State of the art on Composite Indicators

Stefano Tarantola

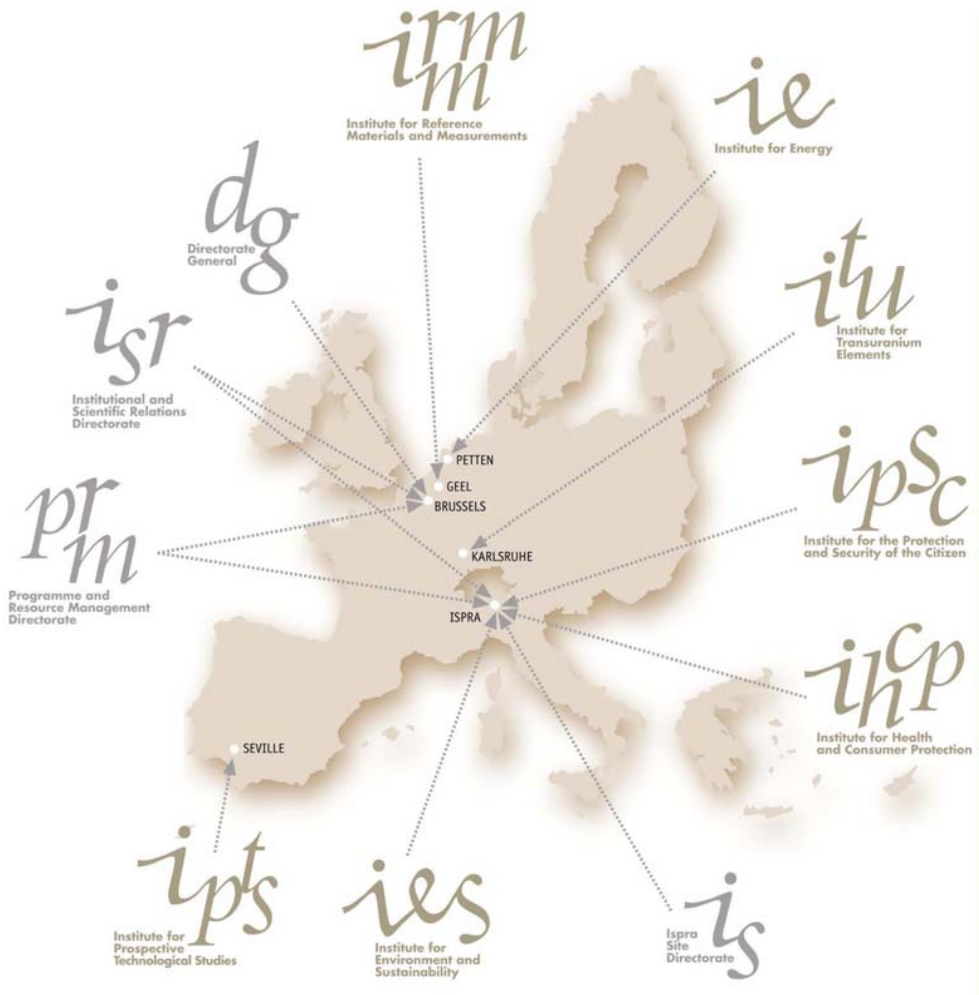
European Commission
Joint Research Centre of Ispra

Tuebingen
March 4th, 2005

As a Directorate General (DG) of the European Commission, the JRC provides scientific and technical support to Community policy making.

7 Institutes in 5 Member States
Total staff: ~ 2200 people.

The Unit Applied Statistics and Econometrics is in Ispra



Prepared with Michela Nardo,

Michaela Saisana

Andrea Saltelli,

and
Niels Schulze

Based on:

Saisana M., Saltelli A., Tarantola S., 2005, Uncertainty and Sensitivity analysis techniques as tools for the quality assessment of composite indicators, *J. R. Stat. Soc. A*, **168**(2), 1-17.

Joint OECD-JRC handbook on good practices in composite indicators building.

Outline of the presentation

- Debate on composite indicators
- Pros and cons
- Correlation and compensability issues
- Composite indicators as models and the critique of models (Rosen)
- Robustness analysis (Edward Leamer)
- Sensitivity analysis

- Composite indicators on the knowledge-based economy

- CI controversy

EU structural indicators – scoreboards versus indices

Database of structural indicators (117)

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
Joint Research Centre

Report from the Commission to the Spring European Council 2004, Annex 1

Joint Research Centre

Relative Performance

Relative performance of the 15 Member States according to the structural indicators on the shortlist

Levels		at	be	de	dk	es	fi	fr	gr	ie	it	lu	nl	pt	se	uk	eu15	us
GDP per capita in PPS (EU 15 = 100)	2003	111,9	107,6	102,2	114,5	85,5	102,3	101,4	68,3	121,8	102,3	138,5	110,4	67	101,8	105,1	100	138,7
Labour productivity (EU 15 = 100)	2003	97,9	114	97,9	99,4	91,2	101,6	109,1	86	116,9	114,7	185,8	95,5	65	94,9	94,7	100	120,1
Employment rate (%)	2002	69,3	59,9	65,3	75,9	58,4	68,1	63	56,7	65,3	55,5	63,7	74,4	68,2	73,6	71,7	64,3	74,6
Employment rate of older workers (%)	2002	30	26,7	38,4	57,8	39,7	47,8	34,8	39,7	48,1	28,9	28,3	42,3	50,9	68	53,5	40,1	:
Educational attainment (20-24) (%)	2002	85	81,1	73,3	83,9	64,9	86,2	81,7	81,3	83,9	69,1	69,8	73,3	43,7	86,7	91	75,6	:
Research and development expenditure (% of GDP)	2002	1,9	2,2	2,5	2,4	1,0	3,5	2,2	0,7	1,2	1,1	1,7	1,9	0,8	4,3	1,8	2,0	2,8
Business investment (% GDP)	2002	20,9	18,3	16,9	17,8	21,8	16	16,3	20,1	17,7	17,8	17,9	17,4	21,6	13,8	15	17,2	:
Relative price levels (EU 15 =100) [1]	2001	98	98	102	126	82	117	99	81	113	92	99	99	74	122	115	100	111
At-risk-of-poverty rate (%)	2000	12,0	13,0	11,0	11,0	18,0	11,0	16,0	20,0	20,0	18,0	12,0	10,0	21,0	11,0	19,0	15,0	:
Long-term unemployment (%)	2002	0,8	3,6	4	0,9	3,9	2,3	2,7	5,1	1,3	5,3	0,8	0,7	1,8	1	1,1	3	0,3
Dispersion of regional employment rates	2002	2,4	8	5,9	:	9,2	7,8	6,2	4,2	:	16,6	:	2,2	3,9	4,6	6,6	12,6	:
Greenhouse gases emissions (Index base year=100) [2]	2000	103	106	81	99	135	96	98	124	124	104	55	103	130	98	87	96	111
Energy intensity of the economy	2001	146	228	169	125	228	263	189	261	168	188	190	201	238	229	227	195	330
Volume of transport	2001	128	70	105	86	118	85	112	90	126	102	129	95	137	87	111	106	95

1. Analysis takes into account relation between GDP per capita and relative price levels. 2. Analysis based on distance to Kyoto targets.

Relative Improvement in Performance (av. since 1999)

Relative improvement in the performance of the 15 Member States according to the structural indicators on the shortlist

Evolution		at	be	de	dk	es	fi	fr	gr	ie	it	lu	nl	pt	se	uk	eu15	us
GDP per capita	Average annual real growth rate 1999-2003 (%)	1,4	1,1	0,8	1,5	2,1	2,3	1,3	3,9	4,8	1,2	2,1	0,3	0,6	1,9	2,3	1,4	1,2
Labour productivity	Average annual real growth rate 1999-2003 (%)	1,2	0,8	1,0	1,9	0,5	1,5	0,5	3,8	3,6	0,1	-0,8	0,4	0,6	1,1	1,8	0,9	1,8
Employment rate	Average annual percentage point change 1999-2002	0,2	0,2	0,2	0,0	1,6	0,5	0,7	0,5	0,7	0,9	0,7	1,0	0,3	0,6	0,2	0,6	-0,3
Employment rate of older workers	Average annual percentage point change 1999-2002	0,0	0,7	0,2	1,1	1,5	2,8	2,0	0,2	1,5	0,4	0,6	2,1	0,0	1,4	1,3	1,0	:
Educational attainment (20-24)	Average annual percentage point change 1999-2002	0,1	1,6	-0,4	-3,1	0,0	-0,2	0,6	0,6	0,5	0,9	-0,5	0,3	1,2	0,1	-0,2	0,3	:
Research and development expenditure	Average annual percentage point change 1999-2002	0,03	0,11	0,03	0,15	0,04	0,09	0,01	:	-0,02	0,03	:	-0,08	0,05	0,31	0,00	0,02	0,05
Business investment	Average annual percentage point change 1999-2002	-0,30	-0,27	-0,90	-0,10	0,33	-0,27	0,03	0,30	-1,00	0,40	-0,50	-0,70	-0,50	-0,07	-0,33	-0,23	:
Relative price levels [1]	Average annual percentage point change 1999-2001	-1,5	-1,8	-1,1	1,9	-0,3	-1,7	-3,1	-0,3	4,9	2,8	0,7	1,0	0,4	-1,7	1,7	0	:
At-risk-of-poverty rate	Average annual percentage point change 1999-2000	0,0	0,0	0,0	0,0	-1,0	0,0	1,0	-1,0	1,0	0,0	-1,0	-1,0	0,0	2,0	0,0	0	:
Long-term unemployment	Average annual percentage point change 1999-2002	0,0	-0,4	-0,1	-0,1	-0,5	-0,2	-0,6	-0,5	-0,4	-0,5	0,0	-0,2	0,1	-0,3	-0,2	-0,3	0
Dispersion of regional employment rates	Average annual percentage point change 1999-2002	0,0	0,0	0,1	:	-0,5	0,1	-0,3	-0,3	:	-0,2	:	0,0	0,1	-0,1	-0,2	-0,5	:
Greenhouse gases emissions	Average annual percentage point change 1999-2000	0,0	0,0	0,0	-6,0	6,0	-3,0	-1,0	6,0	2,0	1,0	0,0	-1,0	-1,0	-2,0	0,0	0,0	2,0
Energy intensity of the economy	Average annual percentage point change 1999-2001	1,4	-7,9	-0,4	-3,6	0,4	-6,3	-1,3	-0,4	-10,0	-2,9	-0,7	-0,4	-4,7	-4,5	-4,0	-1,8	-4,1
Volume of transport	Average annual percentage point change 1999-2001	4,2	-11,0	-1,0	-3,7	7,1	-2,6	-2,9	14,0	1,8	1,5	10,8	-4,2	-16,2	-4,3	-5,3	-2,5	:

1. Analysis takes into account relation between real growth rate of GDP per capita and changes in relative price levels.

Assessing policies: Green – Country policy on a good path; Yellow – Country policy on a bad path (expert judgment)

Levels	y	AT	BE
Labour productivity (EU 15=100)	2003	97.9	114
Employment rate (%)	2003	69.3	59.9
Employment rate of older workers (%)	2003	30	26.7

Enter the FT analysts …

Source: Spring Report, European Commission 2004

Source: **Financial Times** Thursday January 22

Categorization (star rating[*]) in three groups

LEADERS
UK, NL SE, DK, AT,LU

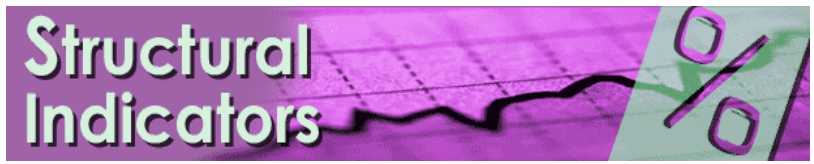
MIDDLE OF THE ROAD
DE, FI, IE, BE, FR

LAGGARDS
IT, GR, ES, PT

done by FT and based likely on same synoptic performance and improvement tables in the Spring Report, 2004, Annex 1 (yellow-green boxes)

[*] Like in the UK NHS hospital rating

Can “league tables” be avoided? Or are they an ingredient of an overall analysis and presentational strategy:

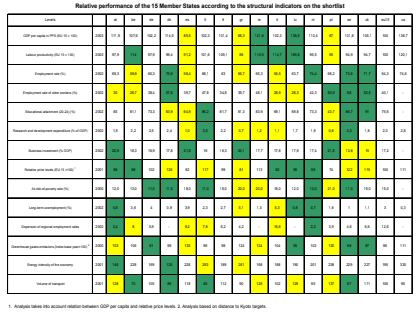


Long list of 117

Short List of 14

Synoptic tables

League tables



<<The aggregators believe there are two major reasons that there is value in combining indicators in some manner to produce a bottom line. They believe that such a summary statistic can indeed capture reality and is meaningful, and that stressing the bottom line is extremely useful in garnering media interest and hence the attention of policy makers.

The second school, the non-aggregators, believe one should stop once an appropriate set of indicators has been created and not go the further step of producing a composite index. Their key objection to aggregation is what they see as the arbitrary nature of the weighting process by which the variables are combined.>>

"Literature Review of Frameworks for Macro-indicators", **Andrew Sharpe, 2004**, Centre for the Study of Living Standards, Ottawa, CAN.

Grupp and Mogee (2004) Indicators for National Science and Technology Policy. How Robust are Composite Indicators? Research Policy 33, Nr. 9, S. 1373-1384

"To investigate the robustness of innovation scoreboards empirically, a sensitivity analysis of one selected case is presented. Composite scores and country rank positions can vary considerably depending on the selection process. The use of scoreboards leaves room for manipulation in the policymaking system. Further research is needed on alternative methods of calculation to prevent their misuse and abuse."

Pros

• Composite indicators can be used to **summarise complex or multi-dimensional issues,** in view of supporting decision-makers (sustainability, competitiveness, welfare).

• Composite indicators provide the **big picture** […]. They facilitate the task of ranking countries on complex issues.

• Composite indicators can help **attracting public interest** […]

• Composite indicators **add a layer** of information to the underlying list of indicators […].

Cons

• Composite indicators **may send misleading, non-robust policy messages** if they are poorly constructed or misinterpreted [… or] may invite politicians to draw **simplistic policy conclusions** […]

• The construction of composite indicators involves stages where **judgement** has to be made: the selection of sub-indicators, choice of model, weighting indicators and treatment of missing values etc. […]

• There could be **more scope for disagreement among Member States** about composite indicators than on individual indicators […].

Pros & Cons
(JRSS paper)

"[…] it is hard to imagine that debate on the use of composite indicators will ever be settled […] official statisticians may tend to resent composite indicators, whereby a lot of work in data collection and editing is "wasted" or "hidden" behind a single number of dubious significance.

On the other hand, the temptation of stakeholders and practitioners to summarise complex and sometime elusive processes (e.g. sustainability, single market policy, etc.) into a single figure to benchmark country performance for policy consumption seems likewise irresistible."

Open issues in CI Building

1 – Variables correlation

(1) A composite constructed on the basis of underlying indicators with high internal correlation is likely to give a very robust CI, whose values and ranking are moderately affected by changes in the selection of weights, the normalisation method and other steps involved in the analysis.

Open issues in CI Building

2 – Variables correlation

(2) When building composite indicators using automated tools such as factor analysis, one seeks to obtain a set of totally uncorrelated new variables. While this can be a powerful tool to benchmark countries performance, or to produce e.g. leading or lagging synthetic indicators, the interpretation in terms of original variables becomes more difficult.

Open issues in CI Building

3, 4 – Variables correlation

(3) At the same time, it would be very difficult to imagine a composite indicator made of truly orthogonal variables.

(4) We would consider the existence of correlation among the attributes of an issue as a feature of the issue, not to be corrected for through re-weighting of components.

However, if two attributes are actually redundant, there might be reasons for using one attribute to characterise the issue.

Example: car beauty and power

Open issues in CI Building

3, 4 – Variables correlation

Example: In European Innovation Scoreboard 2005 USPTO, EPO are kept both in the summary innovation index, though they are correlated at 0.97

-legal differences in the two systems: US data cover grants, EPO data count applications. Granted patents are a subset of all patent applications and do not reflect the total innovative capacity.

Rainer Frietsch
Fraunhofer ISI, Germany
Comments on the European Innovation Scoreboard 2005, February 2005

Open issues in CI Building

3, 4 – Variables correlation

The year of the grants has nothing to do with the point in time when the invention took place: it reflects the processing capacity of the patent office. It is only possible to draw a picture of the situation five to six years before the actual point in time, which is only seldom comparable to the actual situation.

USPTO is a national patent office. A country has a certain home advantage at the national office; US applicants at USPTO are overrepresented in relation to applicants from other countries.

this home advantage does not find its equivalent for European applicants at the EPO, as there are still many national patent offices all over Europe that receive a large number of important and innovative patent applications.

So the inclusion of EPO and USPTO patent data may explain the still increasing gap between the US and the EU

Open issues in CI Building

5 – Aggregation methods

(5) Weights are customarily conceived as 'importance' measures. In additive aggregations

$$\sum_{i=1}^{k} x_i \cdot w_i$$

weights are substitution rates (wi/wj = ratio of substitution of indicator 'i' with indicator 'j.

This implies a compensatory logic, i.e. the possibility of renounce to one point of indicator 'i' with wi/wj points for indicator 'j' . However, when one is not willing to compensate (i.e. literacy with GDP per capita), a non-compensatory multi-criteria approach can be used (Munda and Nardo, 2003).

Multi-criteria approach

(6) With this approach no compensation occurs. To exemplify, a country that does marginally better on many indicators comes out better than a country that does a lot better on a few ones because it cannot compensate deficiencies in some dimensions with outstanding performances in others.

Ongoing work: the OECD JRC handbook

Points touched upon in this brief discussion of open issues in CI building are tackled in a forthcoming joint project from OECD and JRC on composite indicators building.

Handbook on Constructing Composite Indicators: Methodology and User Guide

It aims to be a guide to the construction and use of CI.

Reviews on methodologies and practices on composite indicators :

State-of-the-art Report on Current Methodologies and Practices for Composite Indicator Development (2002) Michaela Saisana & Stefano Tarantola, European Commission, Joint Research Centre

Literature Review of Frameworks for Macro-indicators (2004), Andrew Sharpe, Centre for the Study of Living Standards, Ottawa, CAN.

Composite indicators of country performance: a critical assessment (2003) Michael Freudenberg, OECD.

Measuring performance: An examination of composite performance indicators (2004) Rowena Jacobs, Peter Smith, Maria Goddard, Centre for Health Economics, University of York, UK.

Methodological Issues Encountered in the Construction of Indices of Economic and Social Well-being (2003) Andrew Sharpe Julia Salzman

Methodological Choices Encountered in the Construction of Composite Indices of Economic and Social Well-Being, Julia Salzman , (2004) Center for the Study of Living Standards , Ottawa, CAN.

http://farmweb.jrc.cec.eu.int/ci/

EUROPEAN COMMISSION
DIRECTORATE GENERAL
Joint Research Centre

OECD

Composite Indicators

An information server on composite indicators

- Home
- FAQ
- Indexes
- Papers & Books
- Handbook
- JRC Activities
- Visual Models
- Conferences
- People
- Contact
- Disclaimer

In brief...

"[...] *it is hard to imagine that debate on the use of composite indicators will ever be settled* [...] *official statisticians may tend to resent composite indicators, whereby a lot of work in data collection and editing is "wasted" or "hidden" behind a single number of dubious significance. On the other hand, the temptation of stakeholders and practitioners to summarise complex and sometime elusive processes (e.g. sustainability, single market policy, etc.) into a single figure to benchmark country performance for policy consumption seems likewise irresistible.*"

Andrea Saltelli, JRC

What's New in 2005

A more reliable Environmental Sustainability Index through EU research. The European Commission's Joint Research Centre (JRC) has contributed to the Environmental Sustainability Index (ESI), published by the Yale and Columbia Universities in co-operation with the World Economic Forum and presented in Davos on 28 January 2005.

More on the story:
Press Releases
Main ESI Report (772 Kb)
Full ESI Report (3.3 Mb)

Workshops/Meetings in 2005

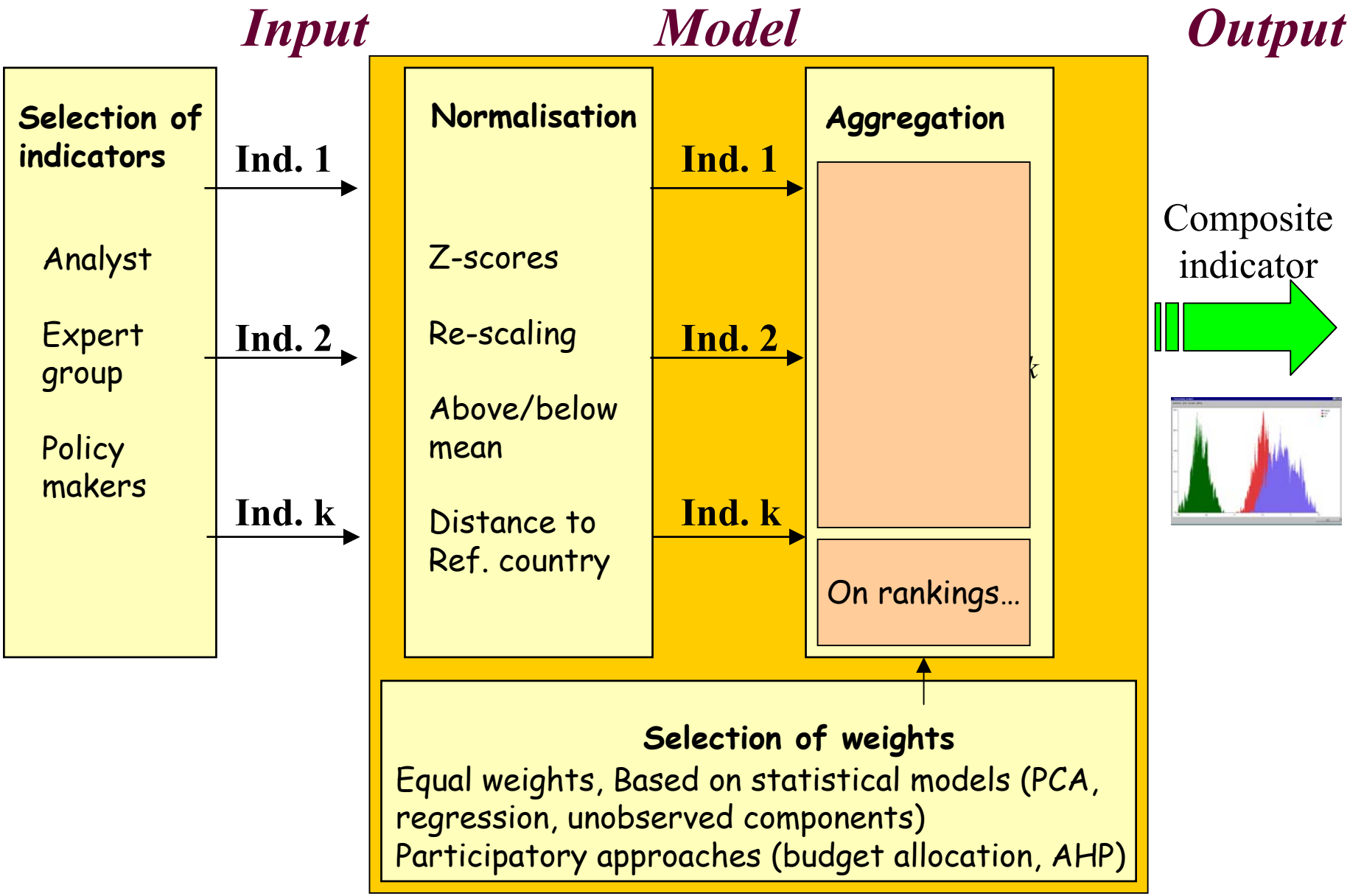
Title/Link	Venue,Date
Workshop on Frontiers in Benchmarking Techniques and Their Application to Official Statistics (organisers: OECD / Eurostat)	Luxembourg 7-8 April
e-2005 eChallenges Conference (supported by the European Commission)	Ljubljana, Slovenia 19-21 October

Please contact Michaela Saisana at [michaela.saisana@jrc.it]
for any comments or suggestions.

Last updated: 02/26/05

- Composite indicators as models … and the critique of models

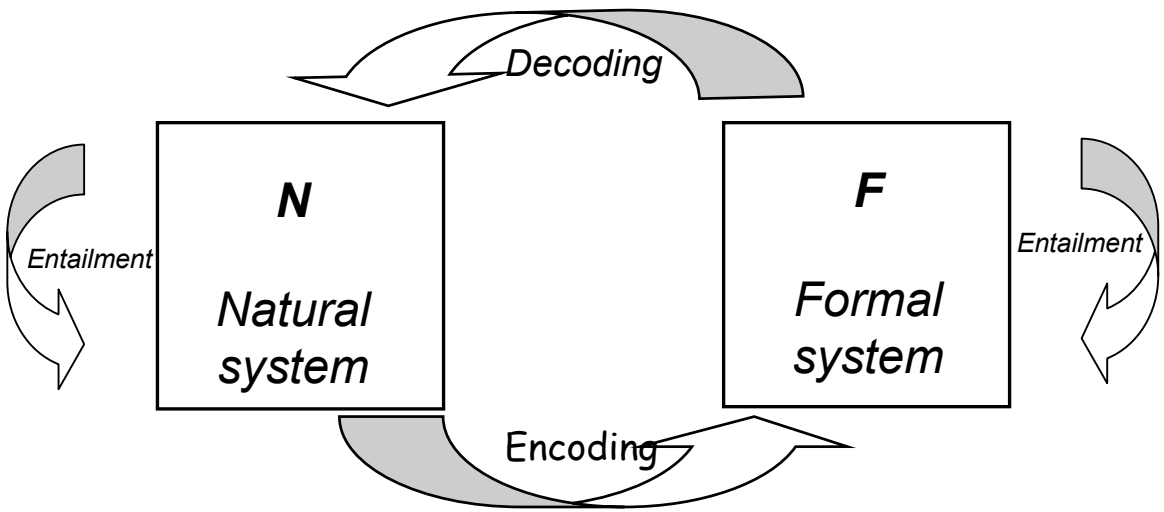
Composite indicators as models …



… and the critique of models (Rosen)

Building a model (encoding) is the result of a craftsmanship

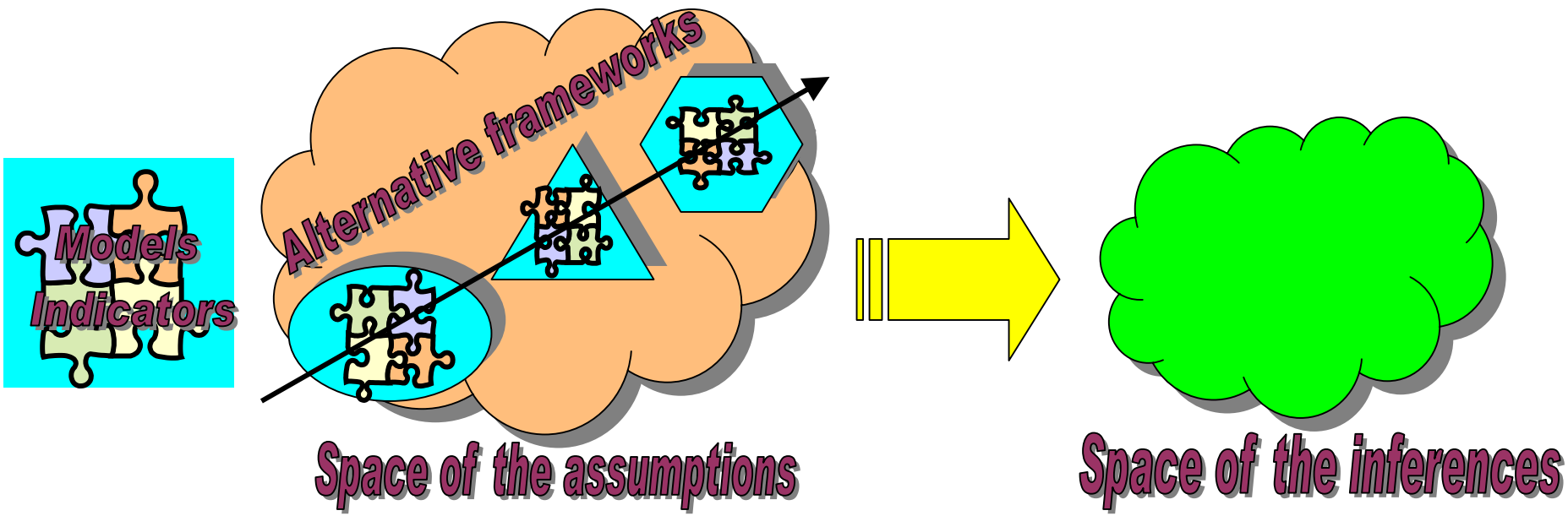
Interpreting the results of the model (decoding) is also a craftsmanship



Need for robustness analysis

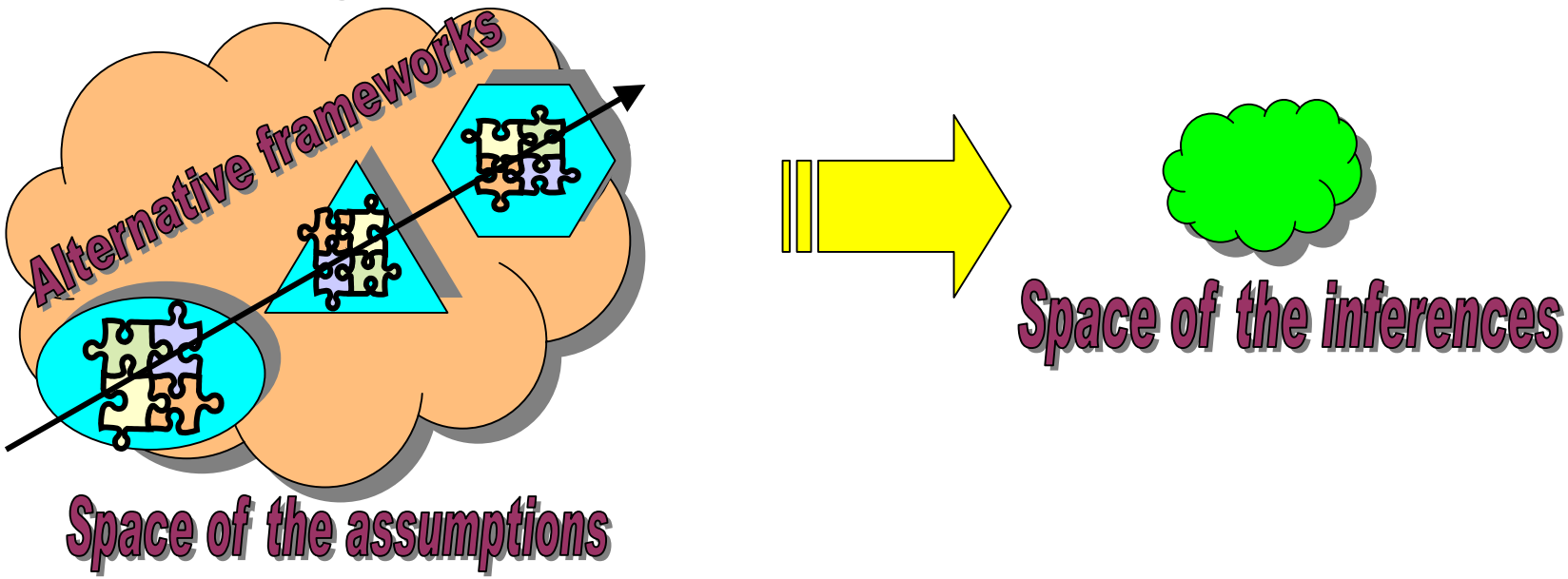
"I propose a form of organised sensitivity analysis in which a neighborhood of alternative assumptions is selected and the corresponding interval of inferences is identified.

Edward Leamer,
Economist at UCLA



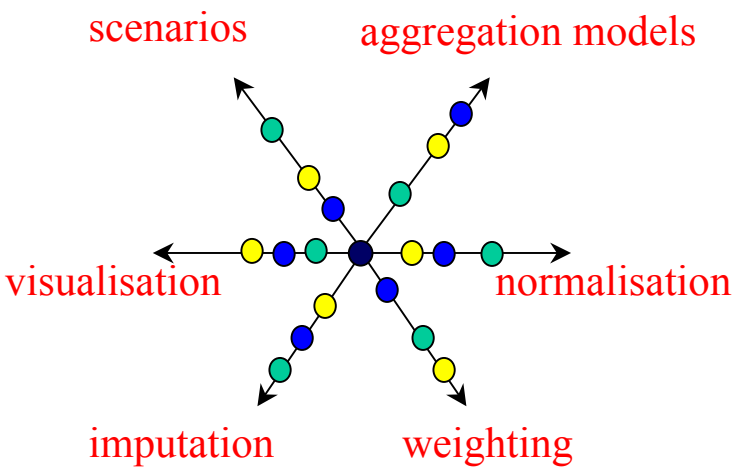
Need for robustness analysis

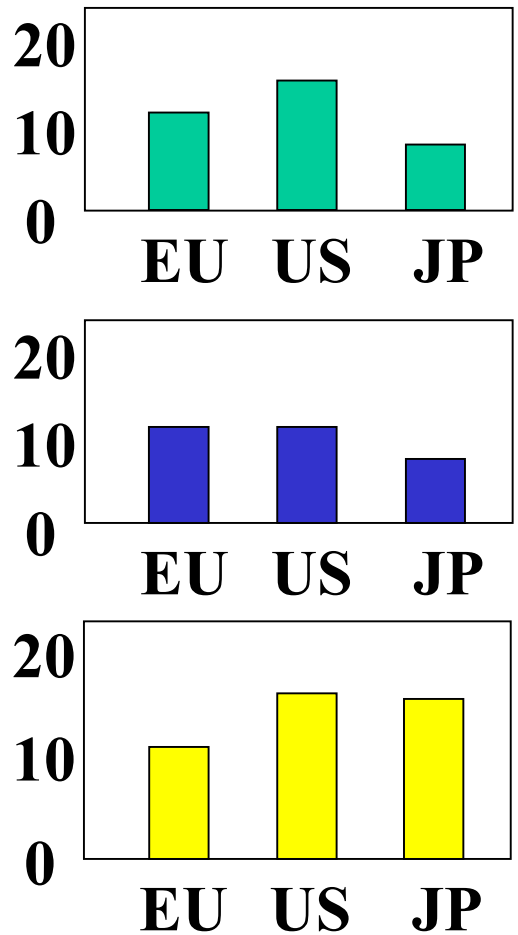
Conclusions are judged to be sturdy only if the neighborhood of assumptions is wide enough to be credible and the corresponding interval of inferences is narrow enough to be useful."



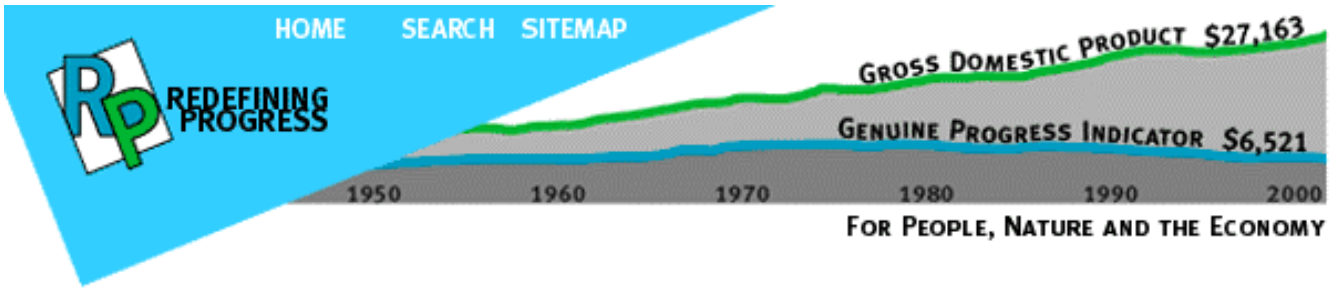
Robustness analysis (scheme)

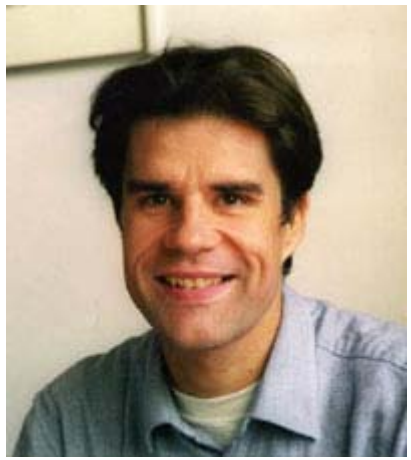
Space of alternatives





Sensitivity analysis (1)





"The [ESI] report would gain from a more extensive peer review and a sensitivity analysis. The lacking sensitivity analysis undermines the confidence in the results since small changes in the index architecture or the weighting could dramatically alter the ranking of the nations."

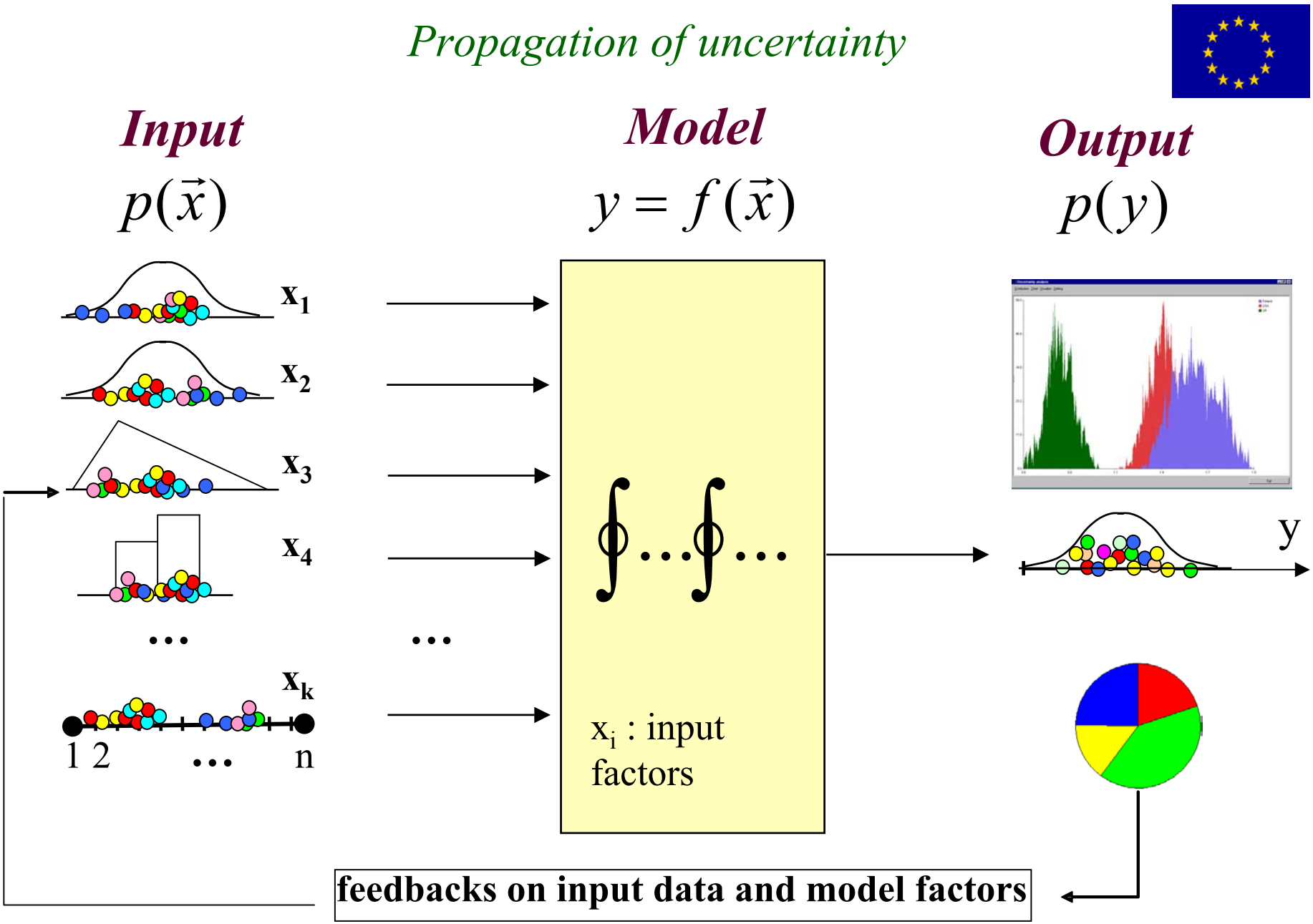
Mathis Wackernagel, mental father of the "Ecological Footprint"

Sensitivity analysis (2)

"The validity and reliability of any indicator has to be tested and the sensitivity of the outcomes and results have to be checked meticulously"

Rainer Frietsch
Fraunhofer ISI, Germany
Comments on the European Innovation Scorel
February 2005

Propagation of uncertainty



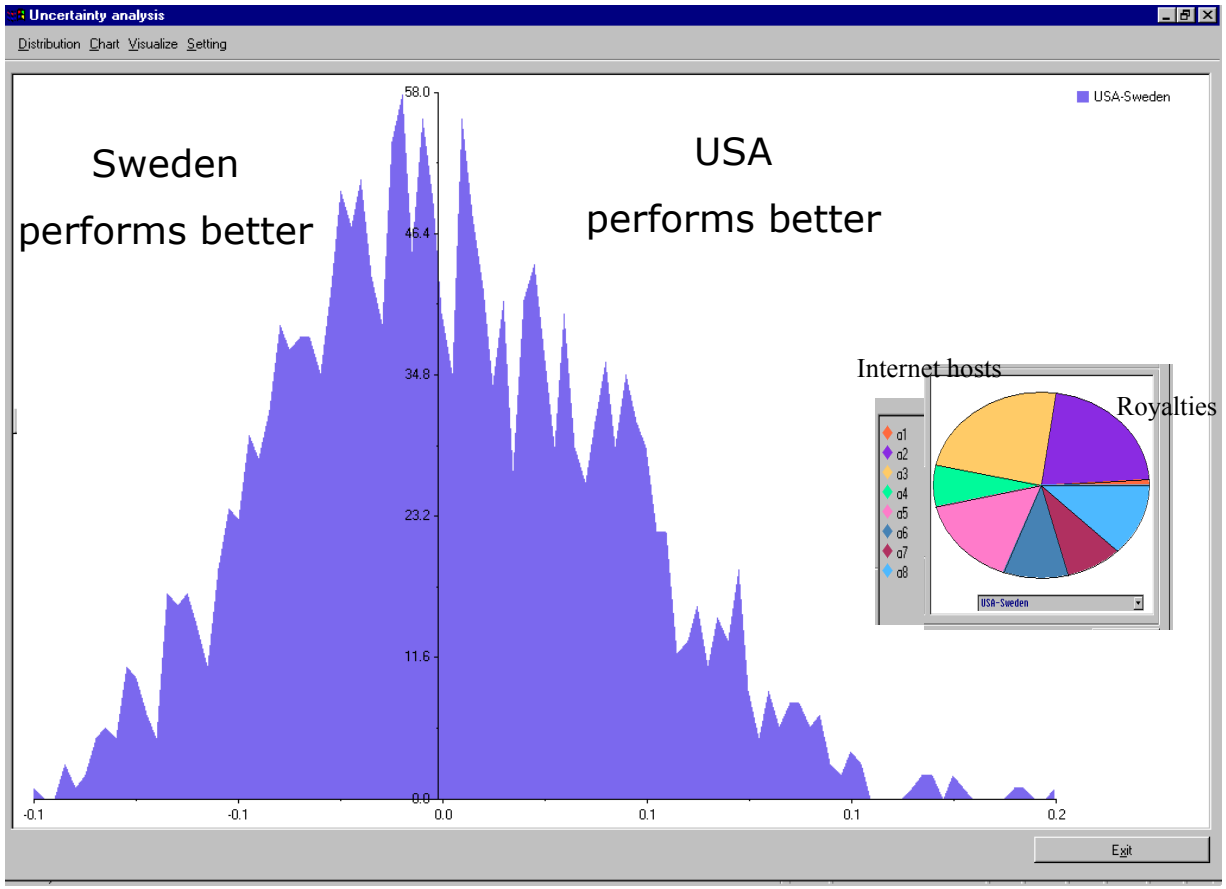
Sensitivity analysis (what for?)

1. capture the plurality of the debate around component indicators, weights and aggregation methods (i.e. acknowledge the opinions of all the stakeholders).

2. to identify robustly groups of countries with similar performance and countries which undoubtedly outperform others.

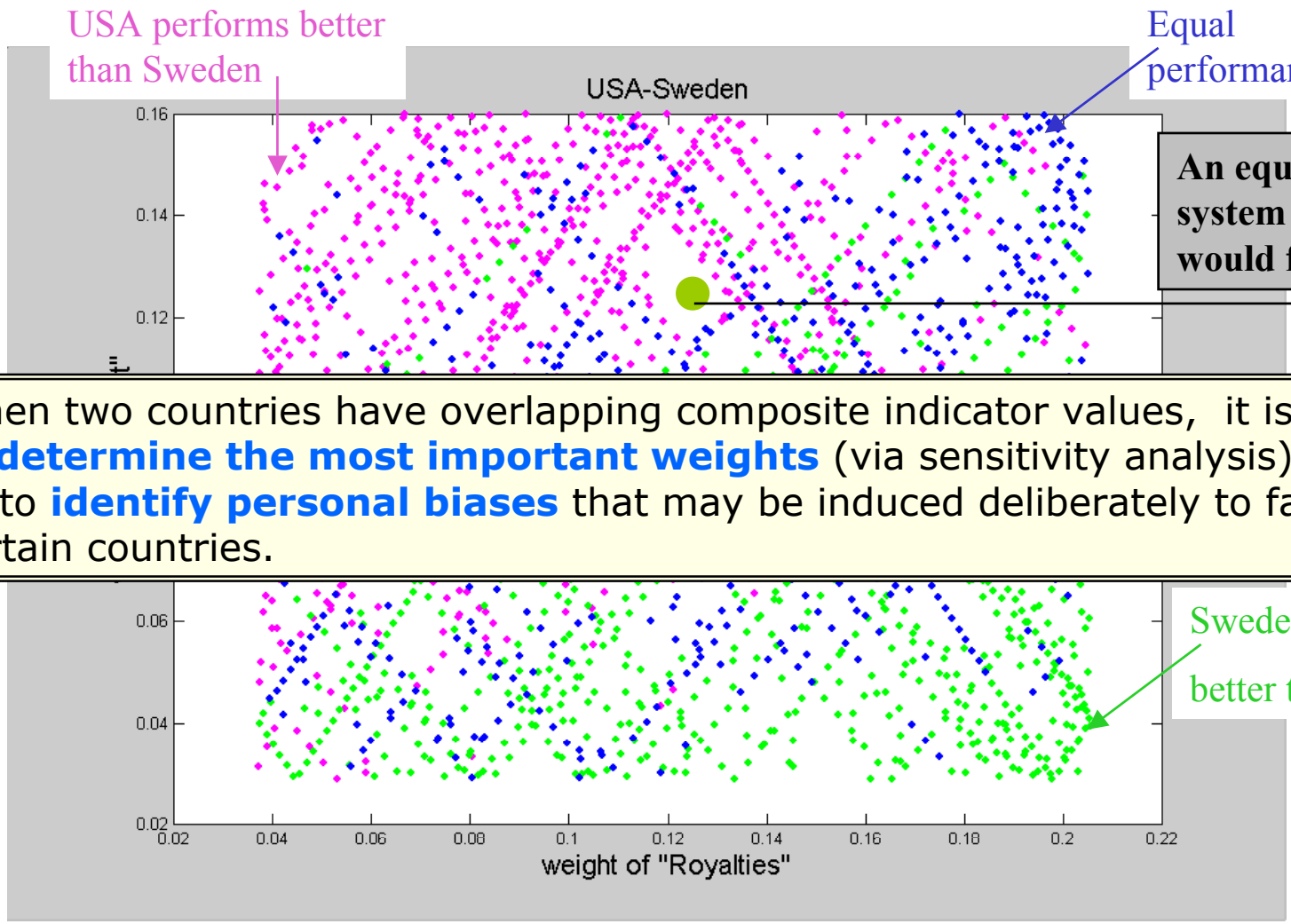
3. to identify the regions in the space of the weights that favor one country with respect to another

Sensitivity analysis (what for?)



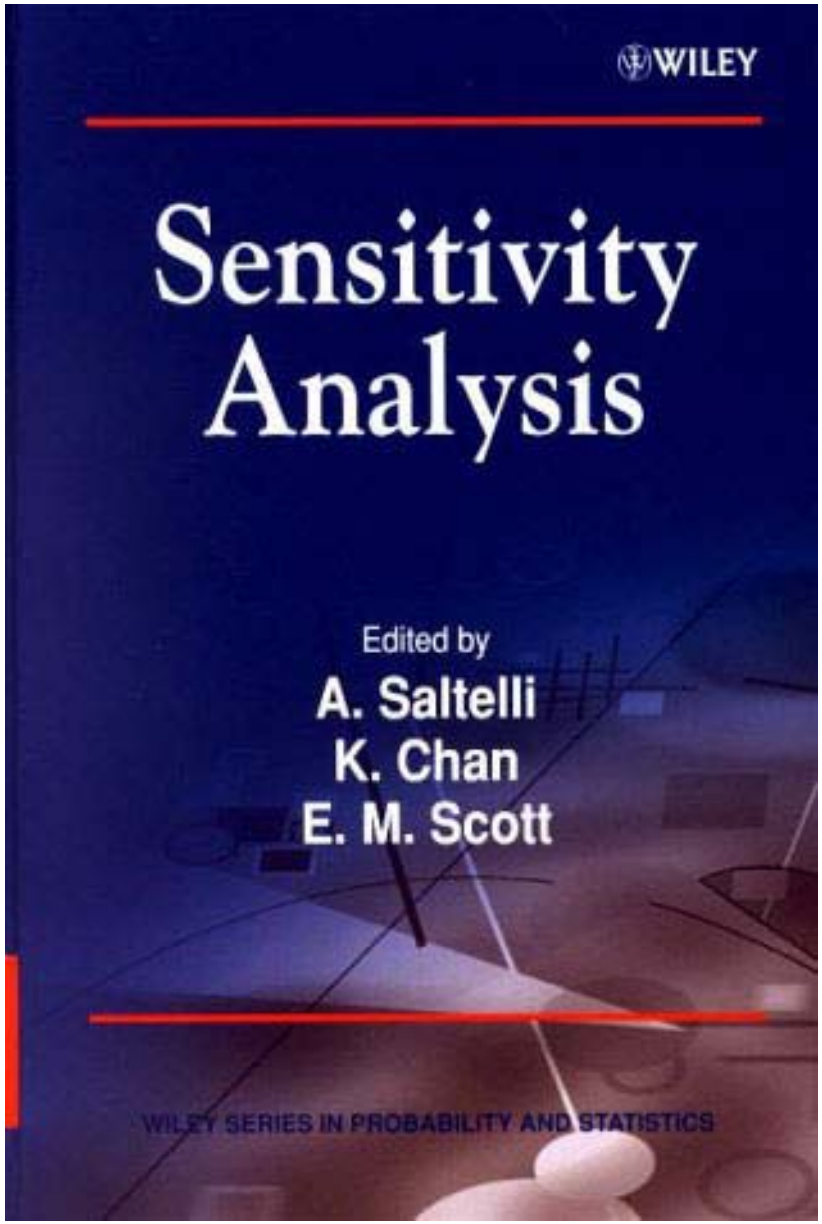
Given the partial overlapping between the two countries, we want to identify the weights that mostly influence this overlap through a sensitivity analysis

Scatterplot of most important weights



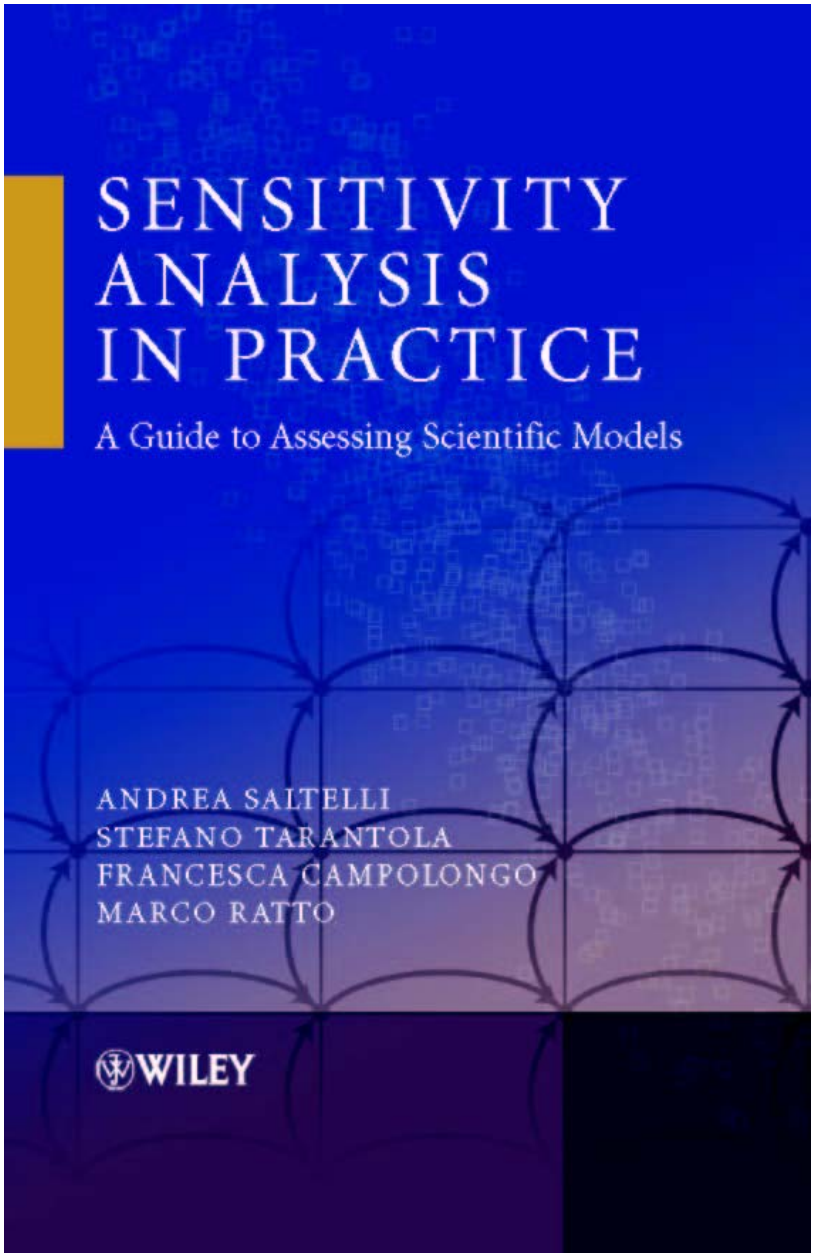
An equal weighting system (0.125) would favor Sweden

When two countries have overlapping composite indicator values, it is wise to **determine the most important weights** (via sensitivity analysis), so as to **identify personal biases** that may be induced deliberately to favor certain countries.



Sources: a multi-author book published in 2000. Methodology and applications by several practitioners.

Chapter1, Introduction and 2, Hitch Hiker guide to sensitivity analysis offer a useful introduction to the topic



Sources: a 'primer', just published by Wiley, is the real introductory book to the topic – its examples are based on a software, SIMLAB that can be freely downloaded from the web.

CI s of the Knowledge-based economy

Investment and Performance on the KBE – DG RTD

Summary Innovation Index 2004 (and 2005 draft) – DG ENTR

Welfare of Nations (Statistics Sweden, 2004)

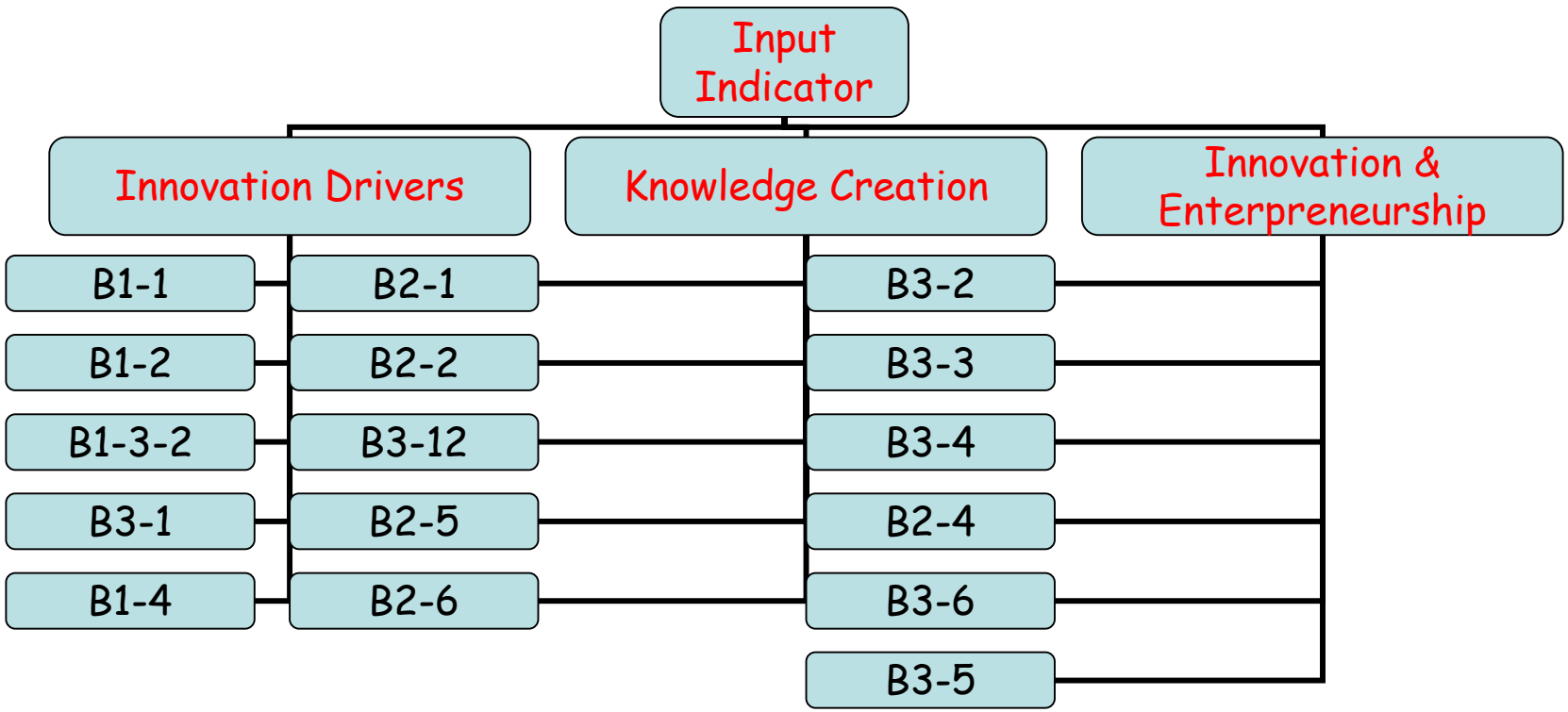
National Innovation Capacity (Porter and Stern)

The 2004 State Technology and Science Index (MilkenInstitute, US)

General Indicator of Science and Technology (NISTEP, Japan)

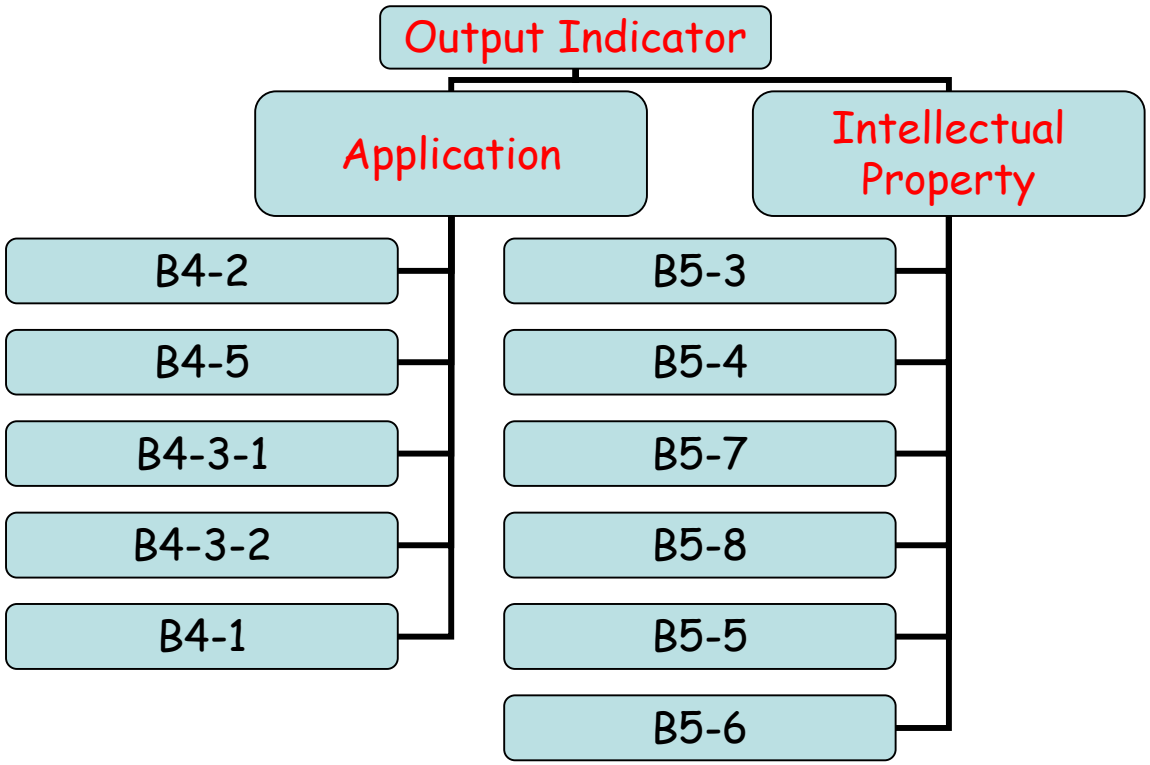
The European Competitiveness Index 2004, Robert Huggins Associates

Summary Innovation Index 2005 (draft) – DG ENTR



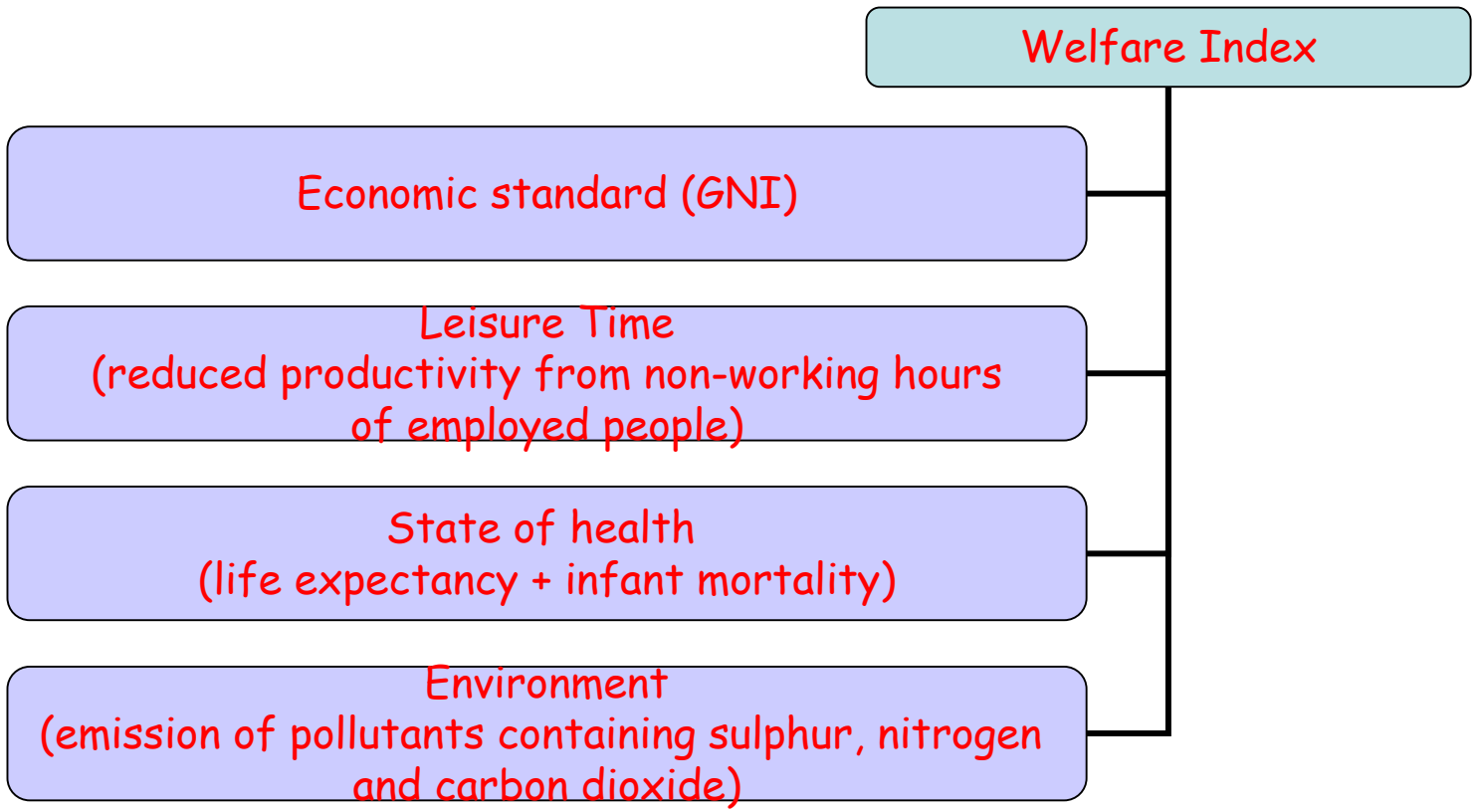
Weights supplied by GSO members in innovation via budget allocation

Summary Innovation Index 2005 (draft) – DG ENTR

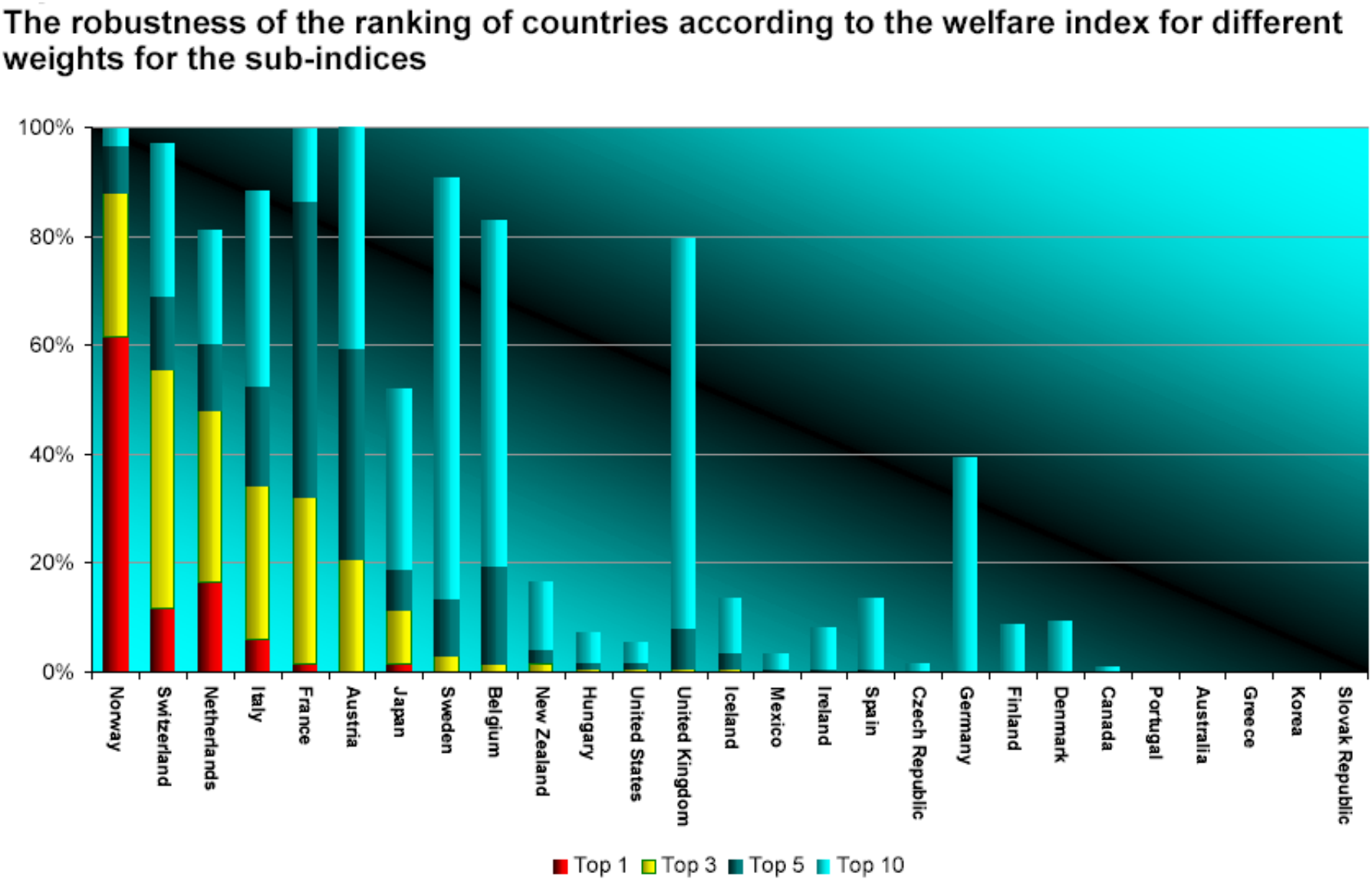


Welfare of Nations (Statistics Sweden, 2004)

OECD countries, min-max normalisation,
Robustness analysis using Monte Carlo with random weights



Welfare of Nations (Statistics Sweden, 2004)



The European Competitiveness Index 2004

Large basket of indicators relating to

1. creativity
2. the knowledge economy
3. sectoral productivity performance
4. economic performance,
5. infrastructure and accessibility.

Covers all European member states and regions, Norway and Switzerland

The gap in competitiveness in Europe's regions is increasing
Uusimaa (Helsinki), Stockholm, Brussels, and Ile de France and Hamburg are the drivers of Europe's knowledge creation and utilisation.
Competitiveness and cohesion are working in contradictory directions.

The 2004 State Technology and Science Index

The index uses 75 indicators in five composite categories to measure how well a US state performs in today's knowledge-based economy:

- Research and development inputs;
- Risk capital and entrepreneurial infrastructure;
- Human capital investment;
- Technology and science workforce; and
- Technology concentration and dynamism.

The index identifies the US regions that can attract firms and industries proficient at converting knowledge into successful innovations, products and services.