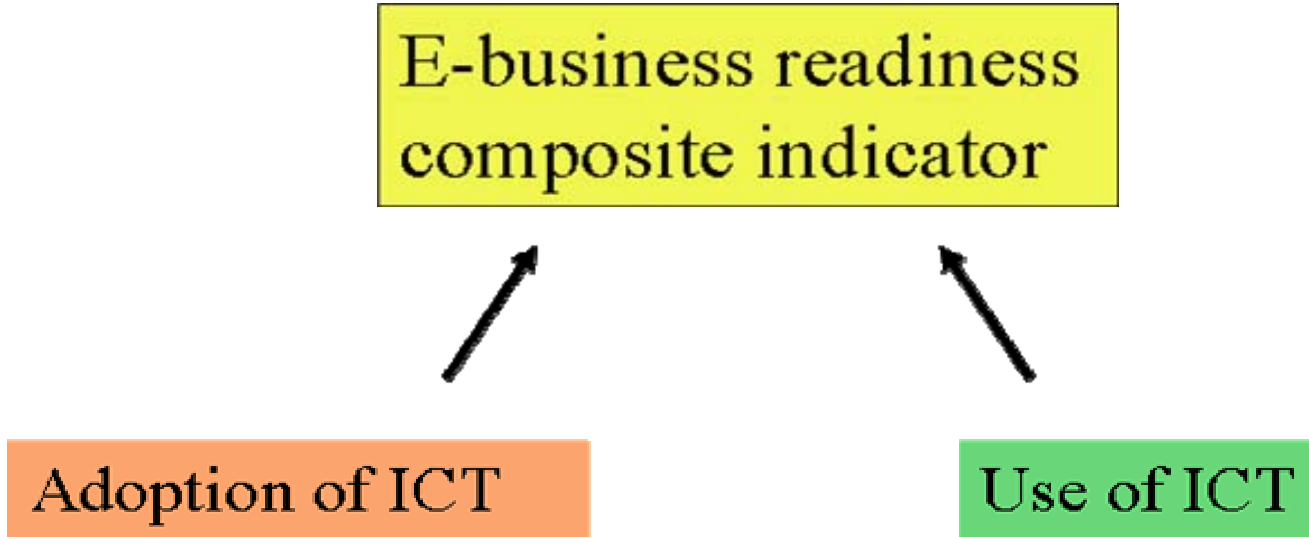
Examples ...



## E-business readiness (EC)

Definition: Capability of a company to **engage in electronic transactions** with the objective of increasing business competitiveness.

Aim: evaluation of a composite indicator to **monitor country progress** in the implementation of the i2010 initiative...

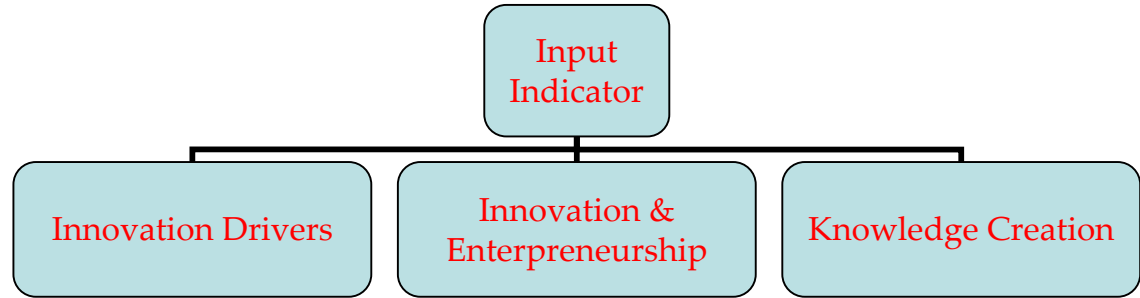






# Summary innovation Index

A theoretical framework consisting of 5 groups was discussed and agreed.



**Innovation drivers**, to measure the structural conditions required for innovation potential

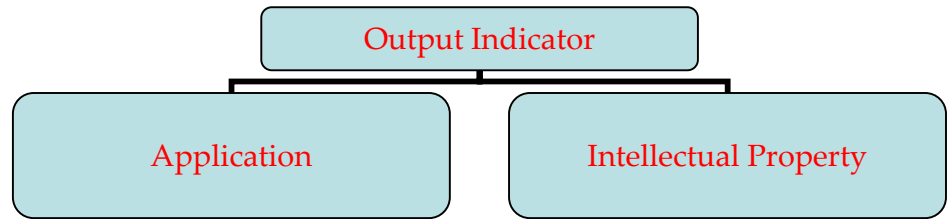
**Knowledge creation**, to measure the investments on human factors and on R&D activities, considered as the key elements for a successful knowledge-based economy

**Innovation & entrepreneurship**, to measure the efforts towards innovation at the microeconomic level

33 countries: EU-27, US, JP, CH, IS, NO, TR



# Summary innovation Index



**Application**, to measure the performance, expressed in terms of labour and business activities, and their value added in innovative sectors

**Intellectual property**, to measure the achieved results in terms of successful know how, especially referred to high-tech sectors

33 countries: EU-27, US, JP, CH, IS, NO, TR



## Step 1. Developing a solid theoretical framework

After Step 1. the developer should have...

- A clear understanding and **definition** of the multidimensional phenomenon to be measured.
- A nested **structure** of the various domains.
- A list of **selection criteria** for the underlying indicators, e.g., input, process, output.

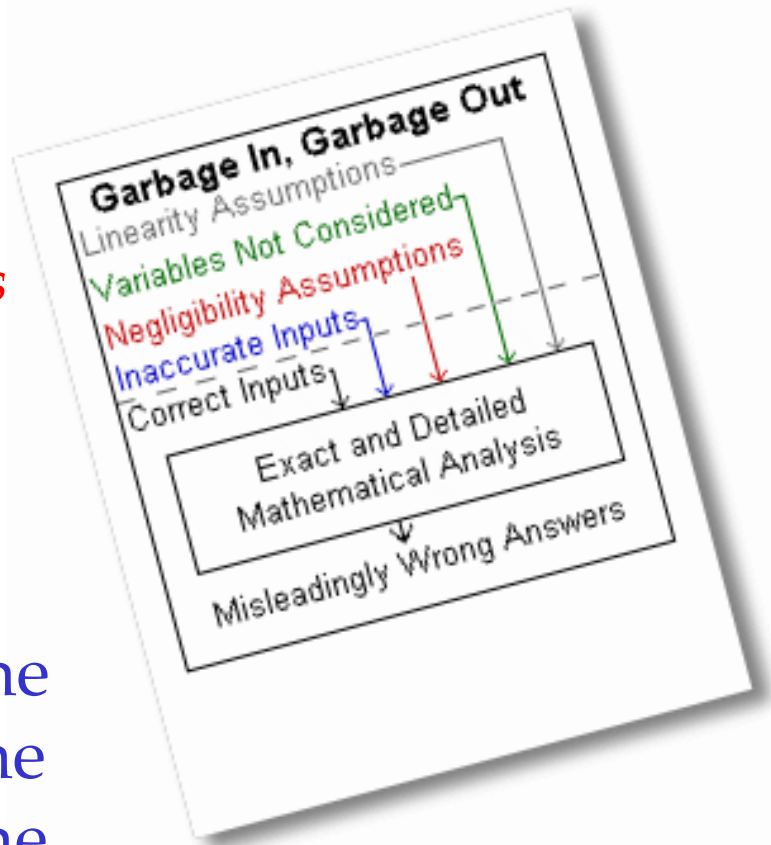




## Step 2. Selecting indicators

*A composite indicator is above all the sum of its parts...*

Excerpt: The strength of a composite indicator can largely depend on the **quality** of the underlying data. [...]. The theoretical **framework** should guide the choice of the underlying indicators. The selection process can be quite **subjective** and therefore should involve **stakeholders**.





Enterprise survey 2007 (**Eurostat**), **yearly** since 2003,  
**Adoption and Use** composite indicators

Joint Research Centre

**E-business readiness  
 composite indicator**

**Adoption of ICT**

**Use of ICT**

- a1** - % of firms that use Internet
- a2** - % of firms that have web/home page
- a3** - % of firms using  $\geq 2$  security facilities
- a4** - % of employees using computer
- a5** - % of firms with broadband connection
- a6** - % of firms with LAN

- b1** - % of firms purchasing products/serv. via Internet
- b2** - % of firms receiving orders via Internet
- b3** - % of firms with IT linked with other internal IT
- b4** - % of firms with IT linked with external IT
- b5** - % of firms using Internet for banking
- b6** - % of firms selling products via Internet



- a1** - % of firms that use Internet
- a2** - % of firms that have web/home page
- a3** - % of firms using  $\geq 2$  security facilities
- a4** - % of employees using computer
- a5** - % of firms with broadband connection
- a6** - % of firms with LAN

Adoption of ICT



- b1** - % of firms purchasing products/serv. via Internet
- b2** - % of firms receiving orders via Internet
- b3** - % of firms with IT linked with other internal IT
- b4** - % of firms with IT linked with external IT
- b5** - % of firms using Internet for banking
- b6** - % of firms selling products via Internet

Use of ICT



# Summary Innovation Index

- Data from Eurostat Science and Technology Indicators and Community Innovation Survey (CIS).

- A set of 53 indicators analyzed (FA)
- A reduced list of 26 indicators was discussed by developers and members of **GSO** (an expert group) with DG ENTR.







## Step 2. Selecting indicators

After Step 2. the developer should have...

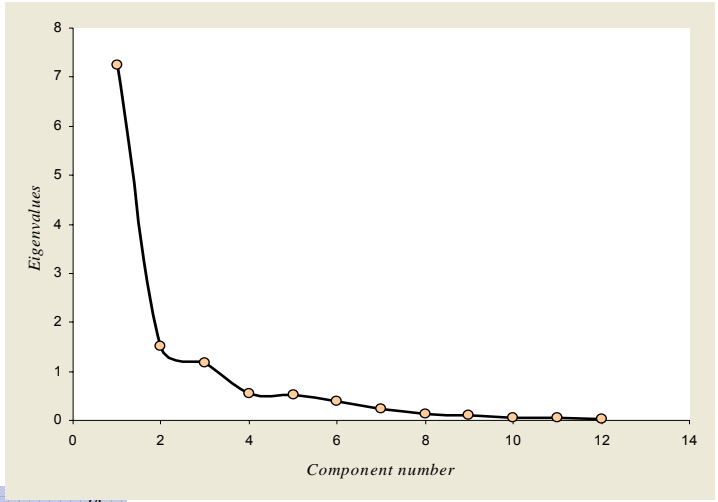
- Checked the **quality** of the available data.
- Discussed the strengths and weaknesses of each selected indicator.
- Metadata



## Step 3. Multivariate analysis

*Analysing the underlying structure of the data is an art ...*

a1	1																							
a2	<b>0.85</b>	1																						
a3	<b>0.83</b>	<b>0.87</b>	1																					
a4	<b>0.77</b>	<b>0.81</b>	<b>0.89</b>	1																				
a5	0.56	0.55	<b>0.72</b>	<b>0.72</b>	1																			
a6	0.64	0.61	0.66	<b>0.73</b>	0.56	1																		
b1	0.50	<b>0.72</b>	0.69	0.63	0.40	0.54	1																	
b2	0.61	<b>0.80</b>	<b>0.75</b>	<b>0.74</b>	0.62	0.54	<b>0.86</b>	1																
b3	0.43	0.39	0.54	0.54	0.39	0.48	0.10	0.29	1															
b4	0.35	0.26	0.41	0.30	0.31	0.37	0.16	0.24	<b>0.75</b>	1														
b5	0.70	0.49	0.48	0.45	0.60	0.31	0.09	0.38	0.21	0.13	1													
b6	0.55	<b>0.75</b>	<b>0.78</b>	<b>0.77</b>	0.59	0.57	<b>0.77</b>	<b>0.83</b>	0.44	0.24	0.29	1												
a1	a2	a3	a4	a5	a6	b1	b2	b3	b4	b5	b6													



Component	Total	Initial Eigenvalues	
		% of variance	Cumulative %
1	7.242	60.35	60.35
2	1.523	12.69	73.04
3	1.178	9.82	82.86
4	0.554	4.61	87.47
5	0.512	4.26	91.73
6	0.385	3.20	94.94
7	0.242	2.01	96.95
8	0.131	1.09	98.04
9	0.098	0.82	98.04
10	0.064	0.54	99.40
11	0.043	0.36	99.76
12	0.029	0.24	100.00



# Step 4. Imputation of missing data.

*The idea of imputation could be both seductive and dangerous ...*

Almost all datasets contain missing data.

E-business



Country	a1	a2	a3	a4	a5	a6	b1	b2	b3	b4	b5	b6
AT	0.97	0.78	0.84	0.53	0.72	0.34	0.42	0.18	0.49	0.19	0.85	0.02
BE	0.97	0.72	0.82	0.59	0.86	0.45	0.43	0.18	0.49	0.18	0.87	0.02
BG	0.75	0.31	0.54	0.20	0.61	0.16	0.03	0.01	0.17	0.05	0.44	0.01
CY	0.88	0.47	0.67	0.42	0.69	0.20	0.12	0.07	0.44	0.06	0.54	0.01
CZ	0.95	0.71	0.77	0.40	0.77	0.31	0.22	0.09	0.31	0.08	0.87	0.01
DE	0.95	0.78	0.90	0.61	0.80	0.47	0.52	0.24	0.52	0.19	0.76	0.05
DK	0.97	0.84	0.93		0.80	0.39	0.36	0.33	0.62	0.21	0.93	0.04
EE	0.94	0.62	0.74	0.39	0.78	0.26	0.13	0.07	0.24	0.11	0.93	0.03
EL	0.93	0.60	0.52	0.37	0.72	0.35	0.08	0.06			0.71	0.01
ES	0.94	0.49	0.76	0.49	0.90	0.27	0.16	0.08	0.30	0.11	0.81	0.02
EU27	0.93	0.63	0.80	0.49	0.77	0.34	0.29	0.15	0.41	0.14	0.77	0.03
FI	0.99	0.81	0.97	0.70	0.91	0.47	0.19	0.15	0.53	0.15	0.91	0.03
FR	0.96	0.57	0.82	0.52	0.89	0.35			0.43	0.11	0.76	0.02



Three common approaches to deal with missing data:

- case deletion (removes either country or indicator from the analysis)
- single imputation (e.g. Mean/Median substitution, Regression, etc.)
- multiple imputation (e.g. Markov Chain Monte Carlo algorithms).



## Step 5. Normalisation of data

*Avoid adding up apples and oranges ...*

- Ranking
- Standardization
- Re-scaling
- Distance to reference country
- Categorical scales
- ...





## Step 6. Weighting and aggregation

*The relative importance of the indicators can become the substance of a negotiation ...*

### **Weights based on statistical models**

- Principal component/Factor analysis
- Data envelopment analysis
- Regression approach
- Unobserved components models



## Step 6. Weighting and aggregation

### **Weights based on opinions: participatory methods**

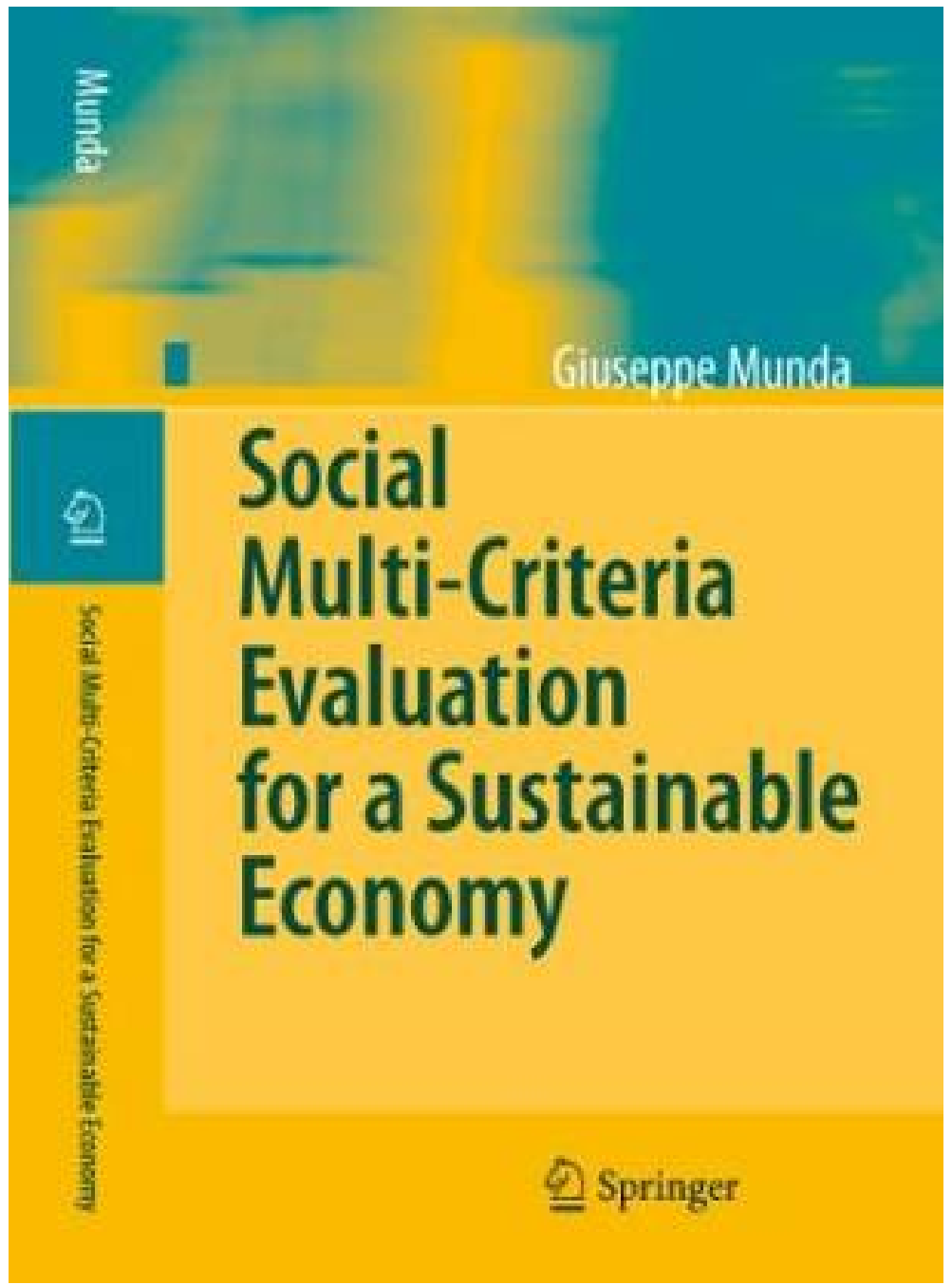
- Budget allocation
- Public opinion
- Analytic hierarchy process
- Conjoint analysis



Step 6. Weighting and aggregation

## Weights based on Multi Criteria Analysis

To know more buy:

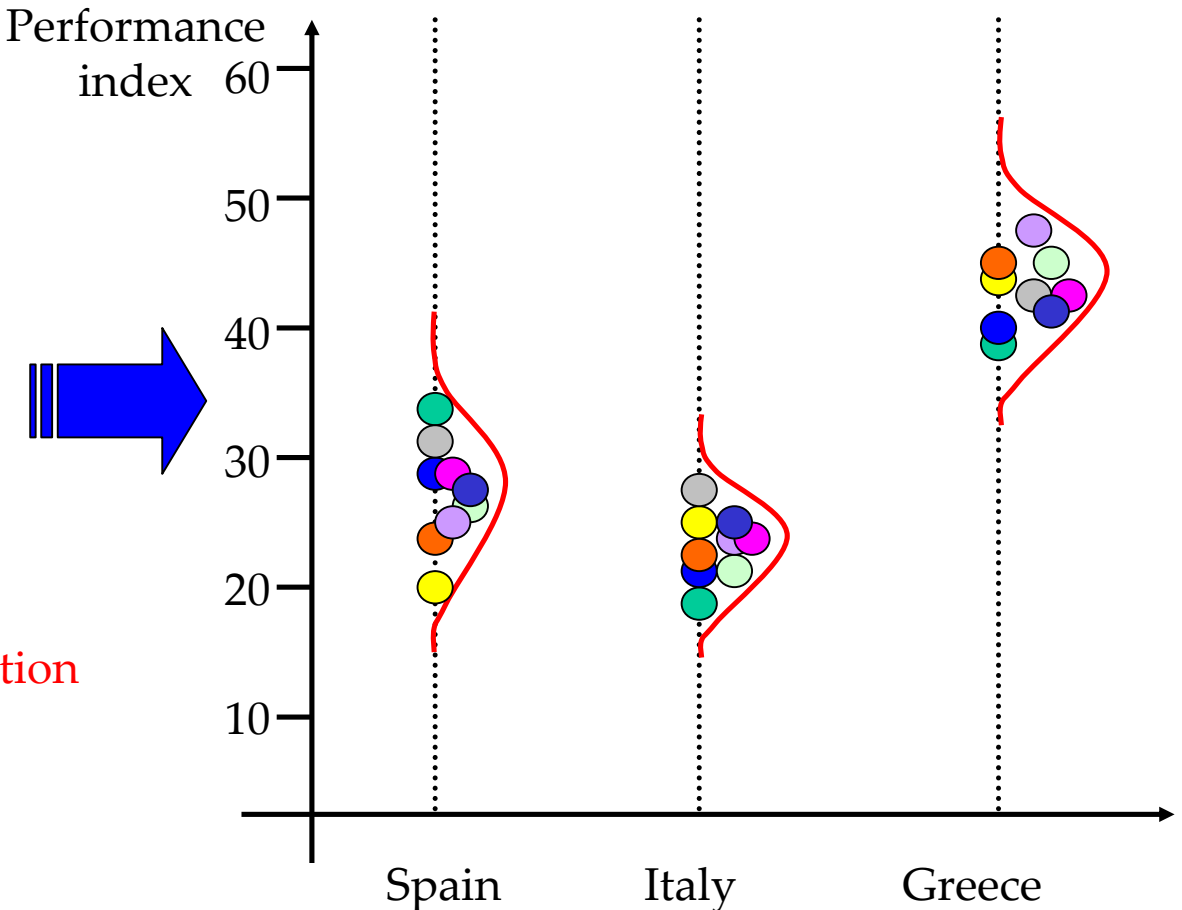
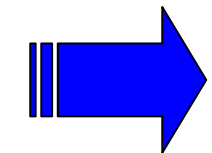
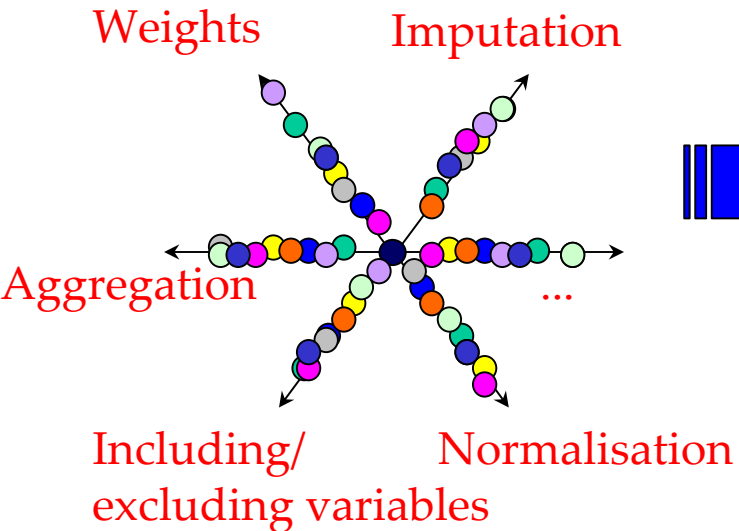




# Step 7. Robustness and sensitivity

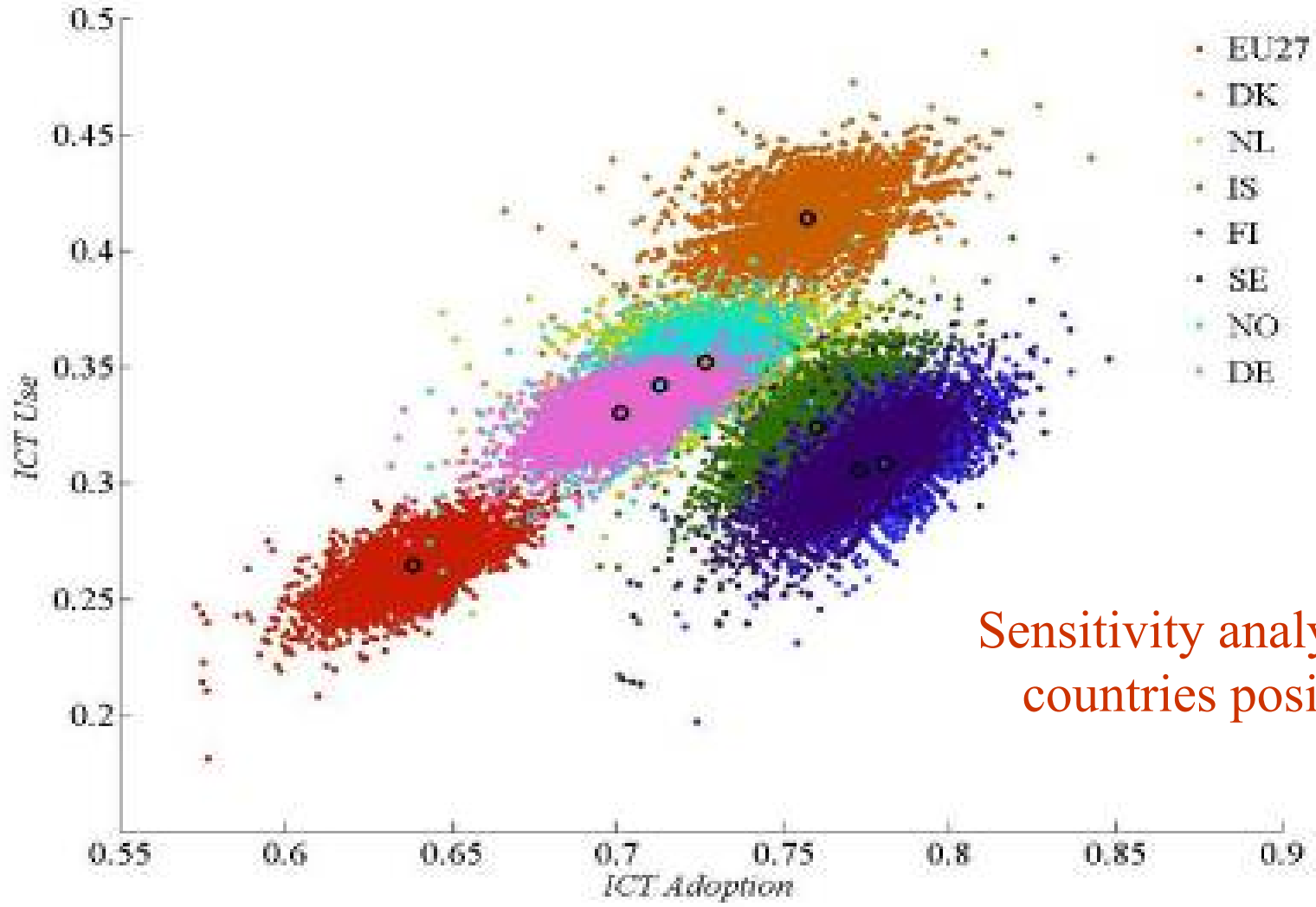
*Uncertainty analysis can be used to assess the robustness of composite indicators ...*

## Space of alternatives





# Monitoring e-Business Readiness of European enterprises



Sensitivity analysis of countries position

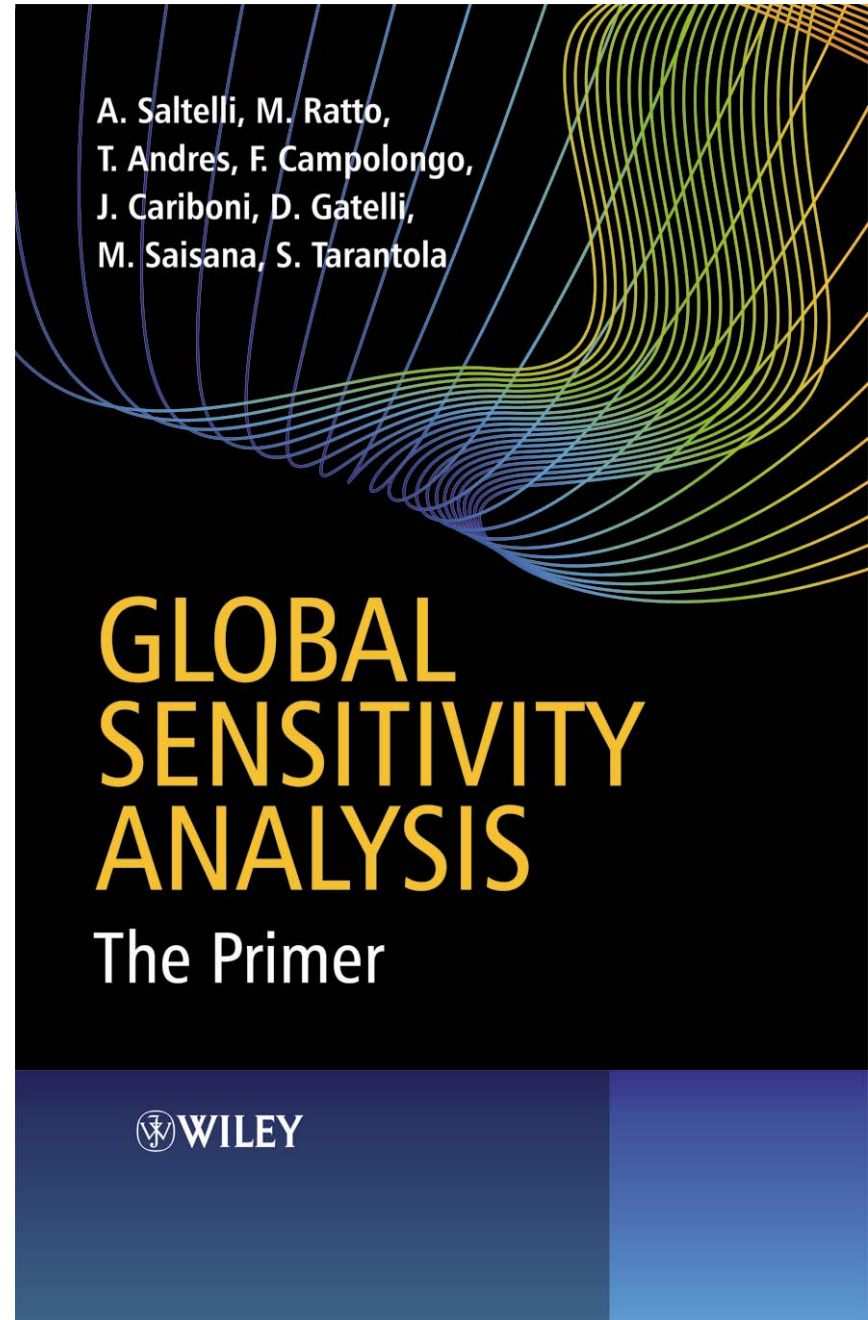




To know more buy:

Or read:

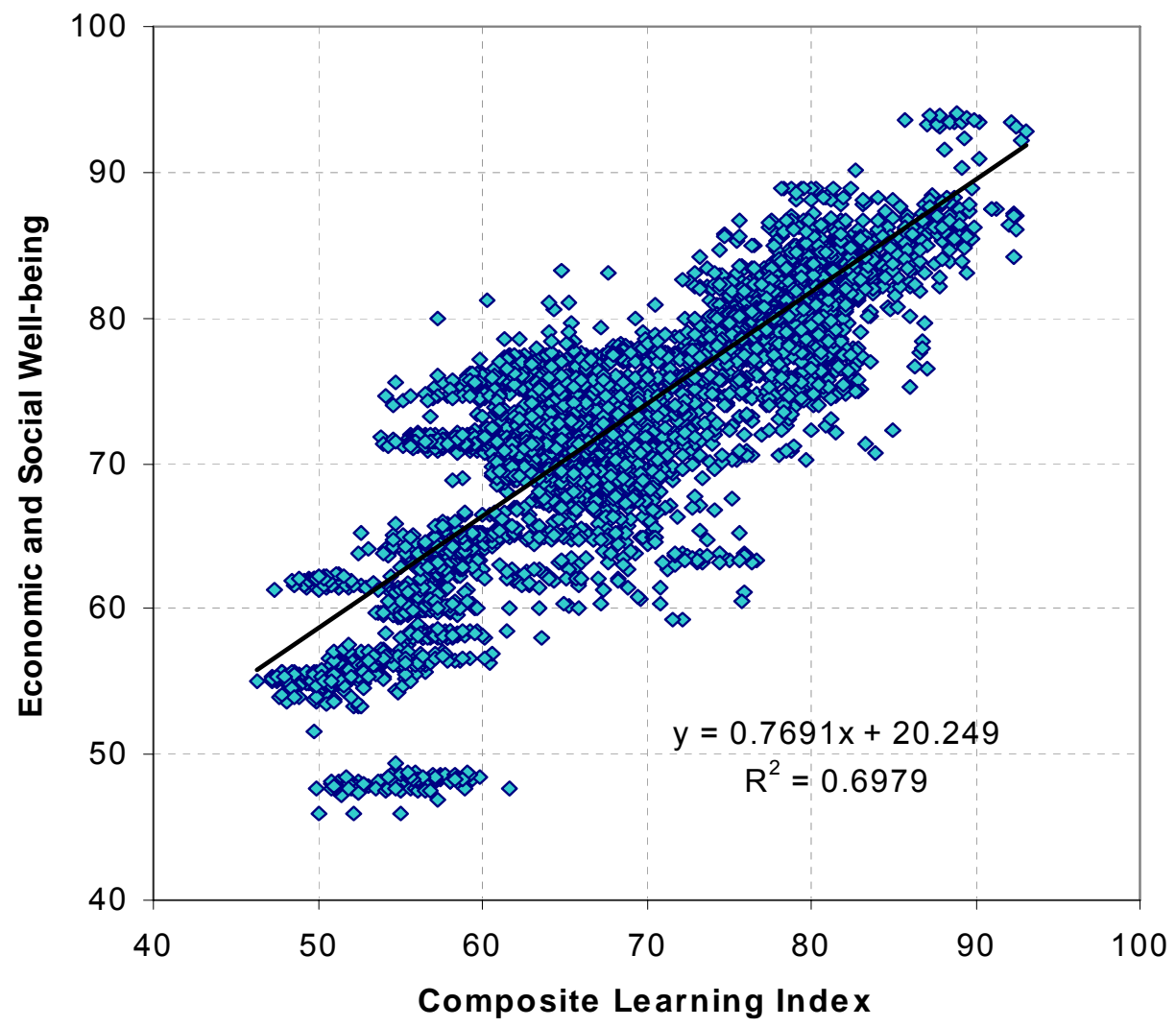
Saisana M., Saltelli A., Tarantola S. (2005) Uncertainty and Sensitivity analysis techniques as tools for the quality assessment of composite indicators, *Journal of the Royal Statistical Society - A*, **168**(2), 307-323.





## Step 8. Links to other variables

*Composite indicators can be linked to other variables and measures*



Comparing effectively complex dimensions: Canadian welfare and Canadian Composite Learning Index



# Improved Test Scores = Improved Growth Performance

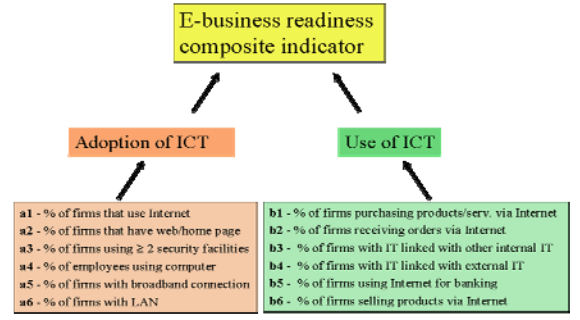
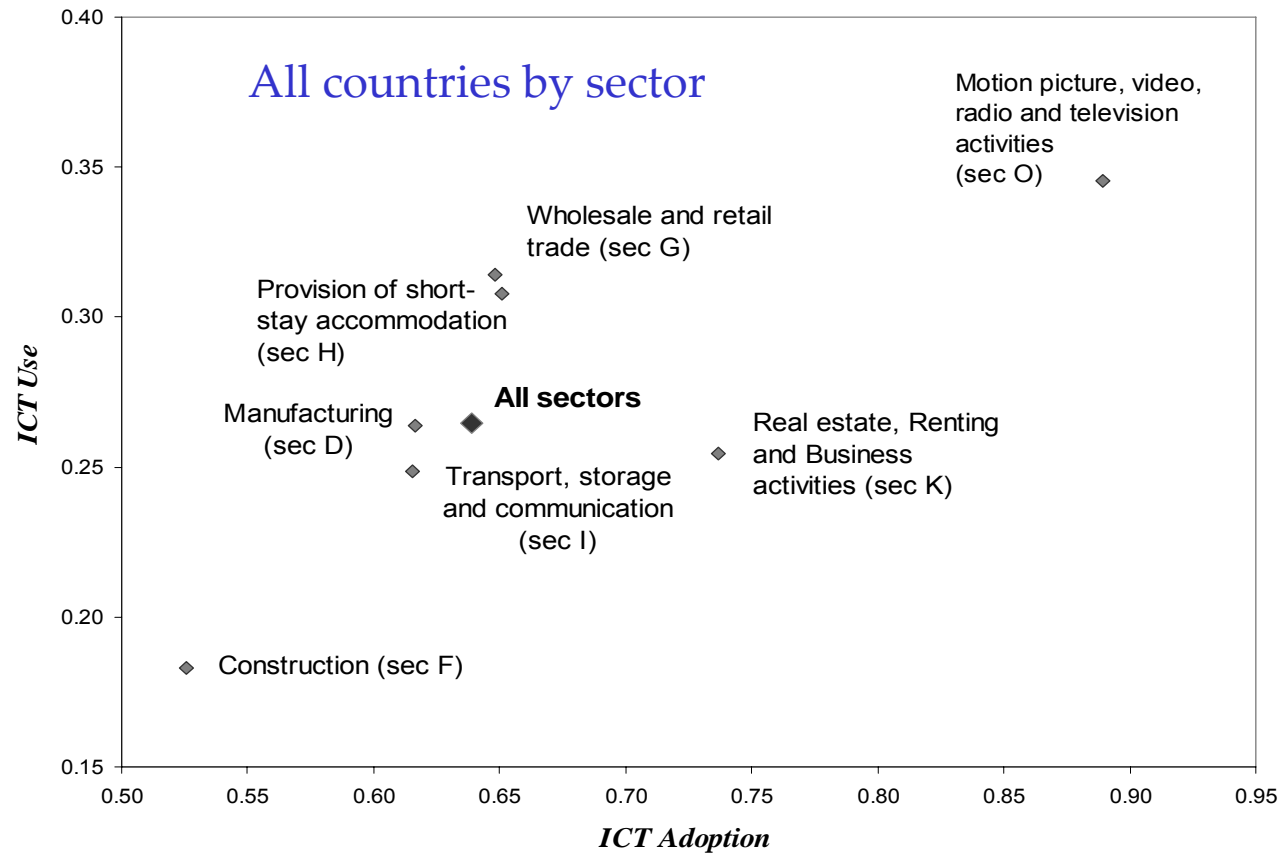
From **Ludger Wößmann**,  
 Contribution of Education and Training to Innovation and Growth, see:  
[www.education-economics.org](http://www.education-economics.org)





# Step 9. Back to the details

*De-constructing composite indicators can help extend the analysis ...*

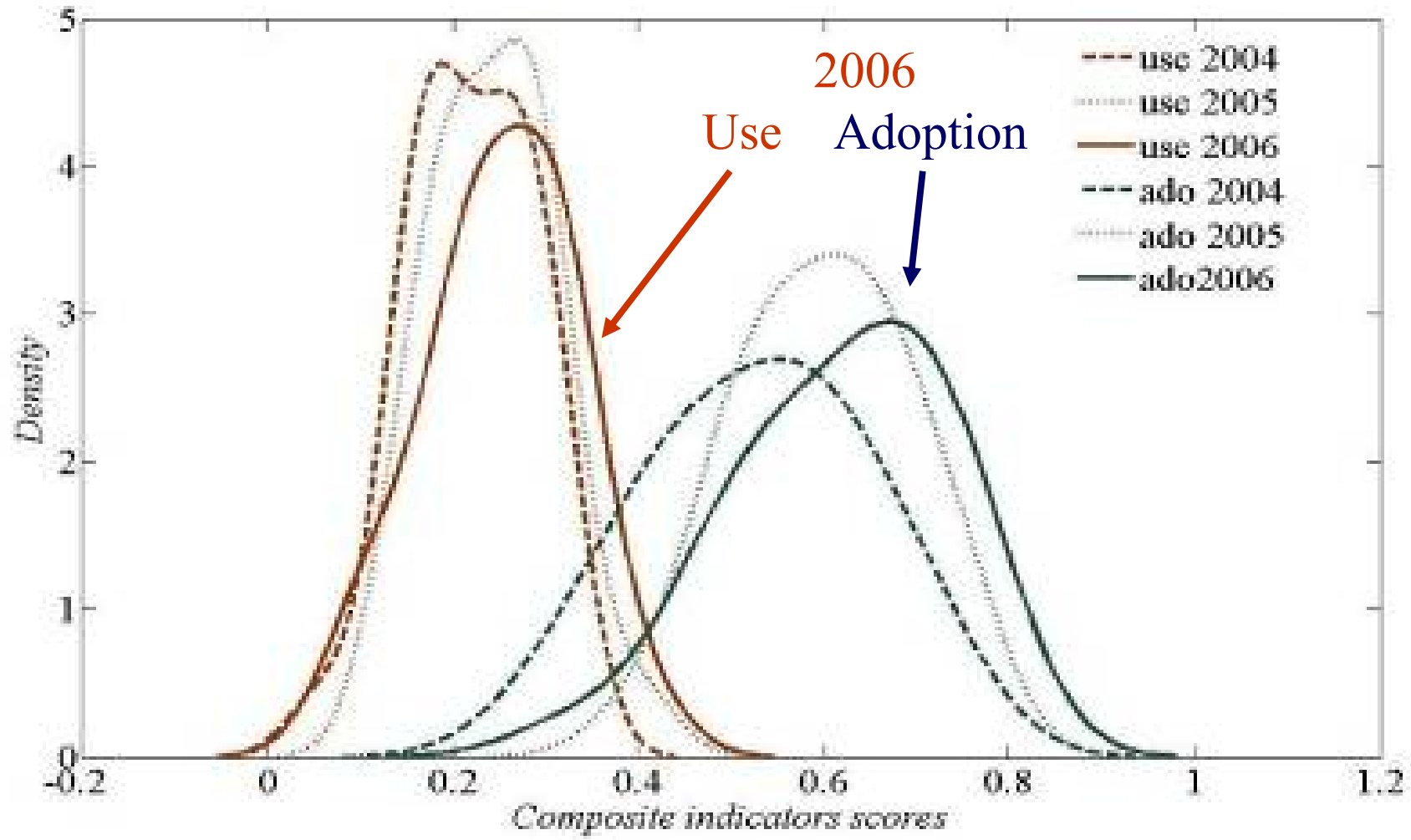




Enterprise survey 2007 (Eurostat),  
yearly since 2003, Adoption and  
Use composite indicators

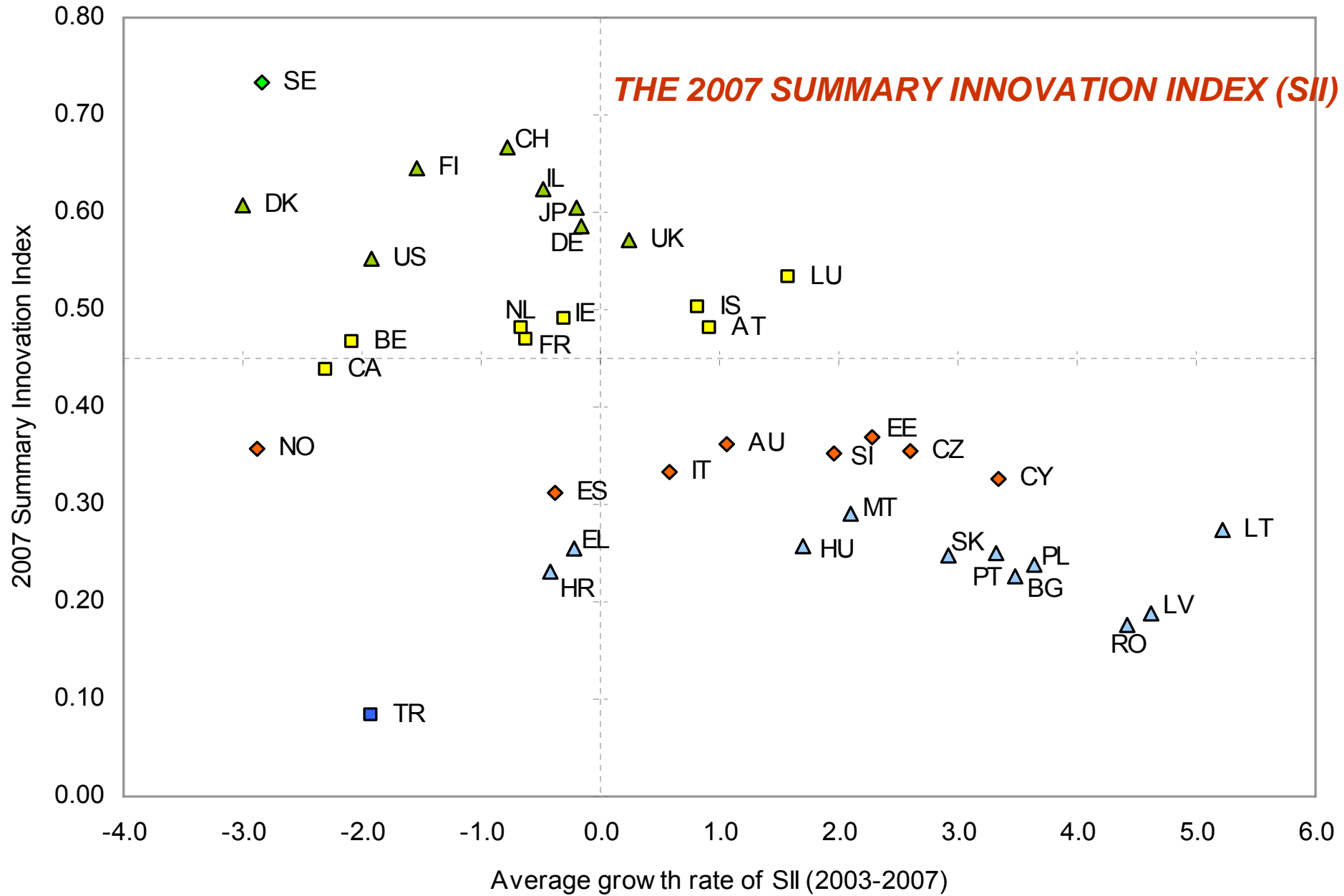
# Monitoring e-Business Readiness of European enterprises

Joint Research Centre





**THE 2007 SUMMARY INNOVATION INDEX (SII)**



- ◆ Sweden
- ▲ Innovation leaders
- Innovation followers
- ◆ Moderate innovators
- ▲ Catching-up countries
- Turkey

*Dotted lines show EU performance.*



# Step 10. Presentation and dissemination

Joint Research Centre

**CLI** Composite Learning Index

**THE 2007 COMPOSITE LEARNING INDEX**  
*Helping Communities Improve their Quality of Life*

*Learning to Know, to Do, to Live Together, to Be*

CANADIAN COUNCIL ON LEARNING **CCL** **CCA** CONSEIL CANADIEN SUR L'APPRENTISSAGE

**2008 Environmental Performance Index**

**Summary for Policymakers**

Yale Center for Environmental Law and Policy  
Yale University

Center for International Earth Science Information Network  
Columbia University

*In Collaboration with:*  
World Economic Forum  
Geneva, Switzerland

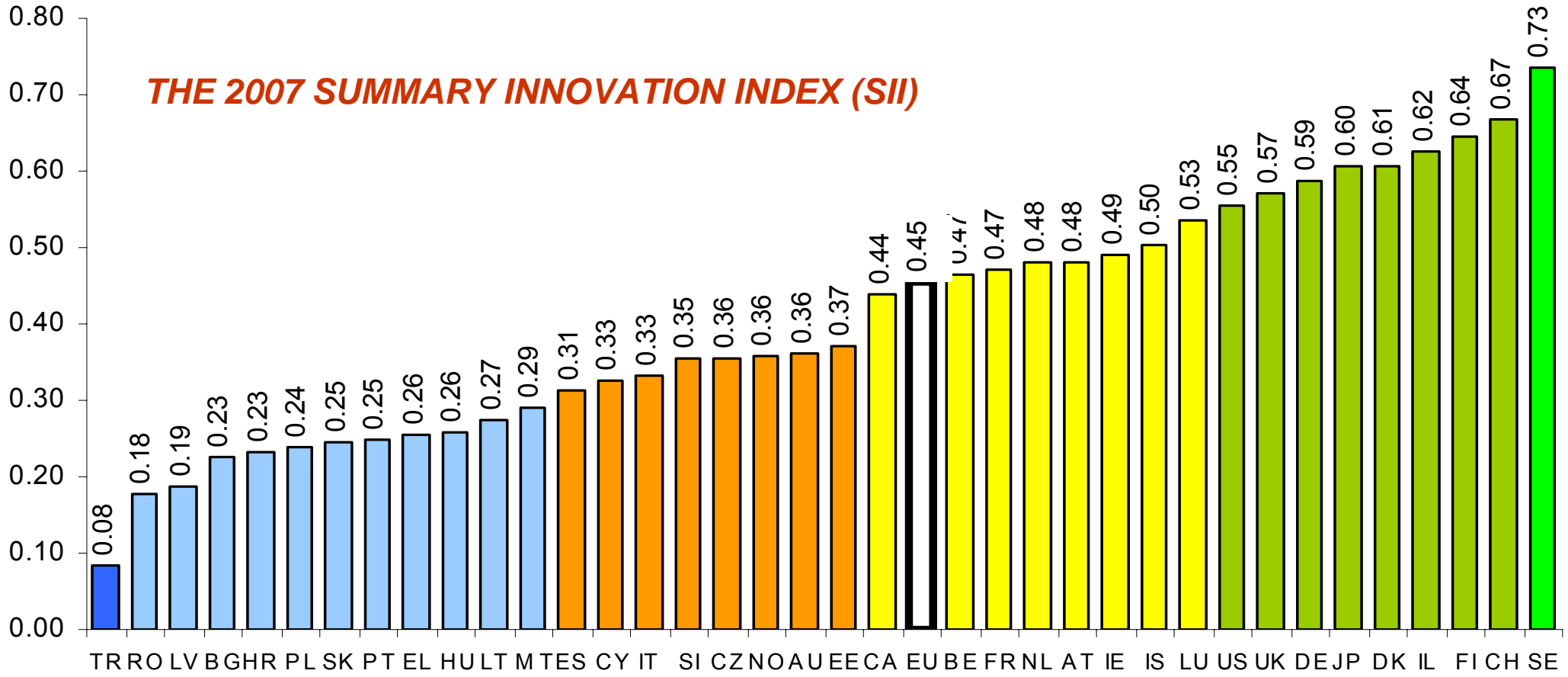
Joint Research Centre of the European Commission  
Ispra, Italy

Full report and additional materials available at:  
<http://epi.yale.edu>

## COMPARATIVE ANALYSIS OF INNOVATION PERFORMANCE OF THE EUROPEAN COUNTRIES

tre

**THE 2007 SUMMARY INNOVATION INDEX (SII)**





## Why the ten steps?

To combat **lack of transparency**, which makes them and the associated the policy inference, arbitrary and objectionable.

The three pillars of a well-designed CI are:

- Solid theoretical **framework**,
- Underlying **data** of good quality,
- Tested **methodology** .



Methodology should include assessment of the CI robustness.

CI can be refuted by practitioners  
(examples of WEF GCI and Ecological Footprint)



- No more methodology. The KEI study



Joint Research Centre



*Giuseppe Munda and Michaela Saisana*





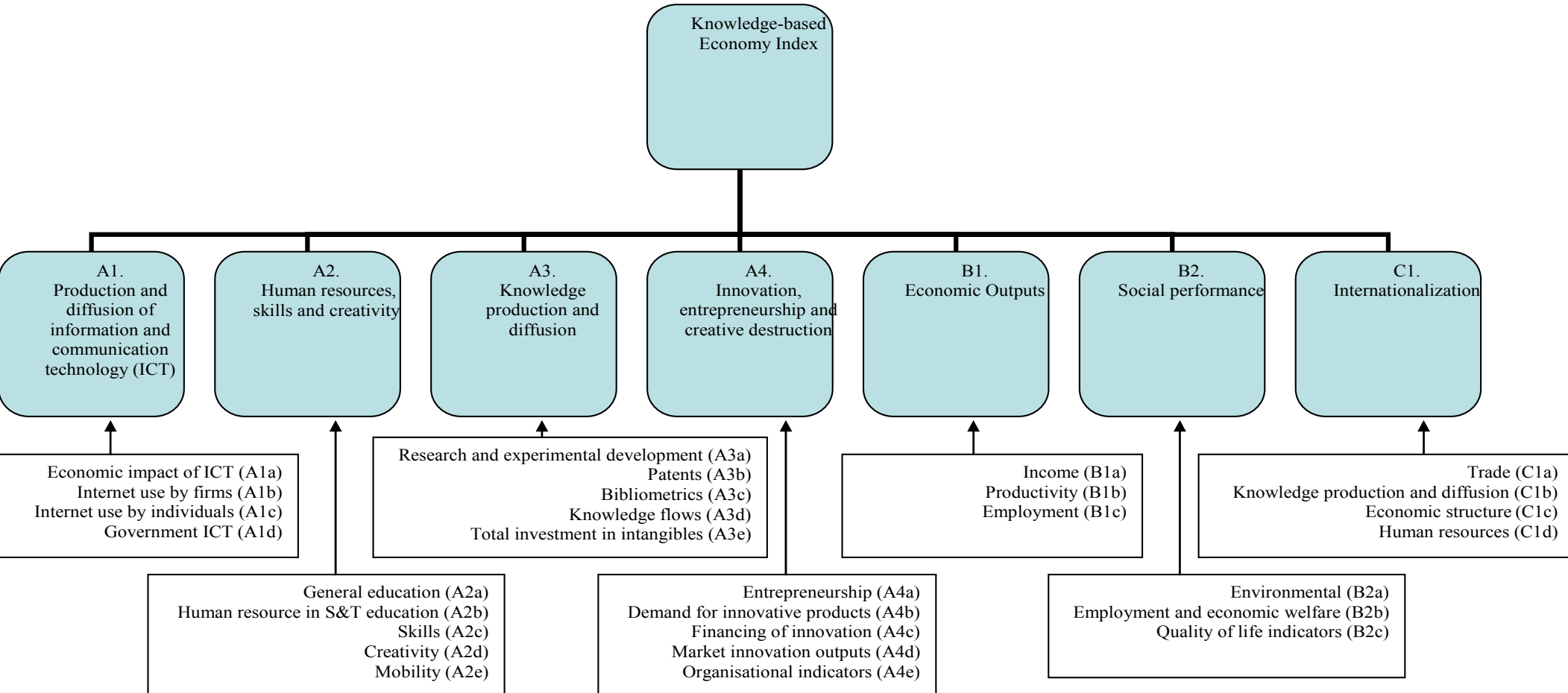
## Main research questions:

- Is it possible to measure the knowledge economy?
- What are the drivers of the knowledge economy?
- How does knowledge economy relate to other complex dimensions?
- Is it possible to reduce the total number of individual indicators of KEI conceptual framework without losing any relevant information?
- Are rankings useful at all for deriving policy suggestions?





**Figure 1. KEI Conceptual framework of a Knowledge-based Economy**





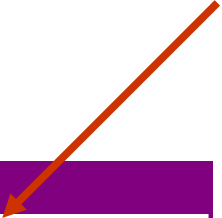
1. Production and diffusion of ICT
2. Human resources, skills and creativity
3. Knowledge production and diffusion
4. Innovation, entrepreneurship and creative destruction
5. Economic outputs
6. Social performance
7. Internationalisation



We do not trust one model to build the Knowledge Economy Index, thus we opt for a “multi-modelling principle”



## Methodological scenarios for the development of the KEI composite

		Aggregation function 		
		Additive (linear)	Multiplicative (geometric averaging)	Non-compensatory multi-criteria analysis
<b>Sub-dimensions included</b>	All (total 29)	•	•	•
	One-at-a-time excluded	•	•	•
<b>Dimensions included</b>	All (total 7)	•	•	•
	One-at-a-time excluded	•	•	•
<b>Pillar Structure</b>	Preserved	•	•	•
	Not preserved	•	•	•
<b>Normalisation</b>	z-scores	•	•	
	Min-max	•	•	
	Raw data			•
<b>Weighting</b>	Factor analysis	•	•	•
	Equal weighting	•	•	•
	Data envelopment analysis	•		

The **frequency matrix** of a country's rank in each of the seven dimensions and the overall KEI was calculated across the **~2,000 scenarios**.



# Knowledge Economy Index

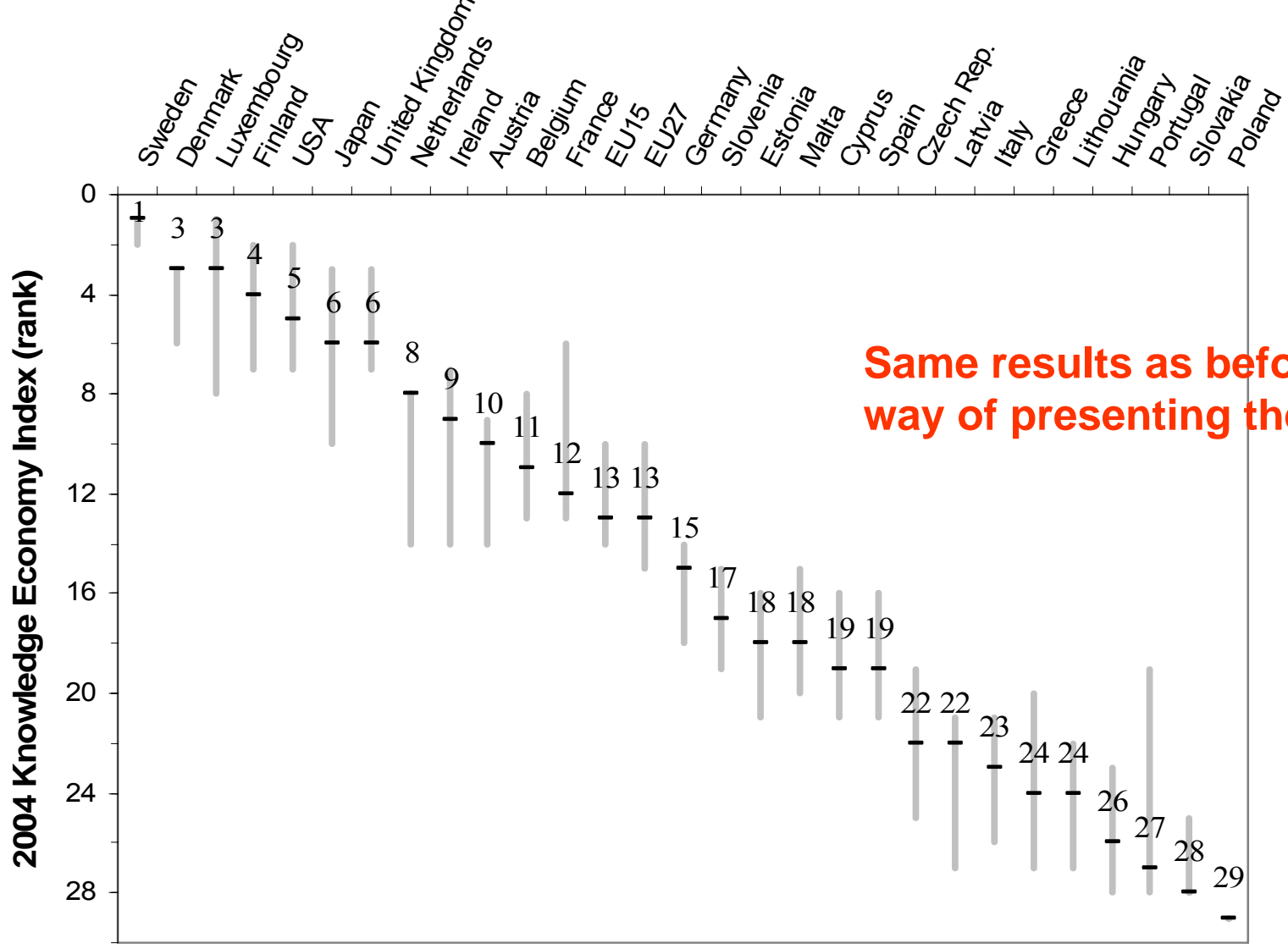
Joint Research Centre

	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	Rank 7	Rank 8	Rank 9	Rank 10	Rank 11	Rank 12	Rank 13	Rank 14	Rank 15	Rank 16	Rank 17	Rank 18	Rank 19	Rank 20	Rank 21	Rank 22	Rank 23	Rank 24	Rank 25	Rank 26	Rank 27	Rank 28	Rank 29	
Sweden	54	46																												
Denmark			55	30	14																									
Luxembourg	36	4	14	25	4	7	7	4																						
Finland		18	23	29	9	11	11																							
USA	11	32	2	4	39	9	4																							
Japan			4	7	18	32	36			4																				
UK			2	5	16	38	39																							
Netherlands								86	4		4			7																
Ireland							4		61	14	4	9	9																	
Austria									18	50	18	7	7																	
Belgium							11	4	11	57	16	2																		
France					4				14	18	11	54																		
EU15											4		57	39																
EU25										4	4	14	32	39	7															
Germany														7	79	4	7	4												
Slovenia														7	41	38	14													
Estonia															4	36	25	21	11	4										
Malta														7	13	9	21	23	27											
Cyprus															36	7	4	23	23	7										
Spain														4	4	32	25	29	7											
Czech. Rep.																			4	7	30	39	5	7	7					
Latvia																					20	36	11	21	7		5			
Italy																					29	18	9	29	9	7				
Greece																			4	4	4	4	29	18	21	7	14			
Lithuania																						4	41	13	32	11				
Hungary																							2	13	13	57	2	14		
Portugal																			4				4		7	11	61	14		
Slovakia																									4	7	18	71		
Poland																														100

*Legend:*  
 Frequency lower 15%  
 Frequency between 15 and 30%  
 Frequency between 30 and 50%  
 Frequency greater than 50%



## Median and associated 5<sup>th</sup> and 95<sup>th</sup> percentiles for the rank distribution



**Same results as before, different way of presenting them**



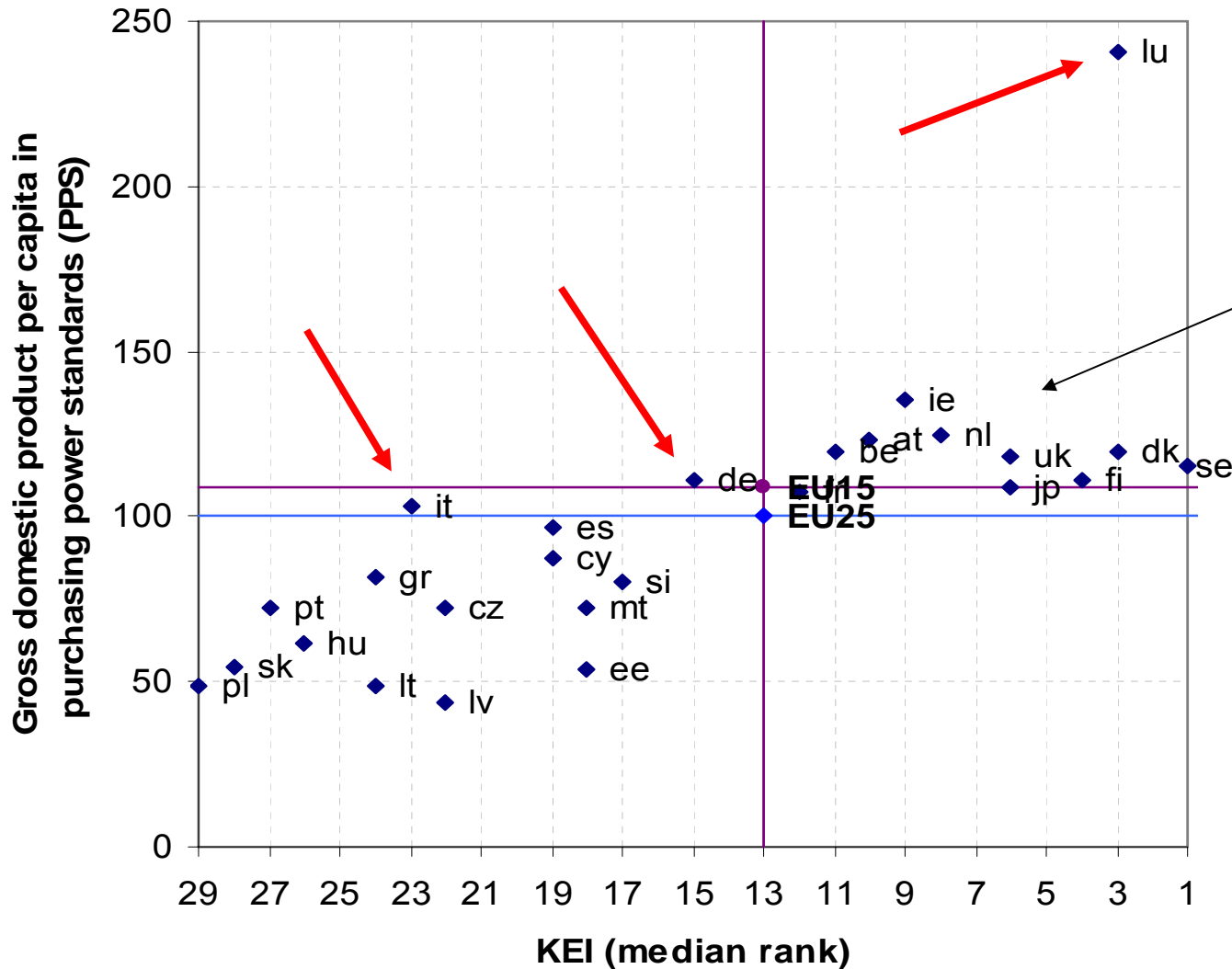
## Spearman rank correlation coefficients between KEI and its seven dimensions (median ranks across ~2,000 simulations)

All dimensions play a role but surely the less influential seems to be the 'innovation, entrepreneurship and creative destruction dimension'.

	Production and diffusion of ICT	Human resources, skills and creativity	Knowledge production and diffusion	Innovation, entrepreneurship and creative destruction	Economic outputs	Social performance	Internationalisation
<i>KEI</i>	0.90	0.95	0.91	0.40	0.68	0.78	0.63
Production and diffusion of ICT		0.87	0.88	0.42	0.54	0.58	0.41
Human resources, skills and creativity			0.90	0.44	0.60	0.68	0.53
Knowledge production and diffusion				0.37	0.42	0.69	0.55
Innovation, entrepreneurship and creative destruction					0.17	0.06	0.04
Economic outputs						0.59	0.48
Social performance							0.61
Internationalisation							1.00

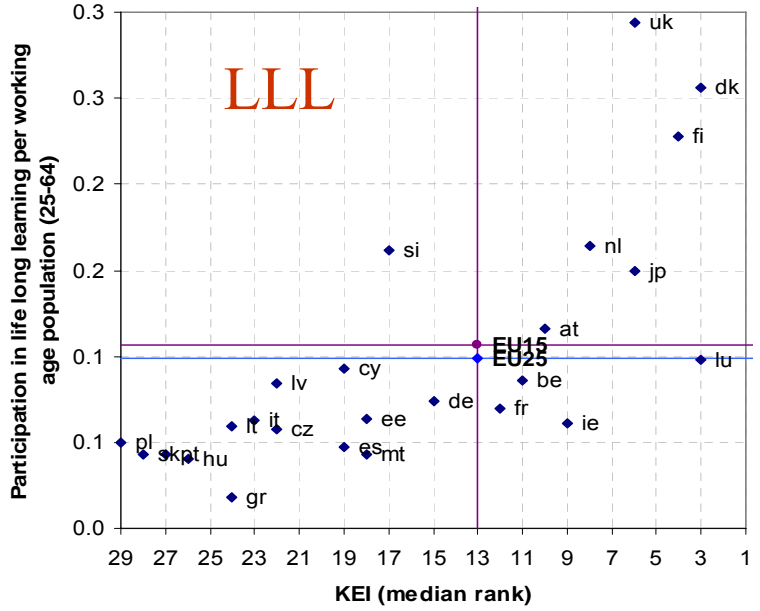
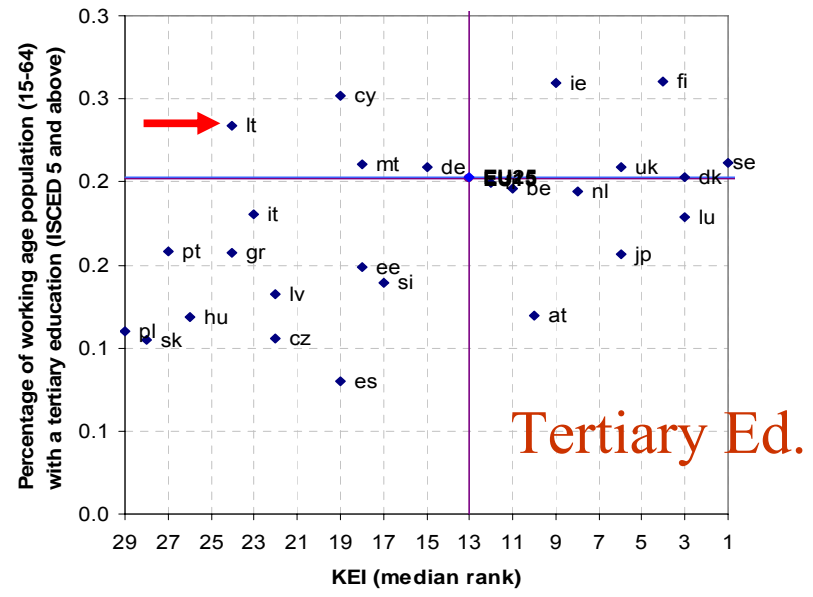
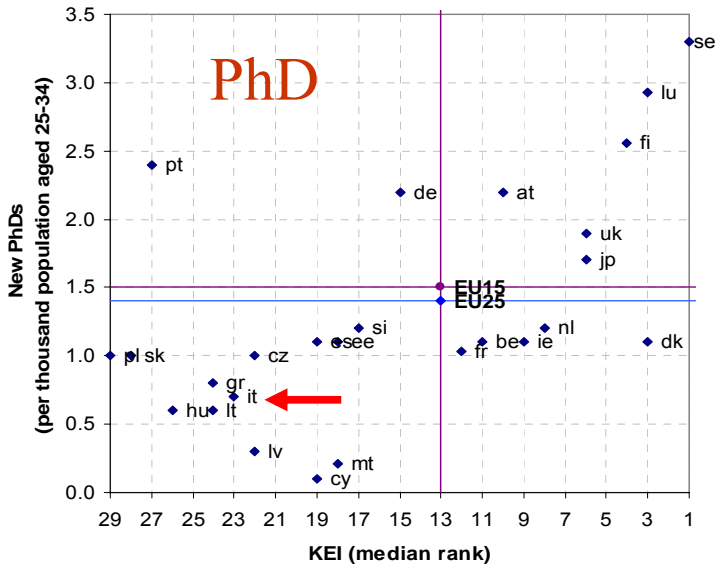
All coefficients are significant ( $p < 0.05, n = 29$ ).





**Is KE relevant at all for a good overall economic performance?**

The answer seems to be **“YES”**. Except Germany, Italy and Spain which have a high level of GDP per capita without any particular good performance on a KEI. All the other high level GDP countries seem to perform well in the KEI composite (Luxembourg probably an outlier)

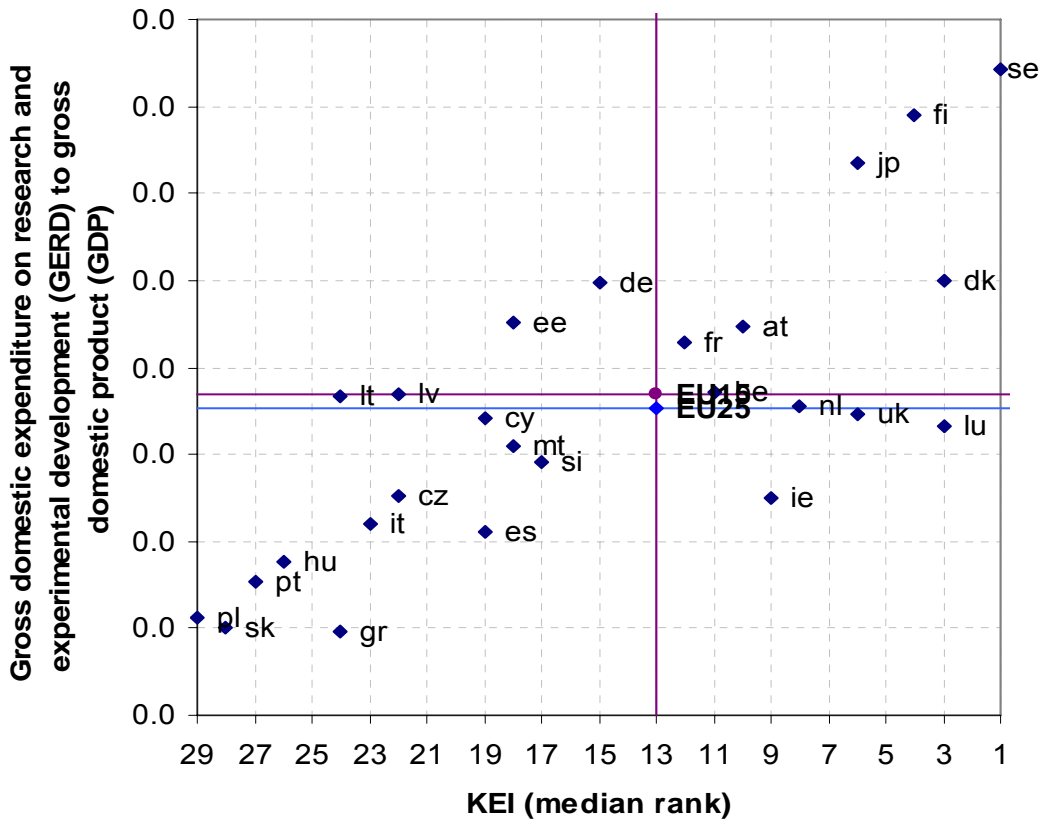


While the number of Ph.Ds seems to play a role in explaining the success of a KE (all countries with an high number of Ph.Ds, except Portugal, are top countries in the KEI median ranking), the same result does not apply to percentage of working population with a tertiary education (very clear the case of Italy where the number of working population with a tertiary education is extremely high, but the number of Ph.Ds is small). Participation to life long learning seems also to be a success factor, although not for all top countries.

**In sum, the human capital theory seems to be corroborated by the KEI results roughly.**

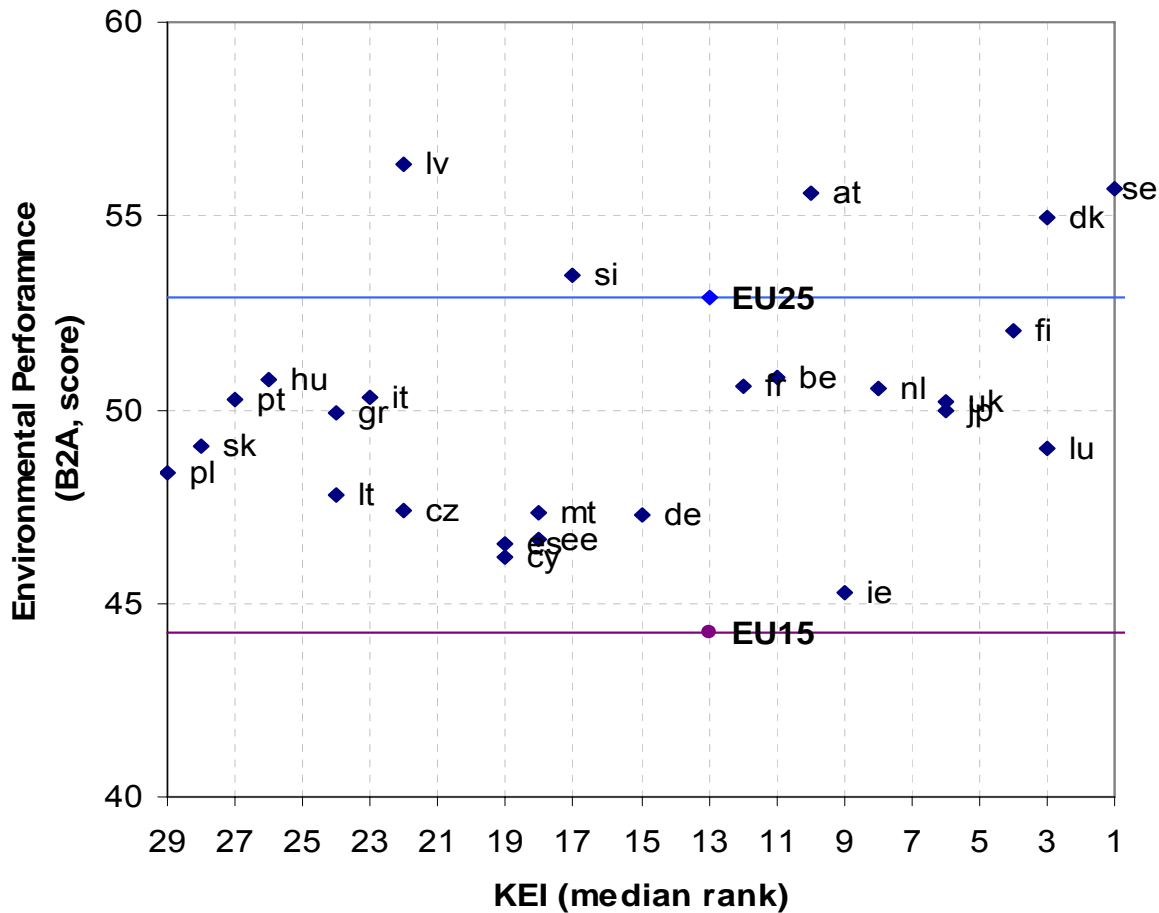
Checking the Schumpeterian theory:

‘Gross domestic expenditure on research and experimental development’ relates with KEI median.



**Countries which invest in research are top countries in a KBE.**

Research is a key driver for a KBE surely, thus the *endogenous growth idea* (the Schumpeterian theory) seems also corroborated.



Is a KE an eco-efficient one?

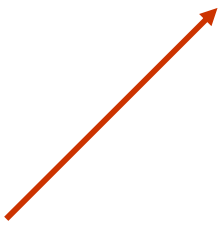
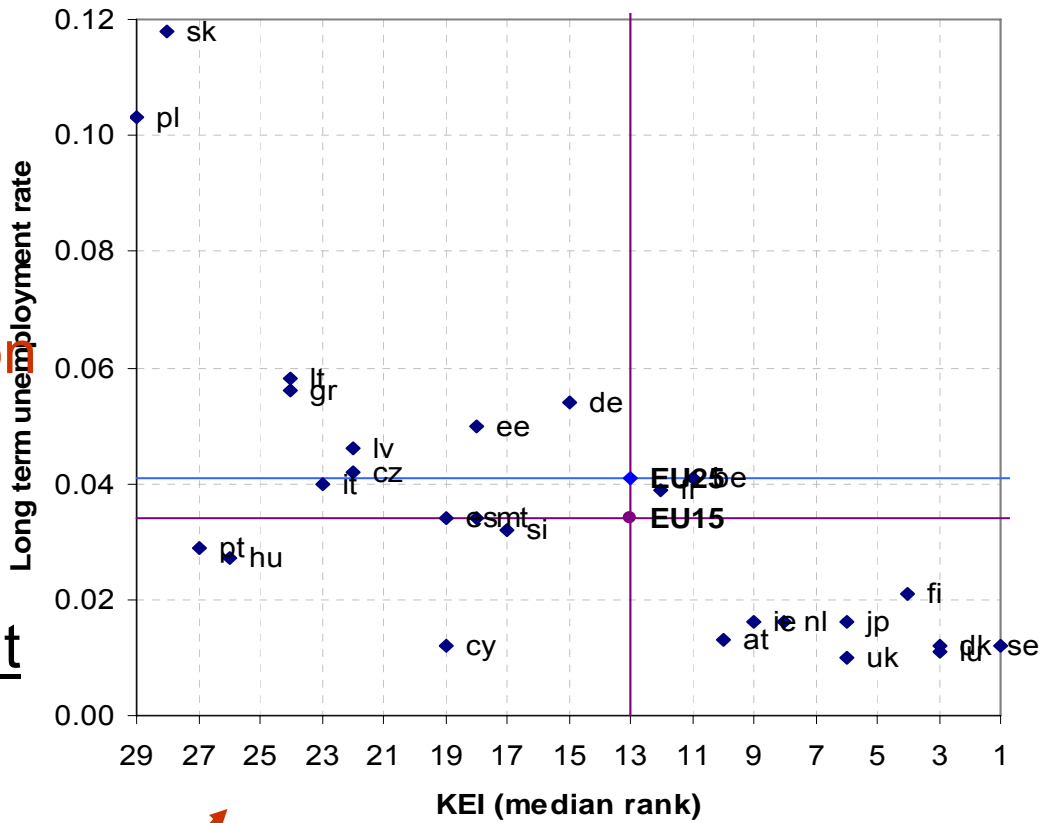
The *Jevons' paradox* teaches us that an ▲ in efficiency in using a resource leads, in the medium to long term, to an ▲ in consumption of that resource (rather than a ▼). No Kuznets curve ...

Join

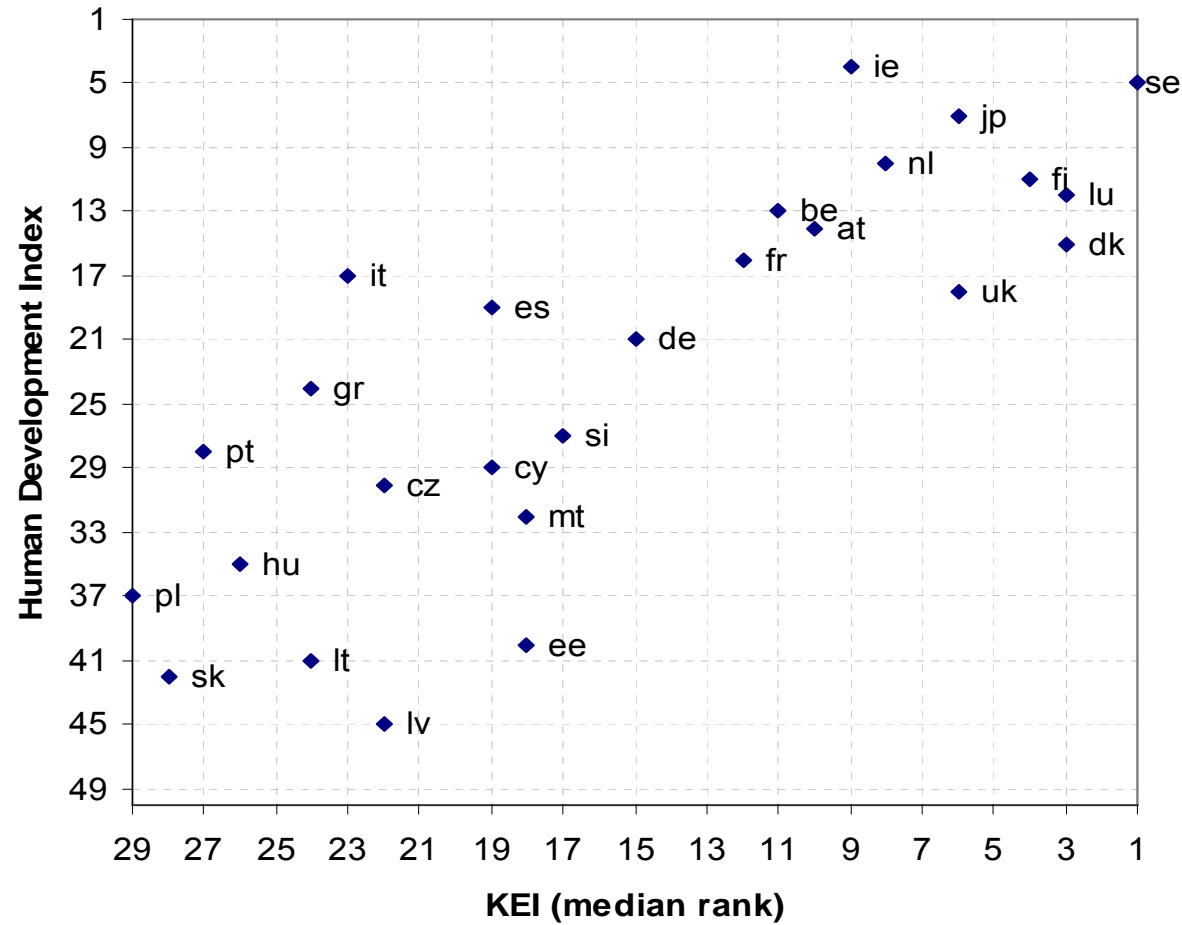
These arguments seem corroborated by the KEI measure. In fact, as one can see, no clear relationship between environmental performance and a KBE exists.

## Does a KE influence unemployment?

In the short term technological progress and job creation are conflicting objectives (**Kok report on Europe**) but they might be compatible in the long period if a right balance (i.e. compromise) between flexibility and employment security is found.



**Compatibility between technological progress and job creation. All top countries in the KEI measure are presenting an extremely low long term unemployment rate.**



External benchmark:  
Human Development Index.

**The relationship found seems again a corroboration of the human capital theories, in fact the correlation between KEI and the HDI is extremely high (though possibly spurious!).**



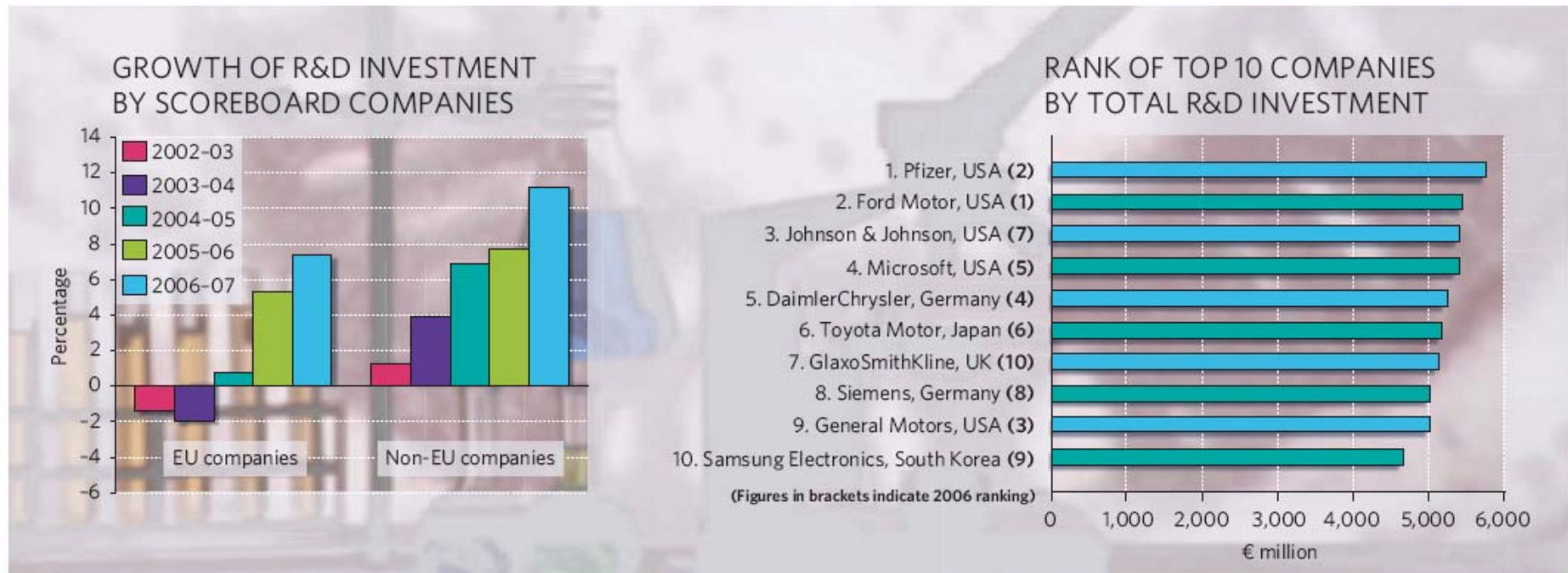
- JRC-IPTS (Seville) work on companies R&D expenditures scoreboard.



## BUSINESS

# Mixed scores for European research

A survey shows positive trends for private-sector research and development in the European Union. But as **Andrea Chipman** reports, there's more to the data than meets the eye.





## Industrial Research and Innovation at JRC - Seville

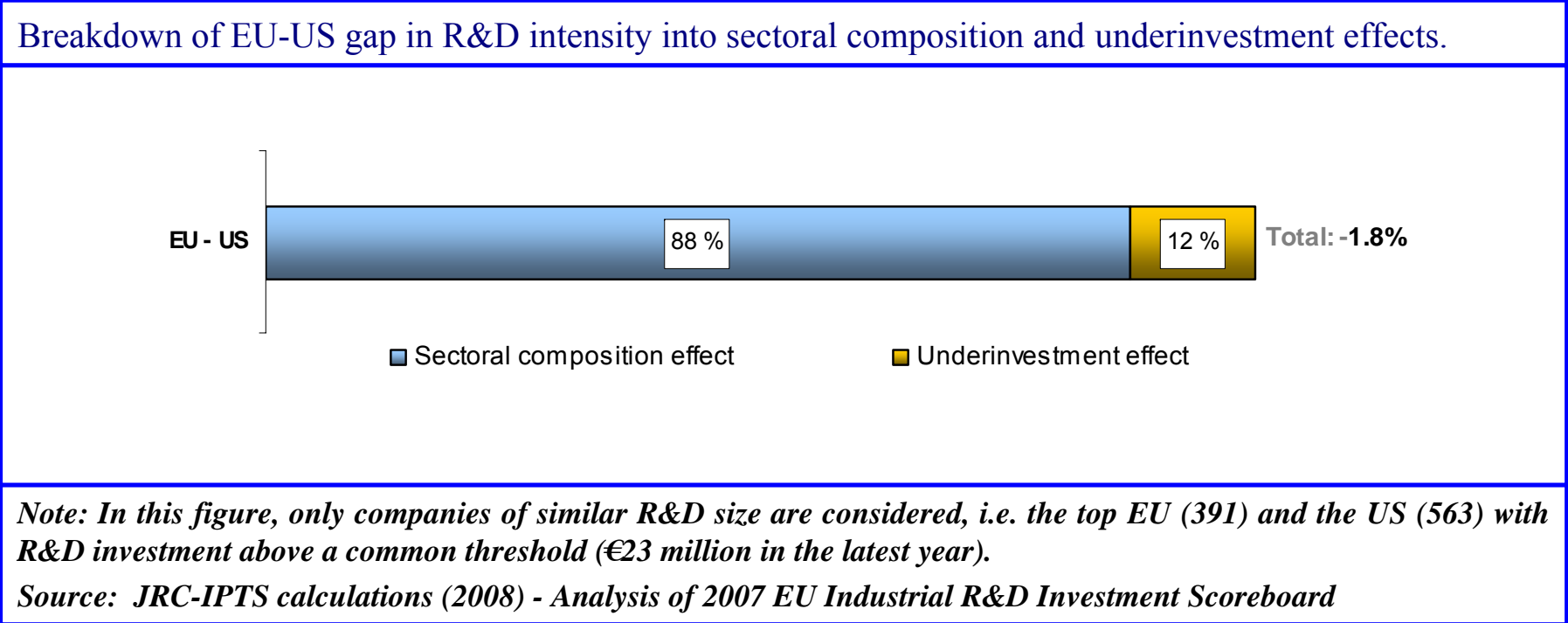
<http://iri.jrc.ec.europa.eu/>

- The EU Industrial R&D Investment Scoreboard: analysis of 1000 EU and 1000 non-EU top investing companies in R&D
- The EU Survey of Business R&D
- Economic and policy analysis of corporate R&D.



## Some Results (1): Nature of the R&D investment gap

EU's R&D intensity deficit is largely explained by the different industrial structure (sectoral composition effect).

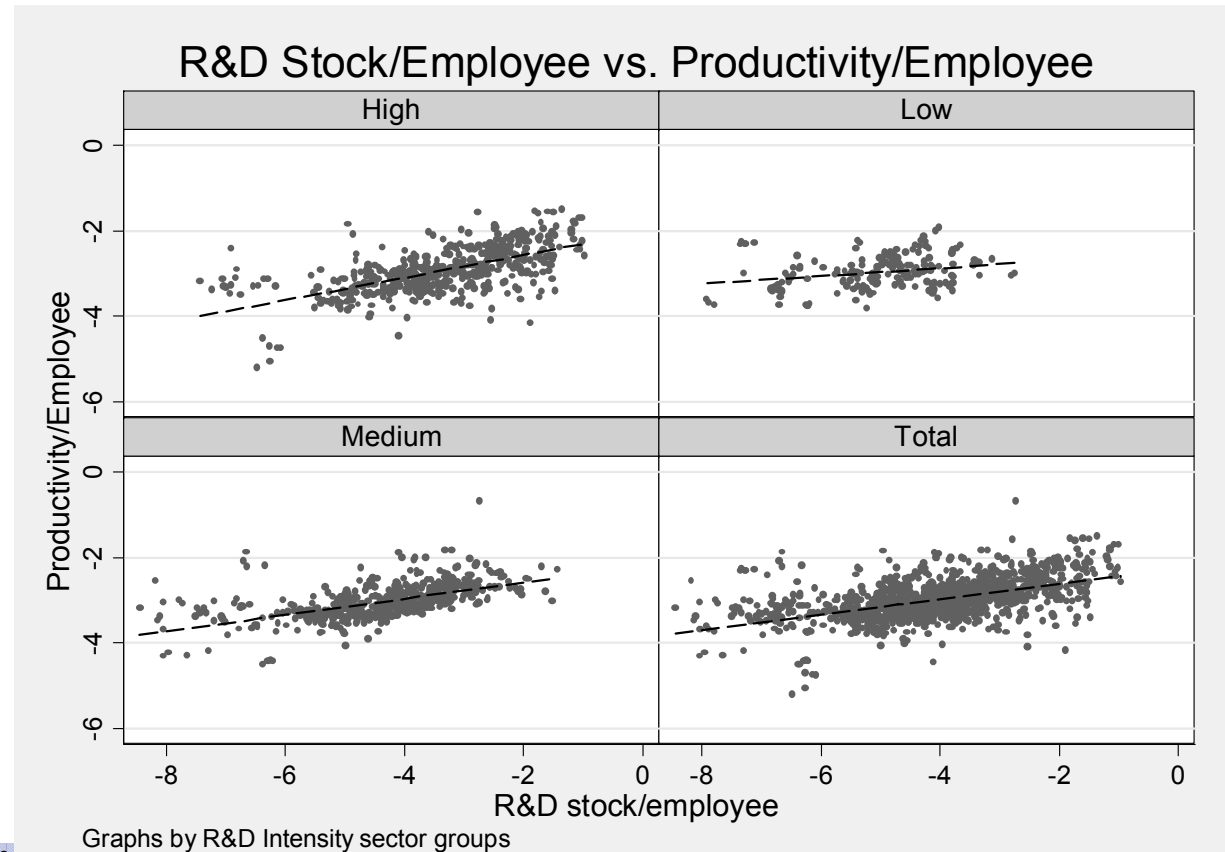


# Industrial Research and Innovation at JRC Seville

Joint Research Centre

## Some Results (2): Econometrics of R&D & firm productivity

- → R&D rises productivity much more in high-tech sectors than in low-tech ones





More info on the R&D investment  
scoreboard from

**Pietro.MONCADA-PATERNO-  
CASTELLO@ec.europa.eu**



More reading at (just google composite indicators):

<http://composite-indicators.jrc.ec.europa.eu>

