

Máster Universitario en Administración y Dirección de Empresas Full Time MBA

Quantitative methods for decision making

Professor Andrea Saltelli

Crunching numbers for sustainability

August 25 2023: The politics of modelling is out!



Praise for the volume

"A long-awaited examination of the role—and obligation—of modeling."

Nassim Nicholas Taleb, Distinguished Professor of Risk Engineering, NYU Tandon School of Engineering, Author, of the 5-volume series *Incerto*.

"A breath of fresh air and a much needed cautionary view of the ever-widening dependence on mathematical modeling."

Orrin H. Pilkey, Professor at Duke University's Nicholas School of the Environment, co-author with Linda Pilkey-Jarvis of *Useless Arithmetic: Why Environmental Scientists Can't Predict the Future*, Columbia University Press 2009.

Mastodon Toots by

@AndreaSaltelli



AndreaSaltelli

2023/08/25 11:03

Thanks to Maria Kozlova of LUT University in Finland for taking and curating this recording. My trajectory from number crunching to thinking about numbers' role in human affairs

[youtube.com/watch?v=...](https://www.youtube.com/watch?v=...)
—@NCC-PolSci UK

View on [mastodon.social](#)

In this set of slides:

- 1 The multiverse of composite indicators

1.

The multiverse of composite indicators

There is now a composite indicator for everything, from the quality of urban life to the freedom of the press passing by the ranking of institutes of higher education. These statistical indicators built aggregating different variables are an excellent vehicle for drawing attention to a particular issue, but should be constructed and / or used with care. The ecological footprint and other examples.

Composite indicators: What are they?



WJP Rule of Law Index 2019

Eight factors further disaggregated into 44 sub-factors



Constraints on Government Powers

- 1.1 Government powers are effectively limited by the legislature
- 1.2 Government powers are effectively limited by the judiciary
- 1.3 Government powers are effectively limited by independent auditing and review
- 1.4 Government officials are sanctioned for misconduct
- 1.5 Government powers are subject to non-governmental checks
- 1.6 Transition of power is subject to the law

One of the eight factors with its 6 sub factors ...

https://worldjusticeproject.org/sites/default/files/documents/WJP-ROLI-2019-Single%20Page%20View-Reduced_0.pdf

Where indicators come
from: from images to
numbers



Ambrogio Lorenzetti (c. 1290 – 1348), Allegory: Effects of good and bad government, Palazzo Pubblico di Siena, (detail: Safety)



Ambrogio Lorenzetti (c. 1290 – 1348), Allegory: Effects of good and bad government, Palazzo Pubblico di Siena, Effects of good government in the city



Ambrogio Lorenzetti (c. 1290 – 1348), Allegory: Effects of good and bad government, Palazzo Pubblico di Siena, Effects of the good government in the countryside



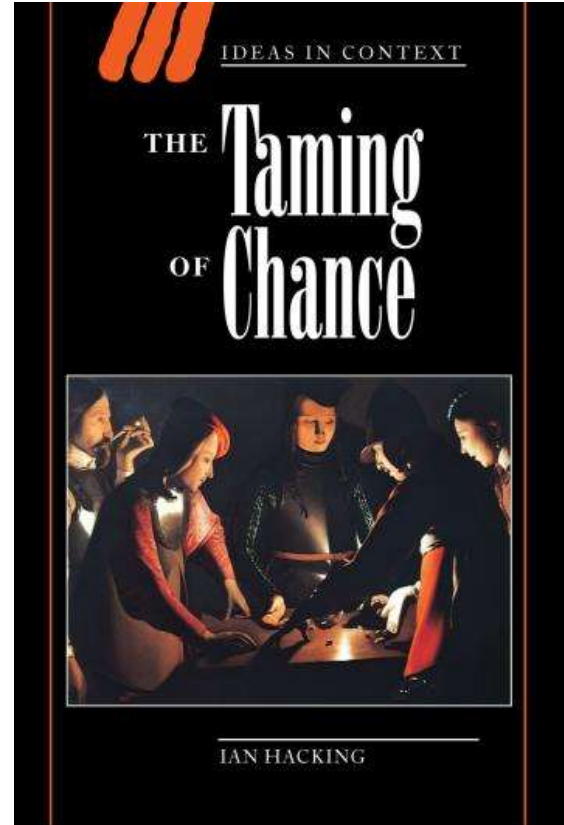
Ambrogio Lorenzetti (c. 1290 – 1348), *Allegory: Effects of good and bad government*, Palazzo Pubblico di Siena, Effects of the bad government in the countryside

“ [⋯] Lorenzetti is modern because he invites the spectator of his frescoes, [⋯] to assess governments not so much on the political principles and values [⋯] than on the practical consequences of their decisions.

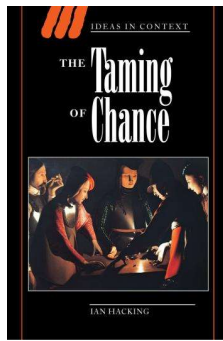
The paintings on the effects of good and bad governance show them were to look, what are the signs, or, in contemporary language, the indicators, to monitor in that respect.”

How numbers flooded the modern world

Ian Hacking, 1990, *The taming of chance*,
Cambridge University Press.



‘Probability’ won an epistemological war between the eighteenth and the nineteenth century.



‘Probability’ became king in adjudicating the credibility of evidence.

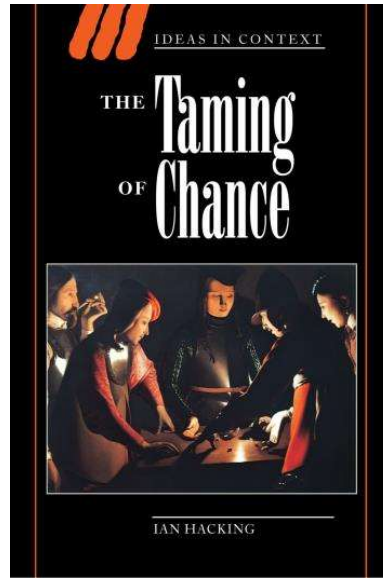
We look at facts mostly through the lenses of statistics – before the enlightenment chance was equated with superstition.

Leibnitz, ‘philosophical godfather of Prussian official statistics’ advising Prince Frederik of Prussia 1700

Leibnitz’s first proposal for a statistical office

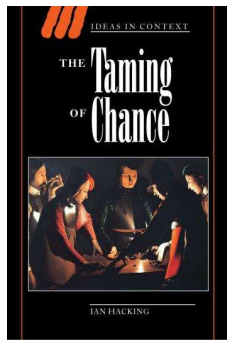


Gottfried Wilhelm
Leibniz (1646–1716)



Statistics ↔ nation state ↔ Modernity?

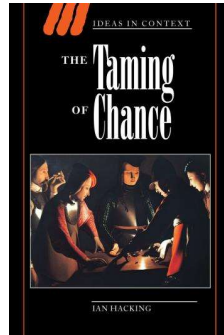
The victory of probability is metaphysical (quantum mechanics), epistemological (statistics as a way of knowing things), logical (statistical inference methods) and ethical (no decision taken without statistical evidence), leading to the ‘imperialism of probability’...



Statistics \leftrightarrow nation state \leftrightarrow Modernity

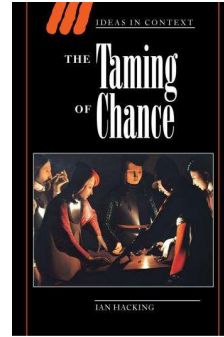
56 categories to ‘measure the power of a state’, the first scoreboard;

- number of marriageable girls,
- able bodied capable to carry arms,
- diseases,
- child mortality,
- ...
- number of Jews



Gottfried Wilhelm
Leibniz (1646–1716)

- number of Jews

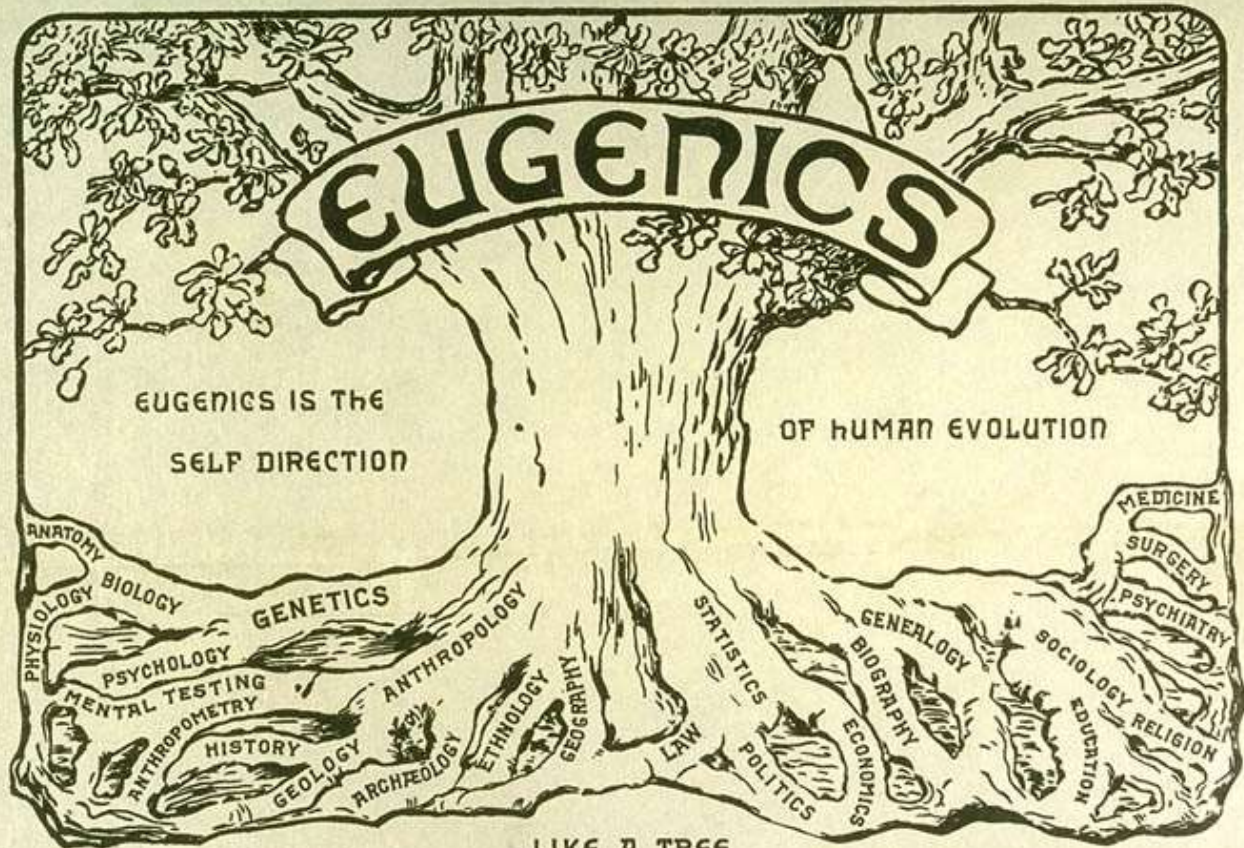


Already in 1745 Jews are being treated as a separate category and counted in Prussian statistics (statistics and antisemitism is one of the chapters in the book)

Statistics has a story in eugenics (Pearson, Galton)



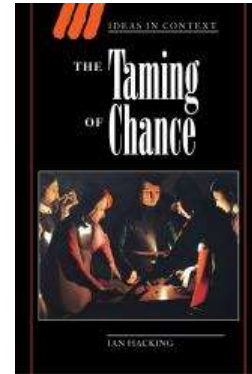
Gottfried Wilhelm Leibniz
(1646–1716)



LIKE A TREE
EUGENICS DRAWS ITS MATERIALS FROM MANY SOURCES AND ORGANIZES
THEM INTO AN HARMONIOUS ENTITY.

Some sins of statistics ...

Francis Galton and Karl Pearson (the one of chi-squared) and their laboratory of biometrics ...
distinguishing army officers from private soldiers
from criminals convicted of murder from non-violent felons from Jews ...

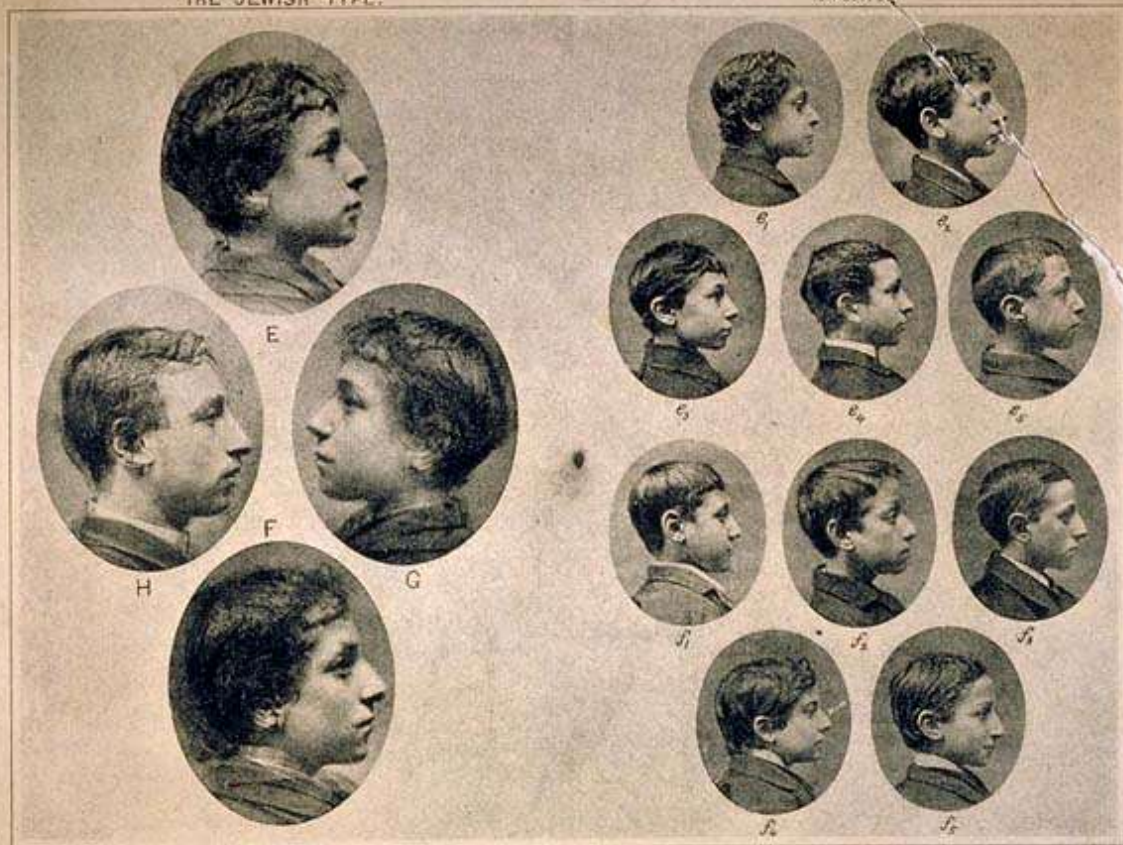


EUGENICS
"IS THE STUDY OF THE AGENCIES UNDER SOCIAL CONTROL, THAT IMPROVE OR IMPAIR THE RACIAL QUALITIES OF FUTURE GENERATIONS EITHER PHYSICALLY OR MENTALLY."
SIR FRANCIS GALTON.

Truman State University. Noncommercial, educational use only.

THE JEWISH TYPE.

Profile.



COMPOSITES.

Components. FRANCIS GALTON, F.R.S. PHOTO.

ILLUSTRATIONS OF COMPOSITE PORTRAITURE.



Coming closer to our times, the story of the first R&D Statistics ever. Benoît Godin (2010) tells us what these researchers thought: Measuring the numbers of sons and daughters of scientists will tell us whether a society degenerates toward stupidity.



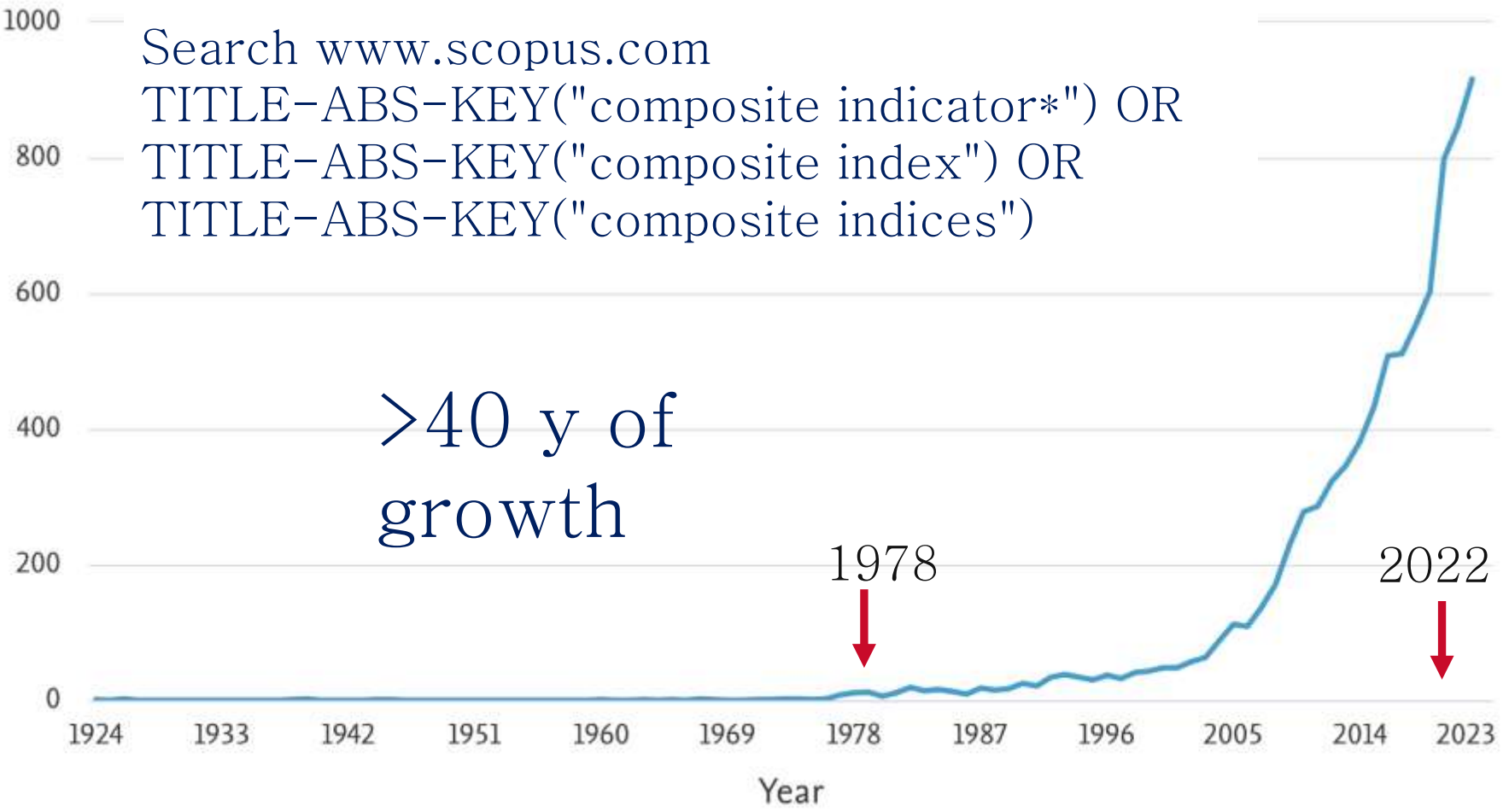
Godin, B., From Science to Innovation, INRS, Montreal, Canada, Communication presented to the Government-University-Industry Research Roundtable (GUIRR) US National Academy of Sciences, Washington, May 21, 2010.

Ubiquity of composite indicators

Search www.scopus.com
TITLE-ABS-KEY("composite indicator*") OR
TITLE-ABS-KEY("composite index") OR
TITLE-ABS-KEY("composite indices")

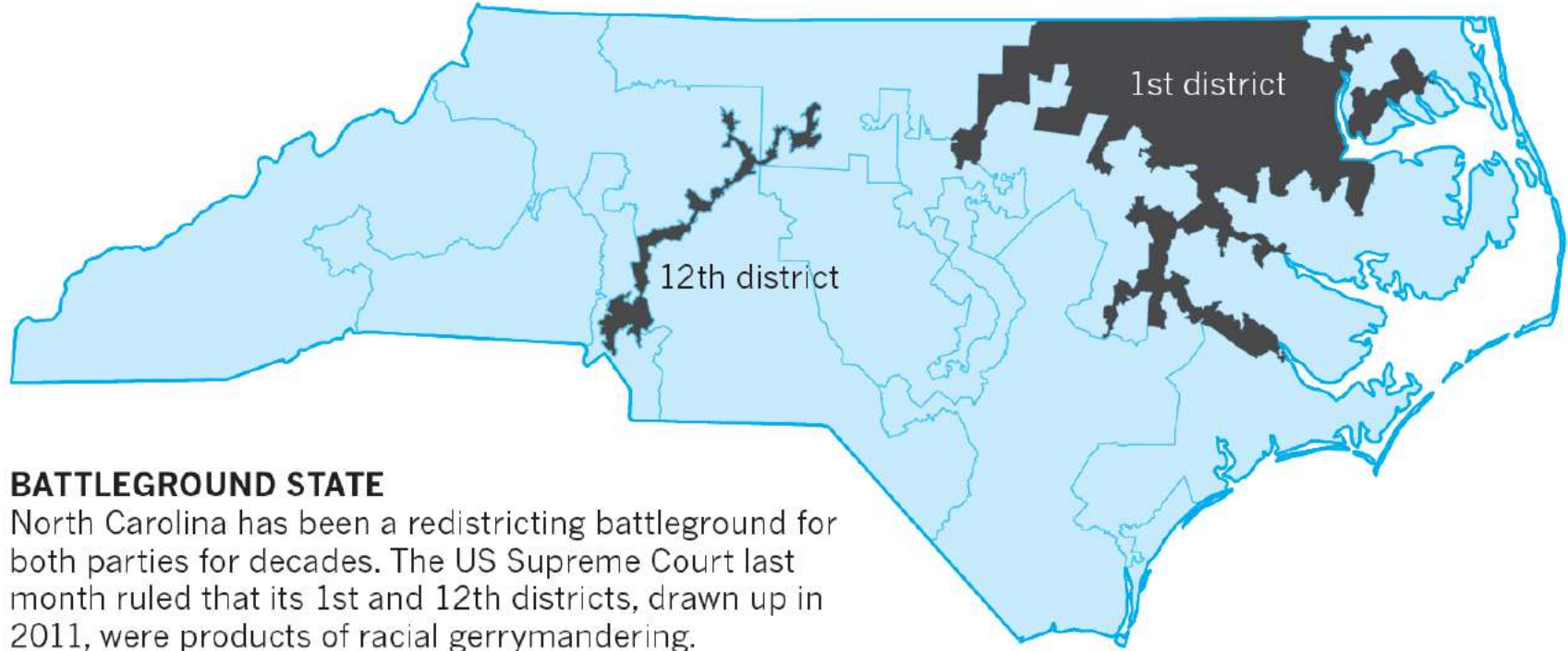
Documents

>40 y of
growth



Useful composite indicators

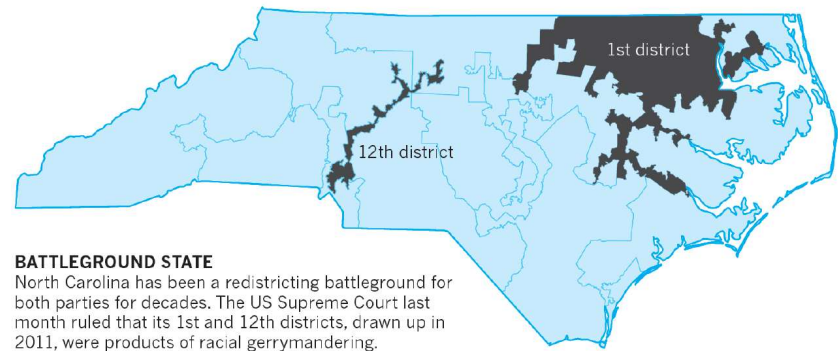
Making the case for gerrymandering?



BATTLEGROUND STATE

North Carolina has been a redistricting battleground for both parties for decades. The US Supreme Court last month ruled that its 1st and 12th districts, drawn up in 2011, were products of racial gerrymandering.

Nature June 2017 article on the mathematics of ‘nailing’ gerrymandering



BATTLEGROUND STATE
North Carolina has been a redistricting battleground for both parties for decades. The US Supreme Court last month ruled that its 1st and 12th districts, drawn up in 2011, were products of racial gerrymandering.

“[US] ranked 55th of 158 nations — last among Western democracies — in a 2017 index of voting fairness (Electoral Integrity Project)”

Carrie Arnold, 2017, The mathematicians who want to save democracy, 200, NATURE, VOL 546, 8 JUNE 2017.

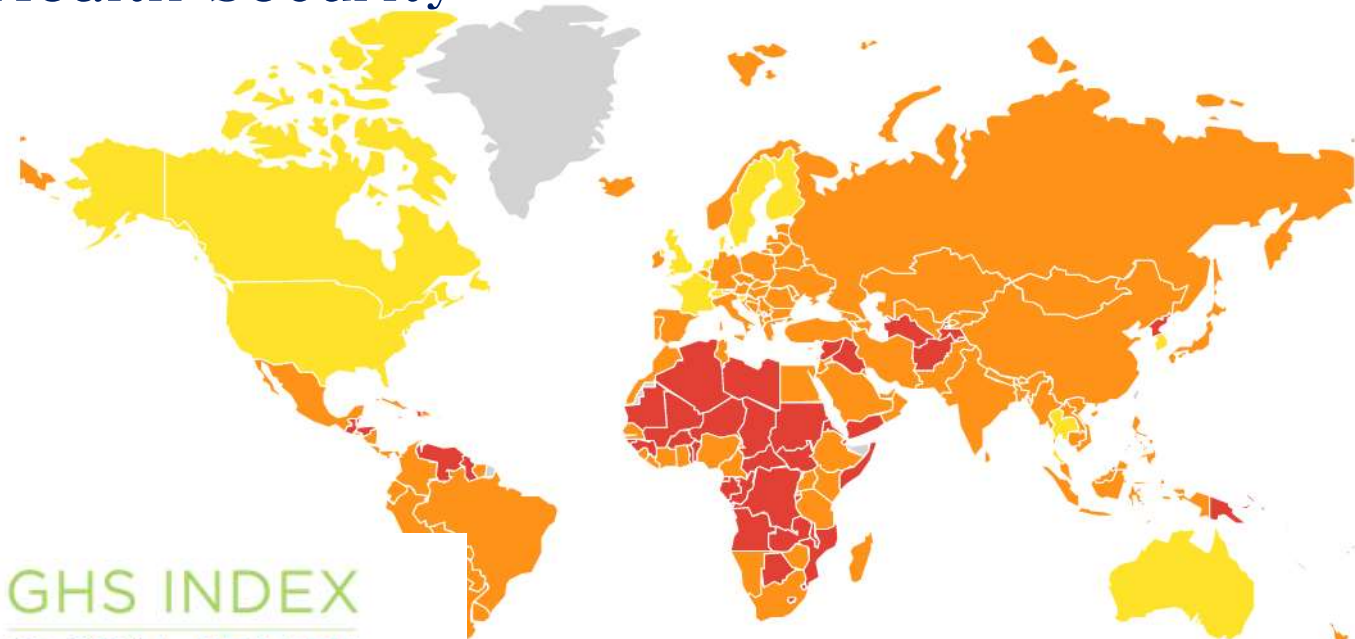
At times wrong

The Global Health Security Index, released 2019 to “spur measurable changes in national health security” in light of “high-consequence and globally catastrophic biological events”



Cameron, E.E. et al., Global Health Security Index. Building Collective Action and Accountability. Nuclear Threat Initiative & Johns Hopkins Center for Health Security (October 2019). Available at <https://www.ghsindex.org/#1-section--map>.

US and UK rank 1 and 2 respectively in Global Health Security



GHS INDEX
GLOBAL HEALTH
SECURITY INDEX

Key

- Most Prepared
- More Prepared
- Least Prepared

Select a country to see Overall Score/Rank and access a full country page.

M. Kaiser, A. T.-Y. Chen, and P. Gluckman, “Should policy makers trust composite indices? A commentary on the pitfalls of inappropriate indices for policy formation,” arXiv.org, vol. 2008.13637, Aug. 2020.

How can a country ranked last in quality of health care, with a raging opioid pandemic, be rated first in 'preparedness'?

See also:

<https://www.nybooks.com/articles/2020/10/22/best-health-care/>

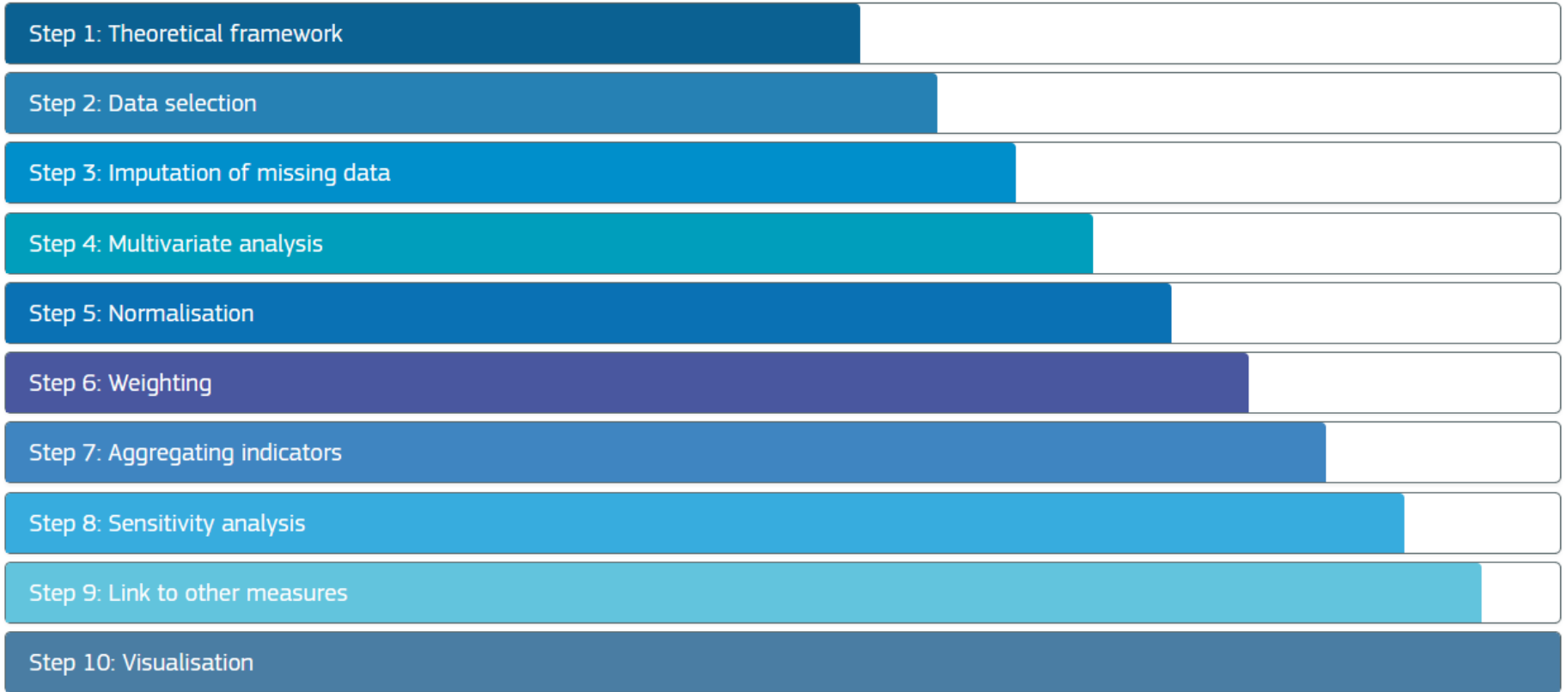
<https://www.commonwealthfund.org/publications/newsletter-article/us-ranks-last-among-seven-countries-health-system-performance>



Quality of composite indicators



Ten steps



Source: <https://composite-indicators.jrc.ec.europa.eu/?q=10-step-guide>



Specific elements of quality for composite indicators

RELEVANCE

In the context of composite indicators, relevance has to be evaluated considering the degree to which it meet current and potential needs of the users

[...] ensure that the **right range of domains is covered in a balanced way**

ACCURACY

The credibility of data products refers to confidence that users place in ... the image of the data producer, i.e., the brand image ...

[crucial] that the data are perceived to be produced professionally and that practices are transparent

(for example, data are not manipulated, nor their release timed in response to political pressure)

COHERENCE

... ensure coherence over time and across countries ...
Coherence across countries implies that from country to country the data are based on common concepts, definitions, classifications and methodology, or that any differences can be justified

Is a theory for composite indicators possible?

Elements for a comprehensive assessment of public indicators



Paul-Marie Boulanger

2014

Editor: Andrea Saltelli



Paul-Marie Boulanger

Paul-Marie Boulanger, 2014, Elements for a comprehensive
assessment of public indicators, Report EUR 26921 EN.
[http://publications.jrc.ec.europa.eu/repository/bitstream/JRC921
62/lbna26921enn.pdf](http://publications.jrc.ec.europa.eu/repository/bitstream/JRC92162/lbna26921enn.pdf)

CI as boundary objects, between analysis and advocacy, as:

- instruments of democratization of expertise;
- instruments of social discovery
- semiotic objects

Paul-Marie Boulanger, 2014, Elements for a comprehensive assessment of public indicators, Report EUR 26921 EN.

<http://publications.jrc.ec.europa.eu/repository/bitstream/JRC92162/lbna26921enn.pdf>

A triadic conception of the sign as structure connecting three elements:

- the sign properly said (S);
- an object (O) and
- an “interpretant”(I)



Charles Sanders Peirce,
the father of semiotics
1839–1914

“This monkey possess a sophisticated repertory of vocal signs for signaling the presence of a predator”



African vervet monkey
(*Cercopithecus aethiops*)

It can distinguish

- a terrestrial stalking one such as a leopard,
- an aerial raptor such as an eagle or
- a ground predator such as a snake



African vervet monkey
(*Cercopithecus aethiops*)

Sign ↔ Cry



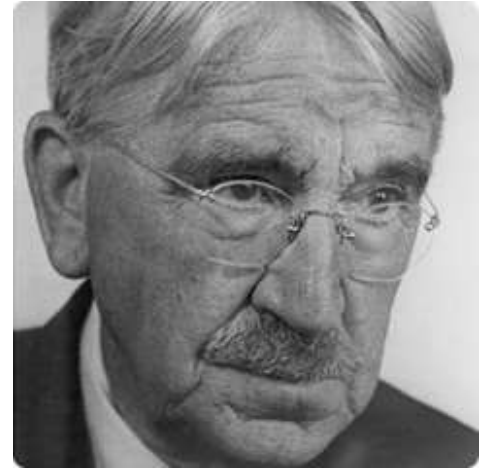
Object ↔ Predator



Interpretant ↔ Behaviour



Composite indicators as
instrumental to the creation of a
new public, through a process
of social discovery (J. Dewey)



John Dewey
1859–1952

Dewey, J., 1938. *The Public and its Problems*, Read
Book Ltd. Edition, 2013.

Why are ‘social discoveries’ needed?

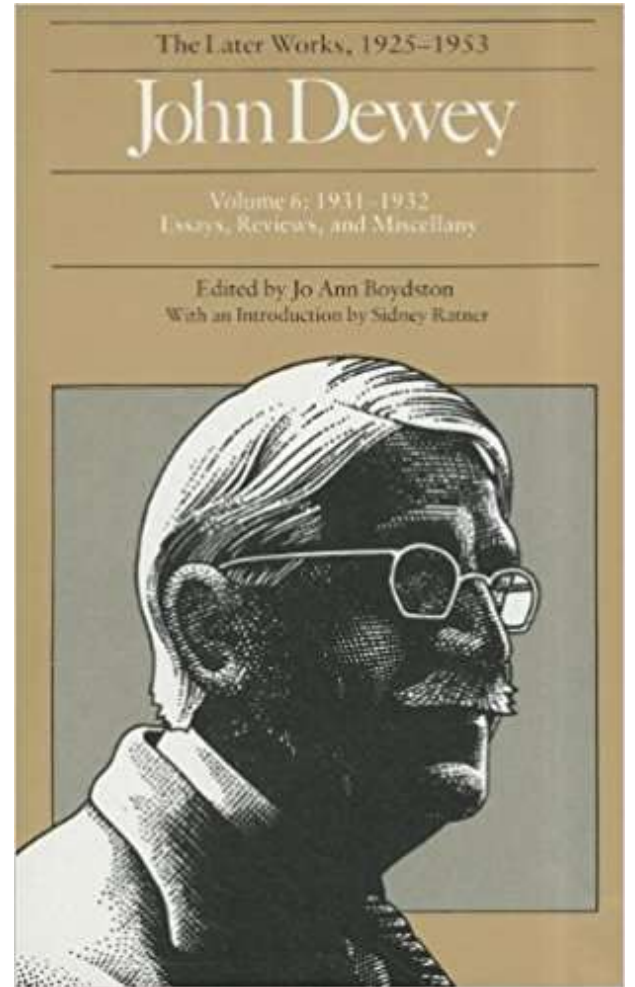
Because there are ‘publics’ affected by transaction taking place somewhere else.

“[...] machine age has so enormously expanded, multiplied, intensified and complicated the scope of the indirect consequences [...] that the resultant public cannot identify and distinguish itself”

Dewey, J., 1938. *The Public and its Problems*, Read Book Ltd. Edition, 2013.

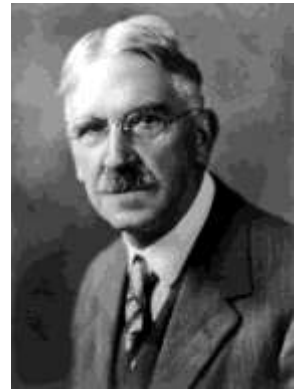
Social facts – unlike physical facts,
are only meaningful in a context of
desired ends

From J. Dewey 'Social Science and Social Control' in John
Dewey: The Later Works, 1925–1953: 1931–
1932, Vol. 6–ExLibrary,



Controversy Dewey- Lippmann (1920's): can the public express agency?

See also controversy between Habermas and Luhmann early 1970's. Today revisited by Philip Mirowski:
https://www.academia.edu/42682483/Democracy_Expertise_and_the_Post_Truth_Era_An_Inquiry_into_the_Contemporary_Politics_of_STS



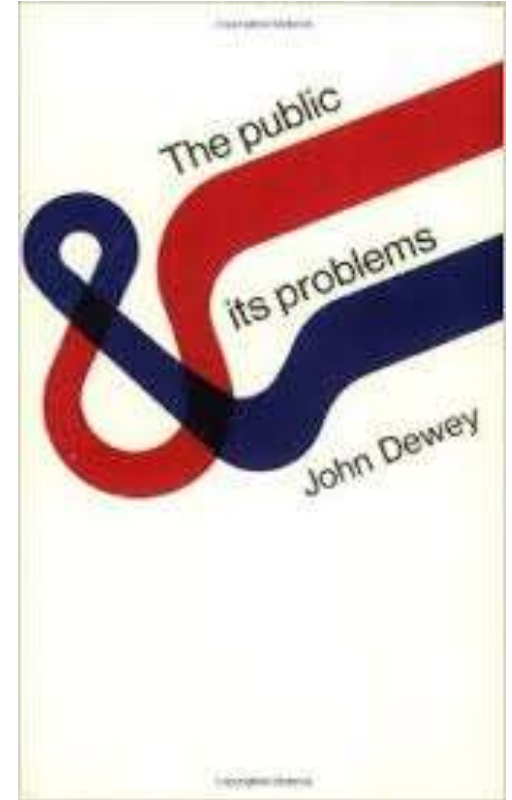
John Dewey
(1859–1952)



Walter Lippmann
(1889–1974)

“When writing “The public and its problem” Dewey was taking stance with Walter Lippmann, [who believed that] the citizens had become only spectators of a political game they were not able to play [because] they could not get informed of the relevant facts.”

Boulanger, Op. cit.



Building a composite indicator can be seen as a process of social discovery for which a model of extended participation comes natural. Frames and indicators are co-produced in the process which must be designed as to have a meaningful ‘interpretant’, or ‘end-in-sight’



Paul-Marie Boulanger, 2014, Elements for a comprehensive assessment of public indicators, Report EUR 26921 EN. <http://publications.jrc.ec.europa.eu/repository/bitstream/JRC92162/lbna26921enn.pdf>

Critique of composite indicators: the Fitoussi–Stiglitz–Sen report

“The role [of statistical indicators] has increased significantly over the last two decades.

This reflects improvements in the level of **education** in the population, increases in the **complexity** of modern economies and the widespread use of **information technology**”



Jean-Paul Fitoussi,
Amartya Sen, Joseph Stiglitz

CMEPSP (2009). Commission on the Measurement of Economic Performance and Social Progress, URL: [http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+ Commission+ report](http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report), last accessed June 2017.

“widespread use of information technology”

Visit https://joint-research-centre.ec.europa.eu/scientific-activities-z/composite-indicators_en
Resources on composite indicators building

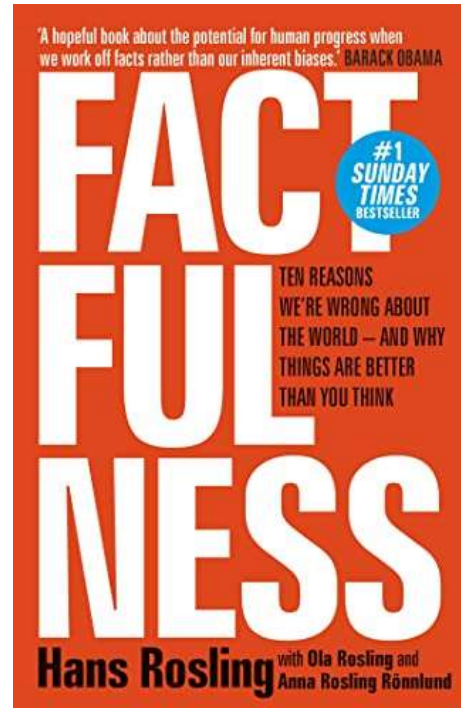
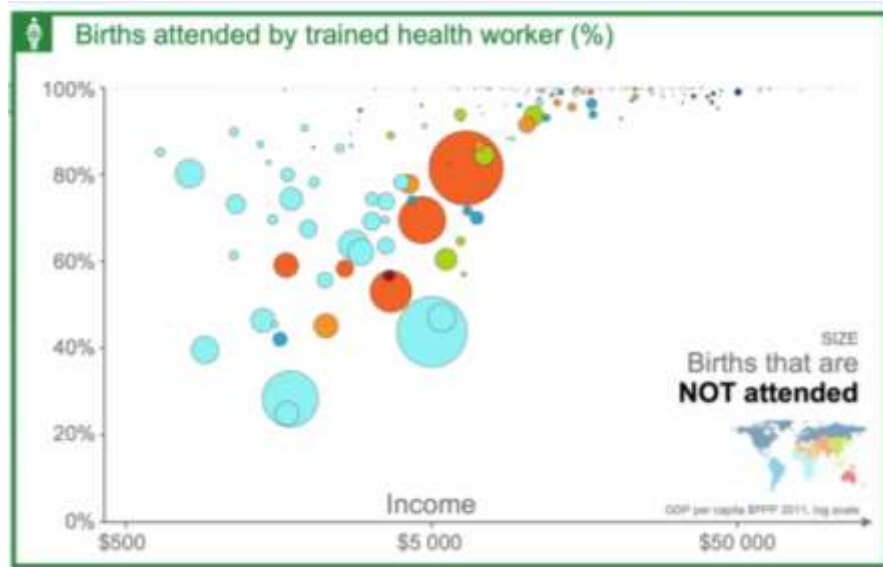


• [Composite Indicators and Scoreboards Explorer One-Pager](#)

• [Mini-Guide Composite Indicators & Scoreboards Explorer](#)

“widespread use of **information technology**”

Visit Hans Rosling’s <https://www.gapminder.org/>



Hans Rosling
1948–2017

“a general criticism ... frequently addressed at composite indicators, i.e. the arbitrary character of the procedures used to weight their various components [...]



Jean-Paul Fitoussi,
Amartya Sen, Joseph Stiglitz

[...] an aggregation procedure always means putting relative values on the items that are introduced in the index”

CMEPSP (2009). Commission on the Measurement of Economic Performance and Social Progress, URL: [http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+ Commission+ report](http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report), last accessed June 2017.

“The problem is **not that these weighting procedures are hidden, non-transparent or non-replicable** – they are often very explicitly presented by the authors of the indices, and this is one of the strengths of this literature.



Jean-Paul Fitoussi,
Amartya Sen, Joseph Stiglitz

The problem is rather that **their normative implications are seldom made explicit or justified**”

CMEPSP (2009). Commission on the Measurement of Economic Performance and Social Progress, URL: [http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+ Commission+ report](http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report), last accessed June 2017.

There are perhaps other driver behind the explosion of CI's

- More hybridization of roles, styles → CI as boundary objects, between analysis and advocacy
- Changing styles of governance → More actors with a voice
- Issues with trust / quality in the scientific enterprise → Do it yourself movements
- More controversy (**wicked issues**) → More 'weapons' developed



Policy Sciences 4 (1973), 155-169
© Elsevier Scientific Publishing Company, Amsterdam—Printed in Scotland

Dilemmas in a General Theory of Planning*

HORST W. J. RITTEL
Professor of the Science of Design, University of California, Berkeley

MELVIN M. WEBBER
Professor of City Planning, University of California, Berkeley

Statistics/indicators for governance: three models

- A **rational-positivist model** for the use of indicators and policy (good quality statistics describe reality and underpin good policies)
- A **discursive-interpretive model** (statistics contribute to a process of framing of and focusing on an issue among the many competing for public's attention)
- A **strategic model** (statistics is used by parties competing for a given constituency).

Boulanger, P-M., Political uses of social indicators: overview and application to sustainable development indicators. *International Journal of Sustainable Development*, 10 (1,2):14-32, 2007.

Critique of composite indicators: Ravallion

There are types two indices: those built on economic theory / monetary aggregates / shadow prices and all others (=mashup indices)



Martin Ravallion

+ existing measures of e.g. development or poverty (Human Development Index, HDI, the Multidimensional Poverty Index, MPI) are bad at coping with tradeoffs

Martin Ravallion, 2010, Mashup indices of development, Policy Research Working Paper 5432 , The World Bank Development Research Group,
<http://documents.worldbank.org/curated/en/454791468329342000/pdf/WPS5432.pdf>

“ To illustrate the distinction, consider two stylized examples of composite indices, both formed from the data on household assets and consumer durables found in the Demographic and Health Surveys (DHS). For index A the variables and their weights are set by the analyst, who has some concept of –economic welfare in mind, and thinks this is related to certain variables in the DHS, which are aggregated based on the analyst’s judgments. For index B, the variables and weights are instead based on a regression model calibrated to another survey data set for which a comprehensive measure of consumption (though still containing measurement errors) could be derived. The model is calibrated to common variables in the expenditure survey and the DHS, and the regression model is used to predict wealth in the DHS. A is a mashup index, B is not.”

(M. Ravallion)



Martin Ravallion

Relevance of all this to composite indicators? Common critiques of CI include:

- Composite indicators as ‘mashup indices’, Ravallion (2010): if only ‘arbitrary weights’ could be replaced with ‘exact’ (shadow) prices ...
- There can be as many indices of sustainability as there are normative definitions of what we want to sustain, CMEPSP, (2009)...
- All too easy to manipulate!

Ravallion M. (2010). “Mashup Indices of Development”. Policy Research Working Paper 5432, The World Bank.

CMEPSP (2009). Commission on the Measurement of Economic Performance and Social Progress, URL: http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf last accessed June 2014.

Human Development Index (HDI): Developed by the United Nations Development Programme (UNDP), the HDI combines indicators of life expectancy, education, and per capita income to assess human development.
<https://hdr.undp.org/data-center/human-development-index>

Environmental Performance Index (EPI): Produced by the Yale Center for Environmental Law and Policy and the Center for International Earth Science Information Network, EPI evaluates environmental performance based on various environmental indicators, including air quality, water resources, and biodiversity.
<https://epi.yale.edu/>

Social Progress Index (SPI): The SPI measures social and environmental performance, covering factors such as basic human needs, foundations of well-being, and opportunity.
<https://www.socialprogress.org/global-index-2022-results/>

Good Country Index: This index evaluates countries based on their contributions to global well-being, including factors related to science and technology, culture, and international peace and security.
<https://www.goodcountry.org/>

Genuine Progress Indicator (GPI): The GPI takes into account economic, social, and environmental factors to provide an alternative to traditional GDP for measuring well-being and sustainability.
<https://dnr.maryland.gov/mdgpi/Pages/what-is-the-GPI.aspx>

Bhutan's Gross National Happiness Index <https://www.grossnationalhappiness.com/>

...

Exercise

Split in groups of five and answer the question:

If I were to use a CI to decide where (in what country) to grow my children what dimensions should the index include

The number of dimensions should not be greater than seven

Appoint a rapporteur to relate in class

Time: 15m



To some observers CI are soft just because a set of pure and hard data corresponding to objective 'facts' has been contaminated by subjective weights.

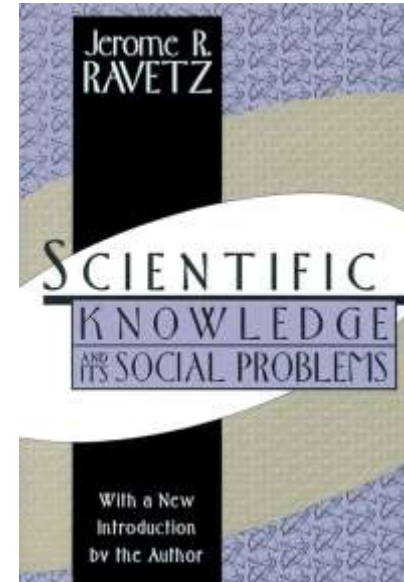
Whatever the shortcomings of CI (and there are many) the neutrality of data as image of facts needs to be questioned.

“Objects of scientific knowledge are intellectually constructed classes rather than ‘things themselves’ ”, Ravetz, 1971, Chapter 4.

“No set of data can be ‘perfect’ as a report of properties of the objects of investigation, nor can it be independent of the plans and expectations for the later stages of the work”, Ravetz, 1971, p. 81.



Jerome R. Ravetz



Ravetz, J., 1971, Scientific Knowledge and its Social Problems, Oxford University Press.

Composite indicators are cherished by the media and the public ... an example where demographic and GDP data are given in the same breath with composite indicators ...

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?

Economist October 16, 2008



More critiques

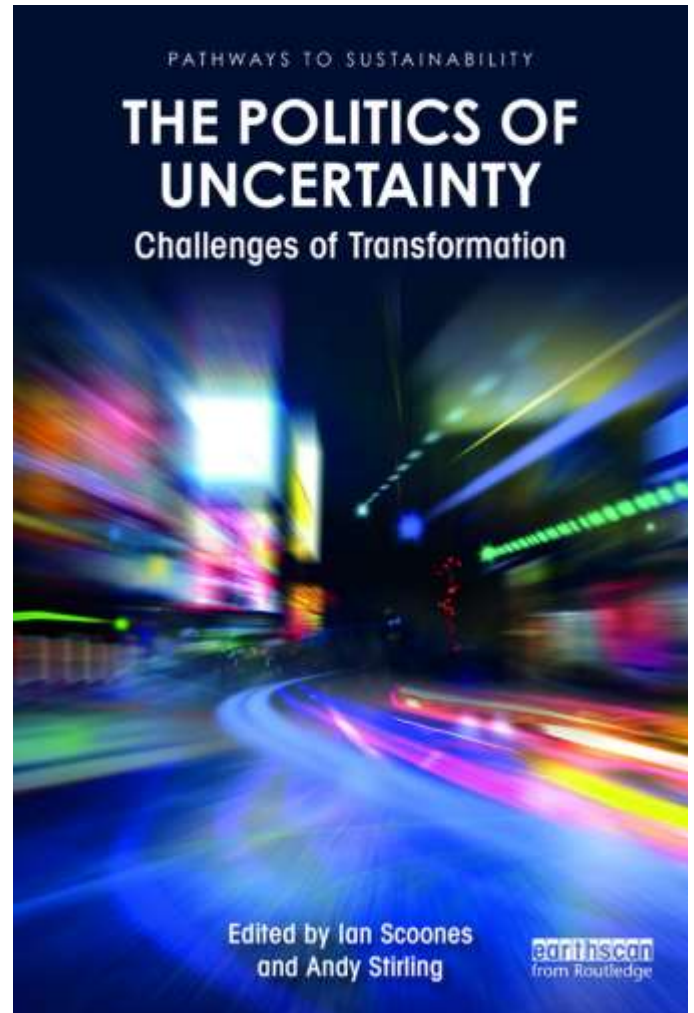
A recent critique of reductionism

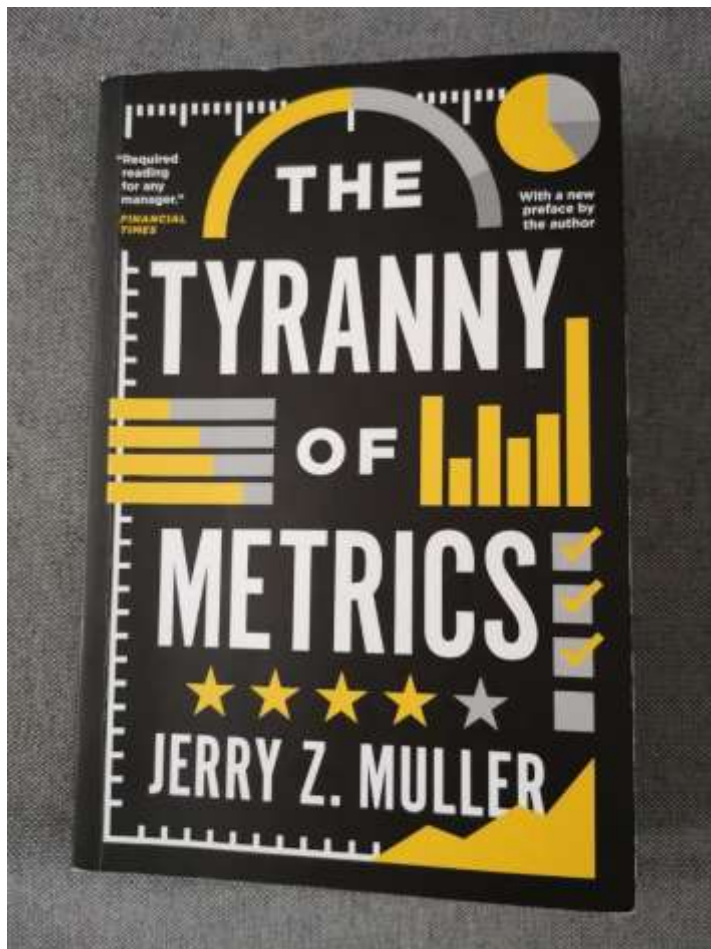
4

THE UNRAVELLING OF TECHNOCRATIC ORTHODOXY?

Contemporary knowledge politics
in technology regulation

Patrick van Zwanenberg

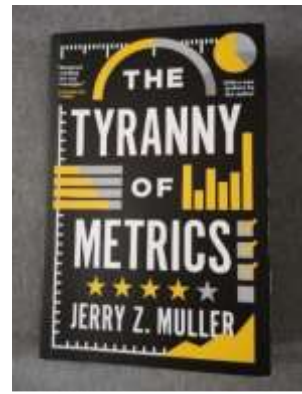




Can composite indicators do harm?

J. Z. Muller, *The tyranny of metrics*.
Princeton University Press , 2018.

Unintended consequences of too many metrics

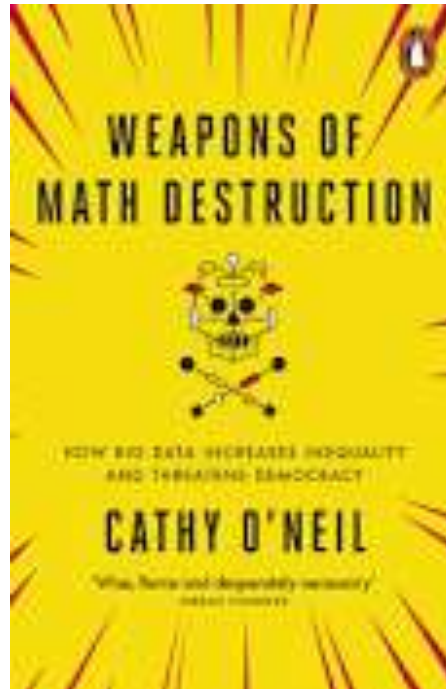


- Goal displacement
- Short termism
- Diminishing utility
- Rule cascade
- Discouraging risk taking
- Discouraging innovation
- Rewarding luck
- Discouraging cooperation and common purpose
- Degrading work
- Time waste
- Loss of productivity

Weapons of math destruction: opaque, do harm, do scale



Cathy O'Neil



Since composite indicators
are here to stay, how can
we make them defensible?

... or how can we
deconstruct them?

Tools for evidence appraisal such sensitivity analysis and sensitivity auditing can be useful to gauge (and possibly deconstruct **or reinforce**) these measures

Sensitivity analysis



PDF



Info

Series A Statistics in Society

[Explore this journal >](#)

Uncertainty and sensitivity analysis techniques as tools for the quality assessment of composite indicators

[M. Saisana](#), [A. Saltelli](#), [S. Tarantola](#)

First published: 3 March 2005 [Full publication history](#)

DOI: [10.1111/j.1467-985X.2005.00350.x](https://doi.org/10.1111/j.1467-985X.2005.00350.x) [View/save citation](#)

Cited by (CrossRef): 181 articles [↻ Check for updates](#)

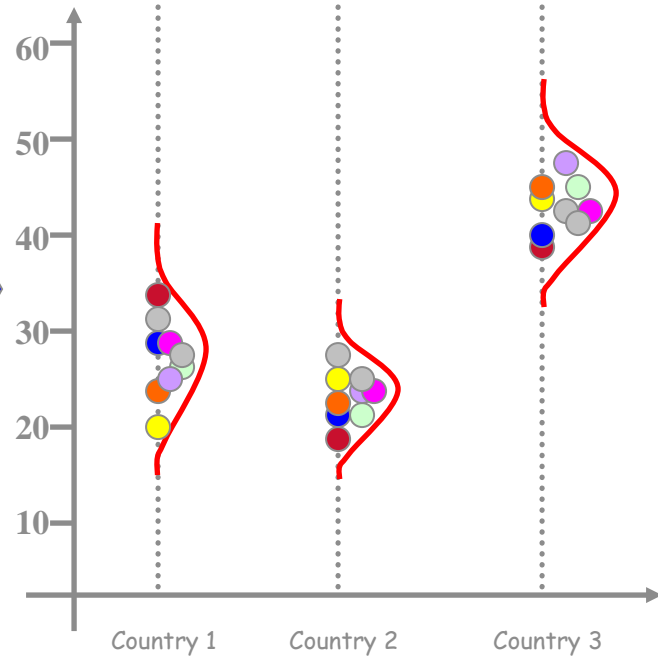
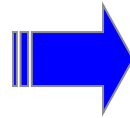
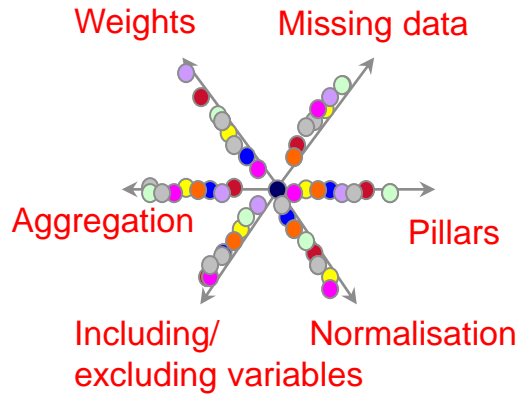
Citation tools ▾



[View issue TOC](#)
Volume 168, Issue 2
March 2005
Pages 307–323

Assumption	Alternatives
Number of indicators	<ul style="list-style-type: none">▪ all six indicators included or one-at-time excluded (6 options)
Weighting method	<ul style="list-style-type: none">▪ original set of weights,▪ factor analysis,▪ equal weighting,▪ data envelopment analysis
Aggregation rule	<ul style="list-style-type: none">▪ additive,▪ multiplicative,▪ Borda multi-criterion

Space of alternatives



Sensitivity analysis to compare volatility of ranking

Research Policy 40 (2011) 165–177



ELSEVIER

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol



Rickety numbers: Volatility of university rankings and policy implications

Michaela Saisana*, Béatrice d'Hombres, Andrea Saltelli

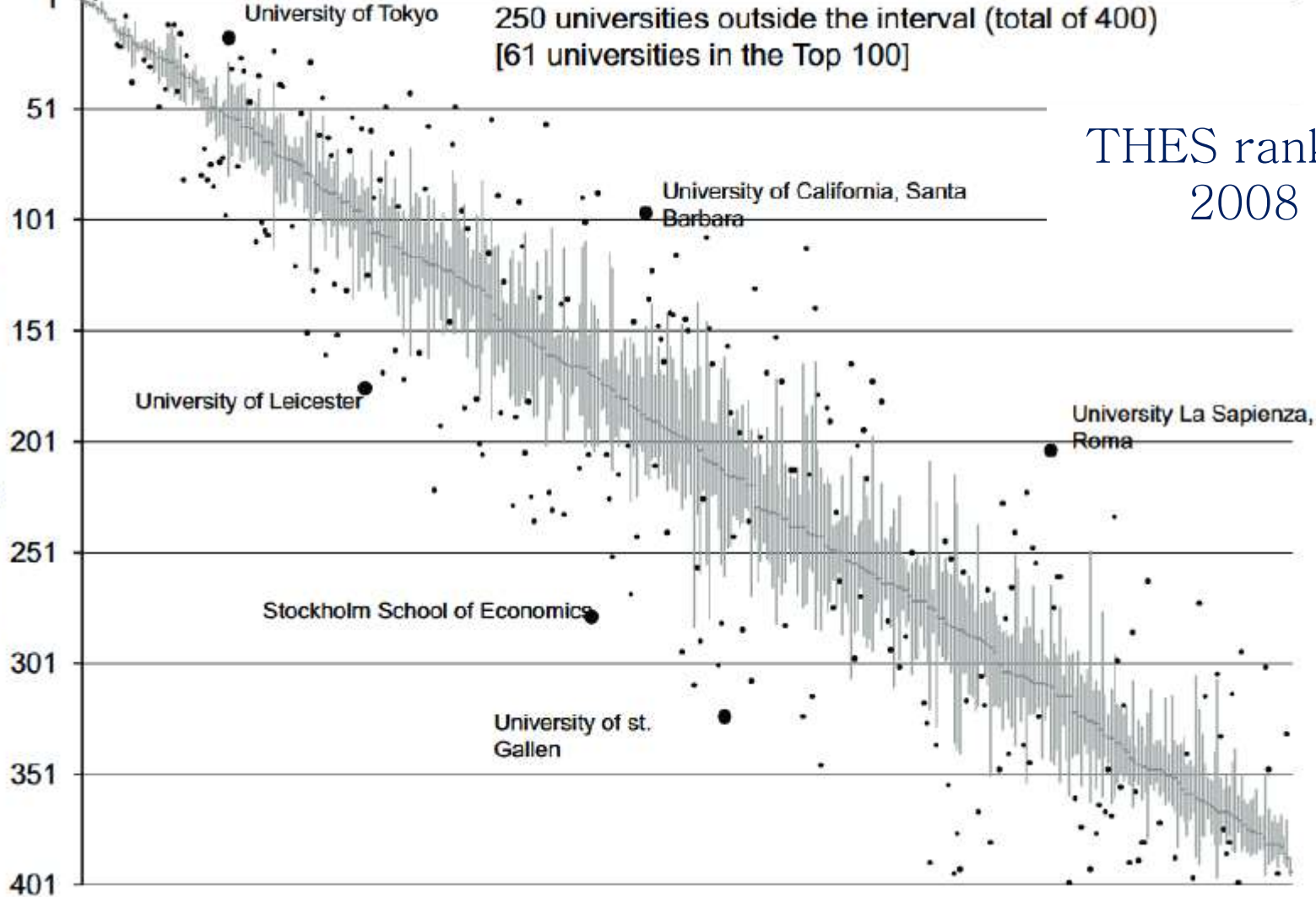
Econometrics and Applied Statistics, Joint Research Centre, European Commission, Enrico Fermi 2749, 21027 Ispra, Italy

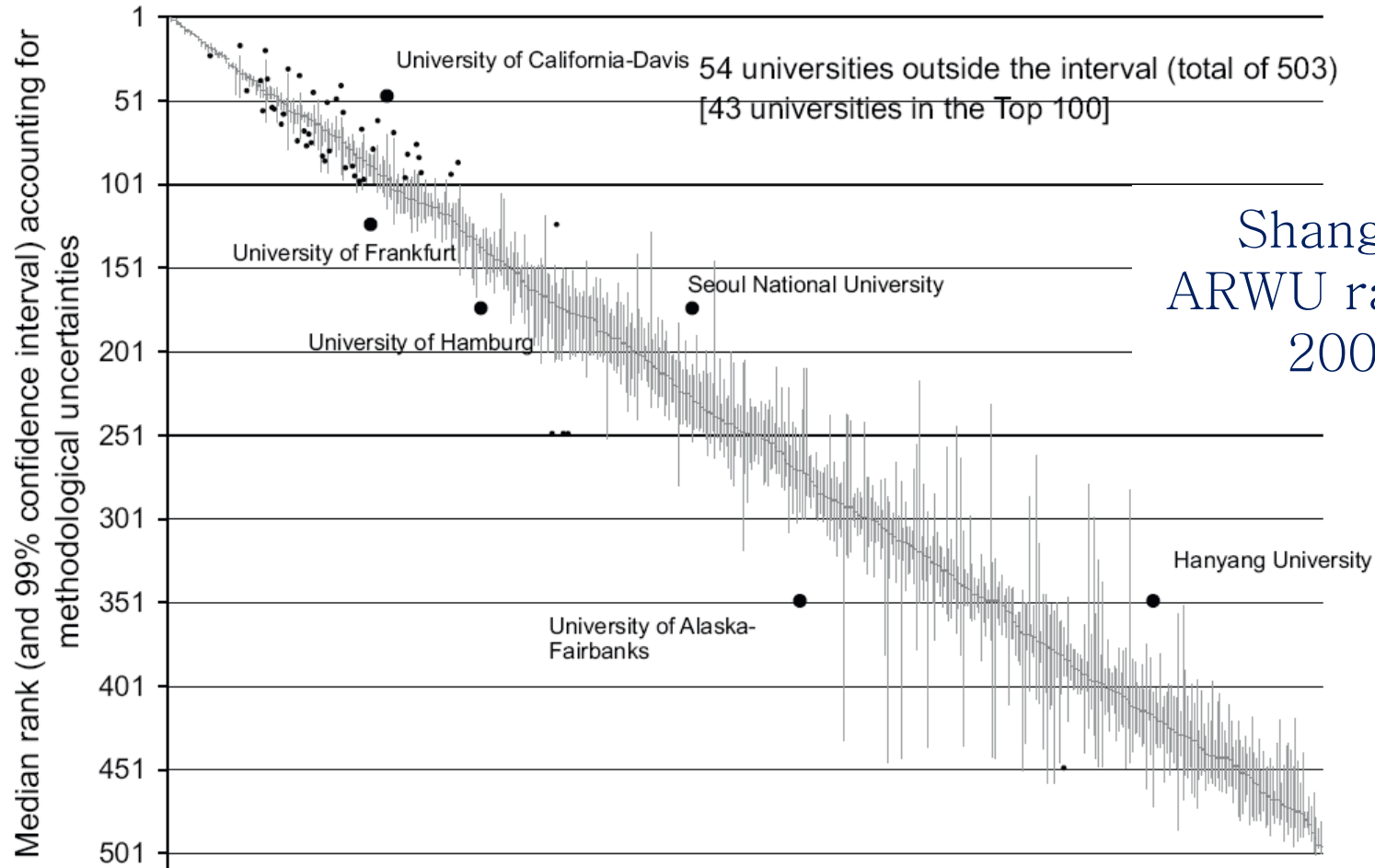
Sensitivity analysis to compare volatility of ranking



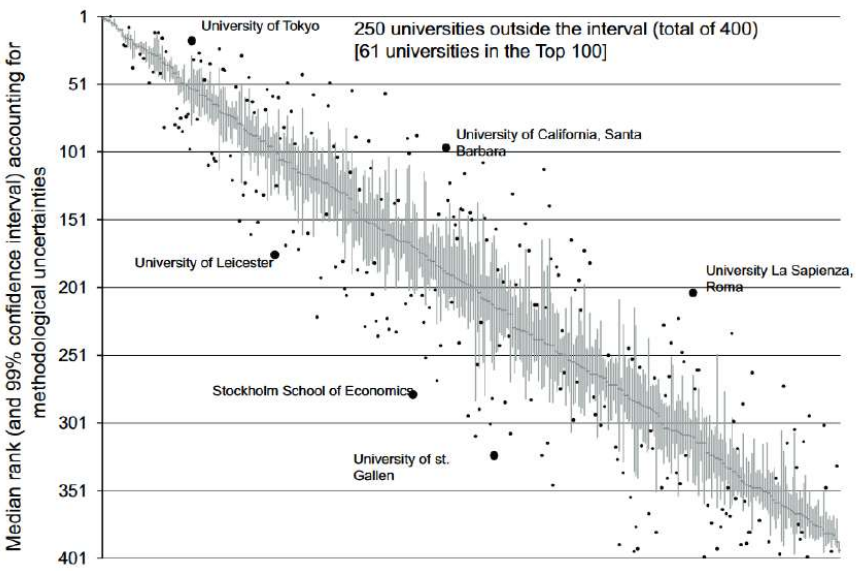
An 'invasive' analysis as the developers' choices are questioned/varied

Median rank (and 99% confidence interval) accounting for methodological uncertainties

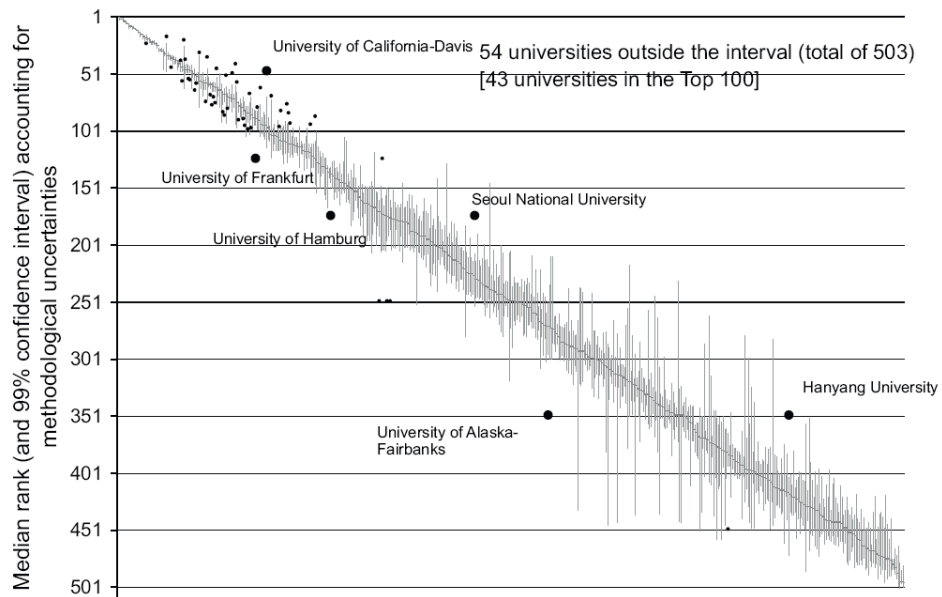




Shanghai
ARWU ranking
2008



THES ranking
2008



Shanghai
ARWU ranking
2008

Incidentally: these university rankings have also damaged the educational systems

« processus de Bologne (en 1999) + stratégie de Lisbonne (en 2000), → passage d'une logique de service public à une logique de marché, concurrentielle et gestionnaire »



Le classement de Shanghai. Histoire, analyse et critique

Fabien Eloire

DANS **L'HOMME & LA SOCIÉTÉ** 2010/4 (n° 178), PAGES 17 À 38

Smashing the glasshouse. Diminishing the prestige of measures of higher education

AUTHORS

Andrea Saltelli, Alexandra Theben, [Rosario Scandurra](#)

AUTHOR ASSERTIONS

Conflict of Interest: No ▼

Public Data: No ▼

Preregistration: No ▼

Abstract: According to many authors the ranking of institutes and universities of higher education (HE) has provoked dramatic consequences in terms of transforming these institutions into a global market, sending the university prices skyrocketing and generally modifying the landscape of HE...

Smashing the glasshouse. Diminishing the prestige of measures of higher education

AUTHORS

Andrea Saltelli, Alexandra Theben, [Rosario Scandurra](#)

AUTHOR ASSERTIONS

Conflict of Interest: No ▾

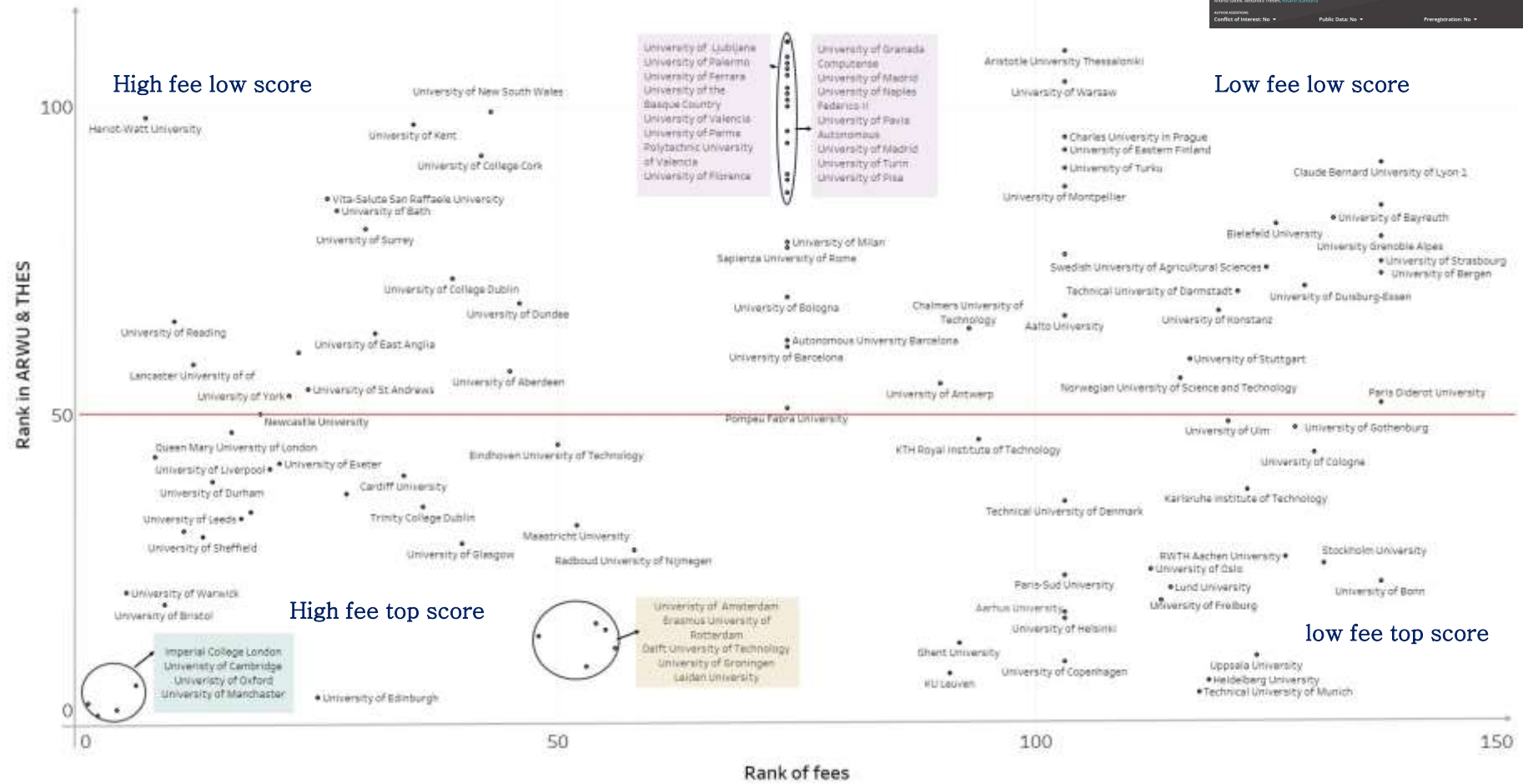
Public Data: No ▾

Preregistration: No ▾

If costs were included the panorama of HE in terms of desirability for students and their family would change considerably.

We review what scholars think of HE rankings.

We use these elements to build a case against rankings of HEIs

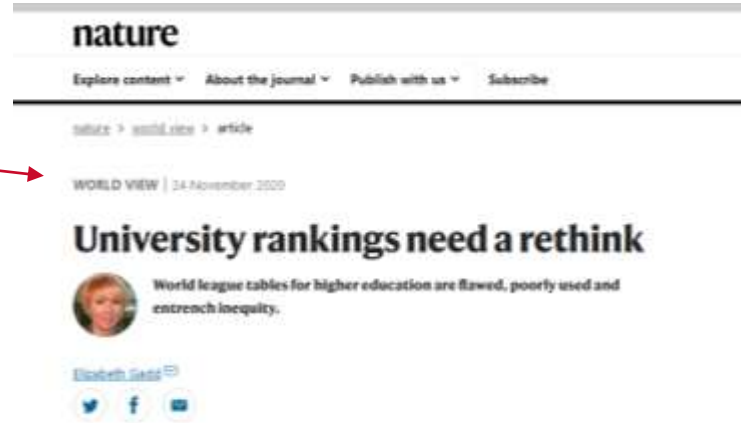
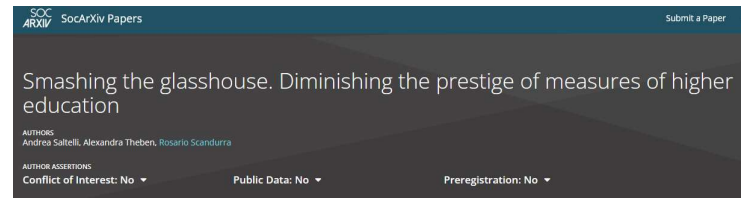


Conclusions

“The University of Saskatchewan is hiring a University Ranking Strategist at an annual salary of \$81K-135K. Reporting into their TWO rankings groups” (Gadd 2020)

Recruiting an expert to game the system undoubtedly represents a new frontier of university behaviour ... “What next?”, and “How did we get there?” appear as reasonable questions.

A cartel sees authors or journals teaming up to cite one another strategically in an effort to improve their metrics.



One can test whether assigned weights correspond to real importance

Journal of the
Royal Statistical Society

SERIES A
Statistics
in Society



J. R. Statist. Soc. A (2013)
176, Part 3, pp. 609–634

Ratings and rankings: voodoo or science?

Paolo Paruolo

University of Insubria, Varese, Italy

and Michaela Saisana and Andrea Saltelli

European Commission, Ispra, Italy

Linear aggregation paradox: **weights are used as if they were importance coefficients** while they are trade off coefficients

An example. A dean wants to rank teachers based on ‘hours of teaching’ and ‘number of publications’ ...

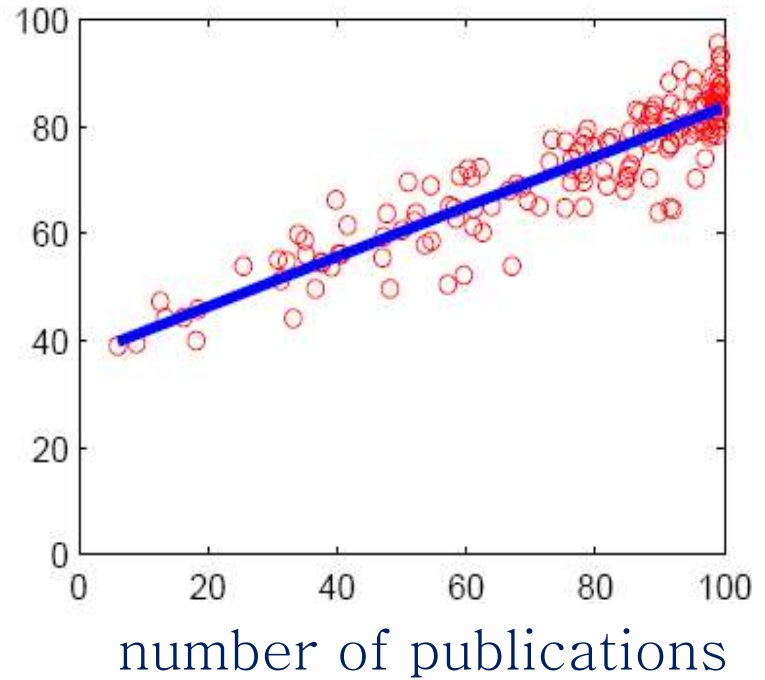
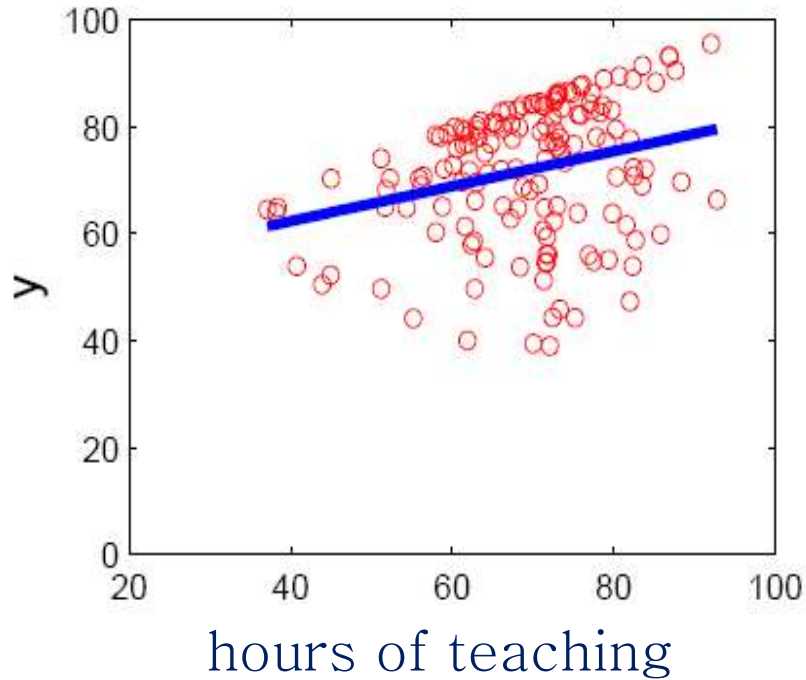


$$Y = 0.5X_1 + 0.5X_2$$

X_1 : hours of teaching

X_2 : number of publications

... adding these two variables up she sees that teachers are practically ranked by publications alone



Dean's example: $y = x_1 + x_2$.

Estimated $R_{ht}^2 = 0.0759$, $R_{np}^2 = 0.826$

To obviate this the dean substitutes the model

$$y=0.5x_1+ 0.5x_2$$

with

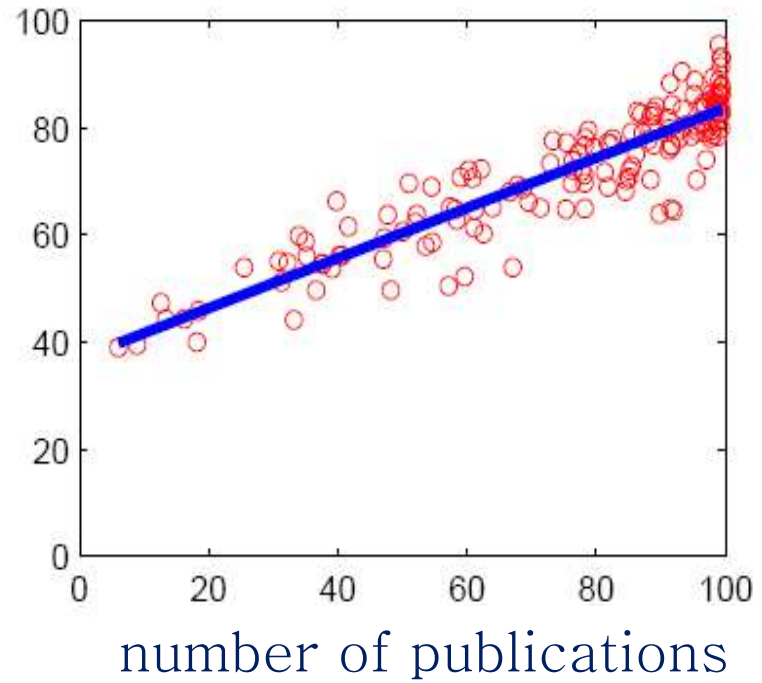
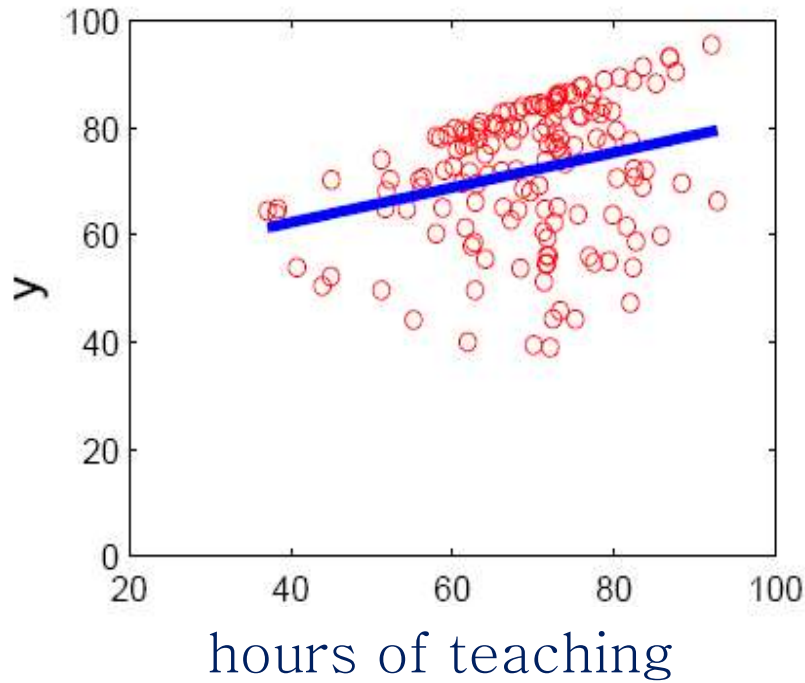
$$y=0.7x_1+ 0.3x_2$$

X_1 : hours of teaching

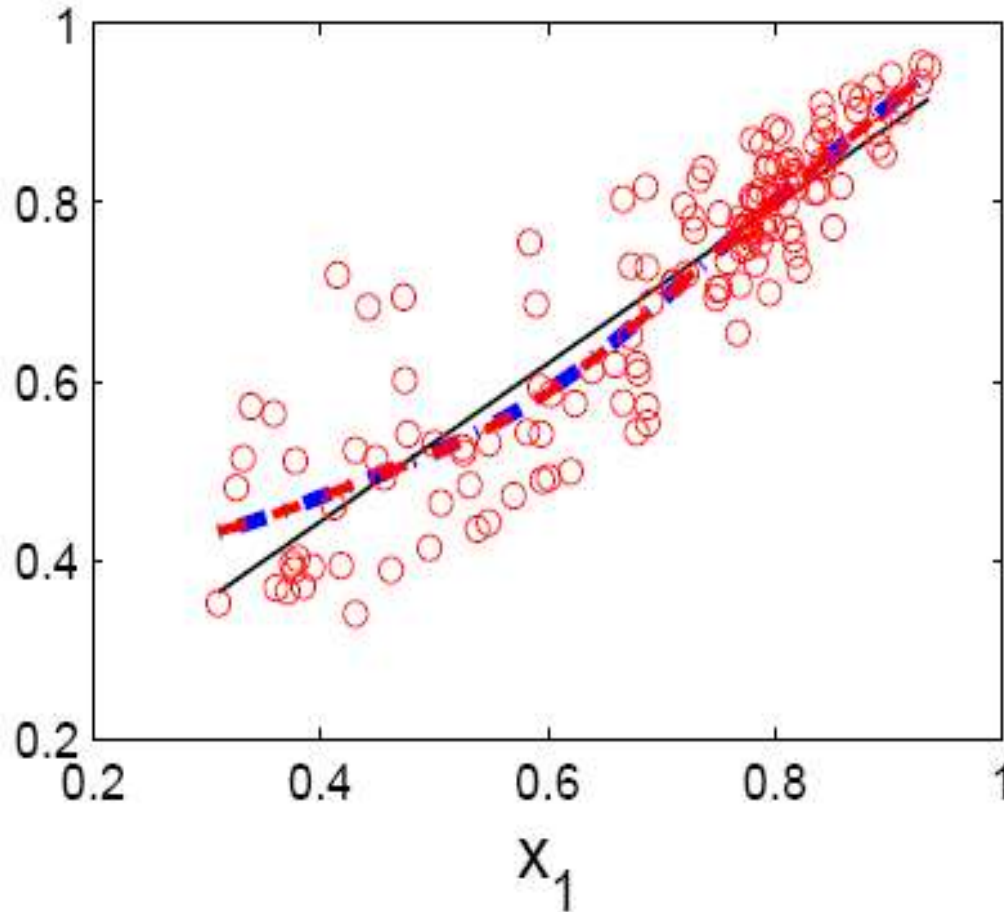
X_2 : number of publication

to rebalance hours of teaching.

A professor comes by, looks at the last formula, and complains that publishing is disregarded in the department ...

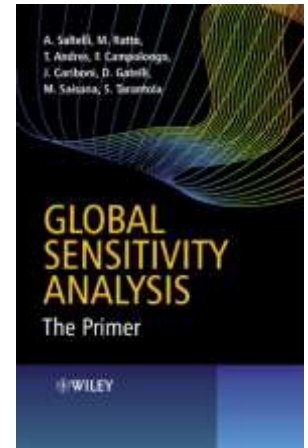


Using a result from sensitivity analysis the scatterplots → numbers reflecting the importance of a variable



The straight line corresponds to R^2

The variance of the moving average is a sensitivity measure



Another paradox

Normalization paradox: **weights are assigned as to add up to one.** This is questionable.

Given a simple CI $Y = w_1x_1 + w_2x_2$

If both x_1 and x_2 are standardized the

importance of x_1 is $S_1 = \frac{w_1^2}{w_1^2 + w_2^2}$

and $S_1 + S_2 = 1$ based on the theory of variance based sensitivity analysis

Thus the relative importance of x_1, x_2

is not $\frac{w_1}{w_2}$ but $\frac{w_1^2}{w_2^2} \dots$

... and the absolute importance are not $\frac{w_1}{w_1+w_2}$

and $\frac{w_2}{w_1+w_2}$

but $\frac{w_1^2}{w_1^2+w_2^2}$ and $\frac{w_2^2}{w_1^2+w_2^2}$

Implications?

	x_1	x_2
Presumed importance	10%	90%
Real importance	1.2%	98.8%

$$\frac{w_1}{w_1 + w_2}, \frac{w_2}{w_1 + w_2}$$

$$\frac{w_1^2}{w_1^2 + w_2^2}, \frac{w_2^2}{w_1^2 + w_2^2}$$

This holds if we use our definition of importance (what expected fraction of the variance of Y would be reduced on average if x_1 could be fixed) – what a sensitivity measure means if translated in plain English

Comparing assigned weights versus measured importance for the 2009 and 2010 versions of the Human Development index

Journal of the
Royal Statistical Society

SERIES A
Statistics
in Society



J. R. Statist. Soc. A (2013)
176, Part 3, pp. 609–634

Ratings and rankings: voodoo or science?

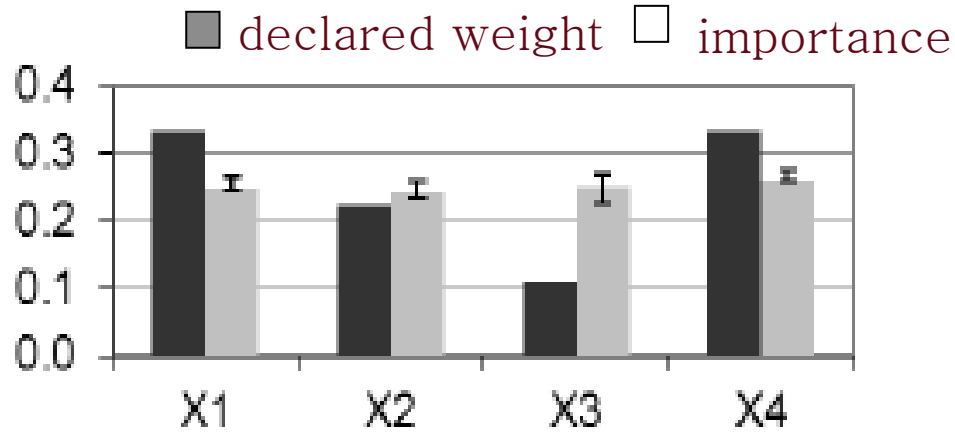
Paolo Paruolo

University of Insubria, Varese, Italy

and Michaela Saisana and Andrea Saltelli

European Commission, Ispra, Italy

HDI
2009



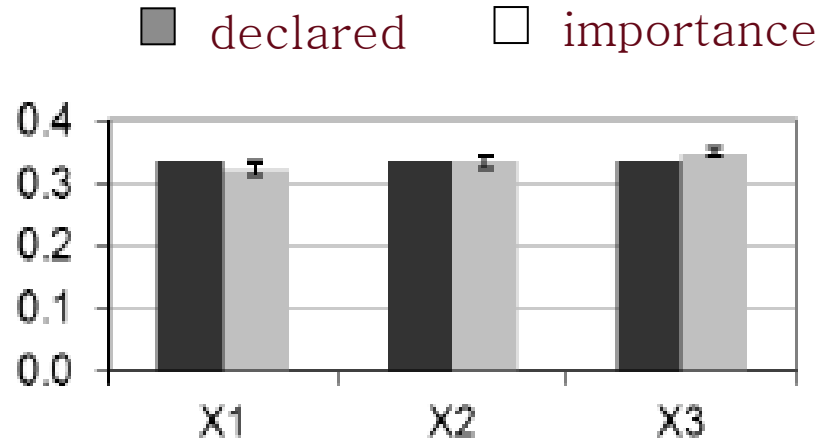
Life expectancy, 33%

Adult literacy, 22%

Enrollment education, 11%

GDP per capita, 33%

HDI
2010



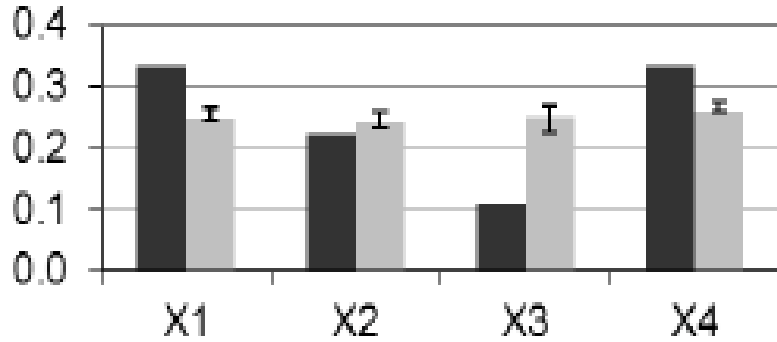
Life expectancy, 33%

Education, 33%

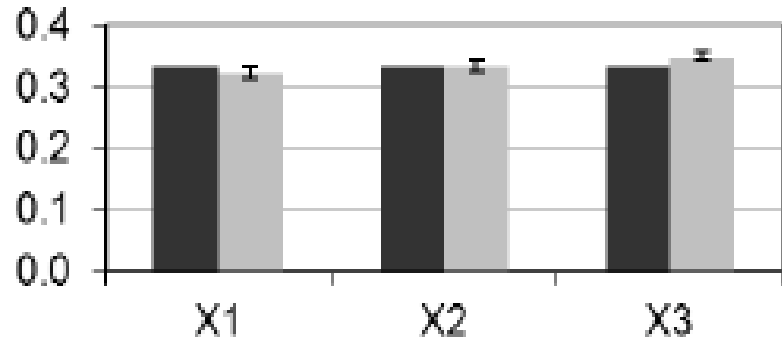
GNI per capita, 33%

■ declared weight □ importance

HDI2009



HDI2010



HDI 2010 more coherent than HDI 2009

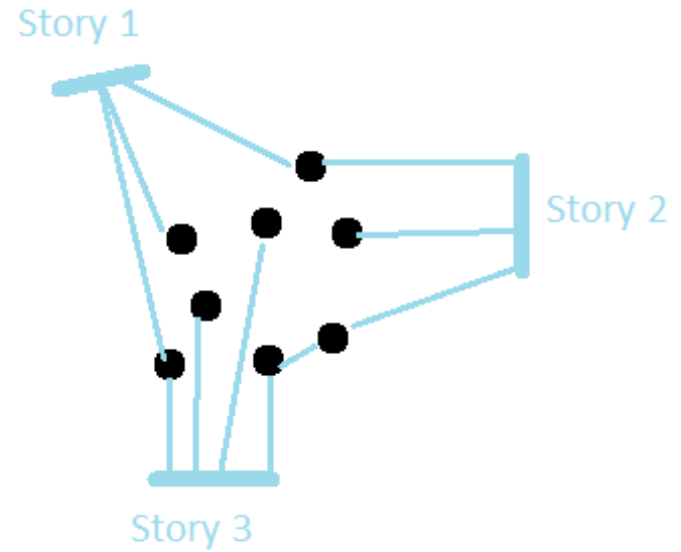
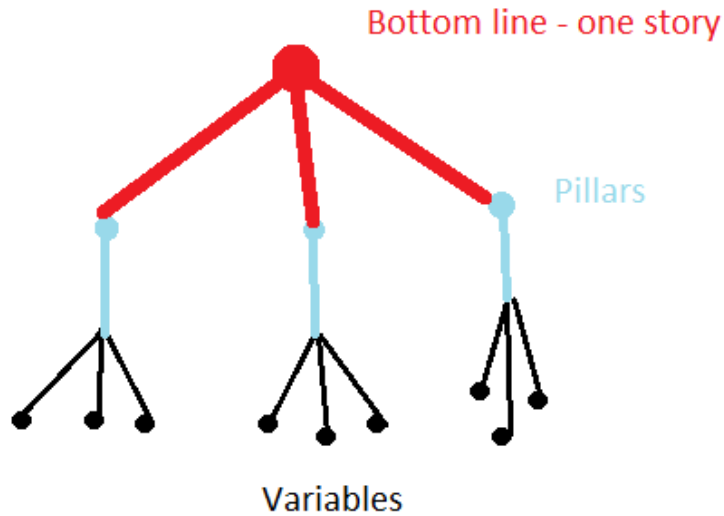
What if different stakeholders have different preferences? A test case of EU convergence analysis;

Four different viewpoints are compared

Kuc-Czarnecka, M., Lo Piano, S. and Saltelli, A. (2020) 'Quantitative storytelling in the making of a composite indicator', *Social Indicators Research*, accepted,
http://www.andreasaltelli.eu/file/repository/SIR_TEMP.pdf



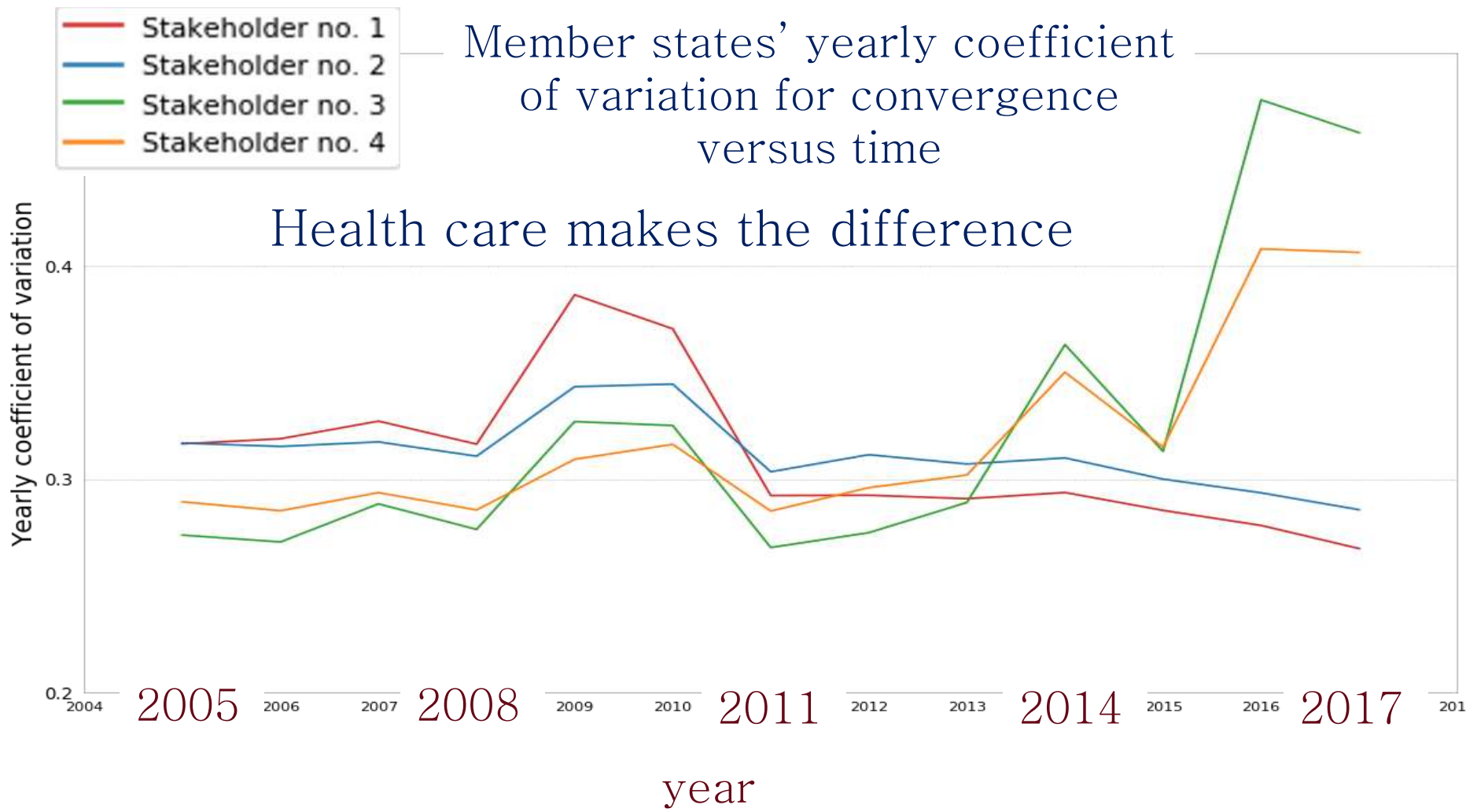
Classical vs. Variable geometry CI



Stakeholder 1	Stakeholder 2	Stakeholder 3	Stakeholder 4
Access to labour market	Access to labour market	Access to labour market	Access to labour market
Fair working conditions	Fair working conditions	Fair working conditions	Fair working conditions
Social protection	Social protection	Social protection	Social protection
	Fairness	Health care	Fairness
			Health care

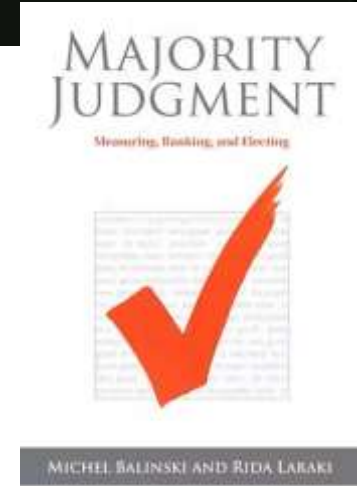
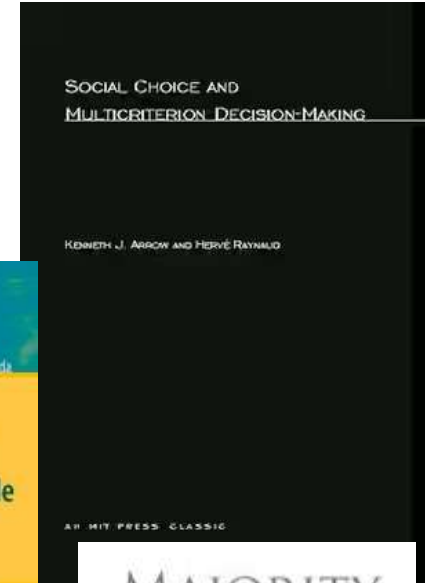
Member states' yearly coefficient of variation for convergence versus time

Health care makes the difference



Better ways to build CI: multicriteria analysis

Building CI using using methods such as Borda, Condorcet, Balinski-Laraki ...





Some of these methods have a long history
(including in Catalonia)



Ramon Llull (Catalan, ca. 1232 – ca. 1315) proposed first what would then become known as the method of Condorcet. **Nicholas of Kues** (1401 – August 11, 1464), also referred to as Nicolaus Cusanus and Nicholas of Cusa developed what would later be known as the method of Borda. **Nicolas de Condorcet**, (17 September 1743 – 28 March 1794) developed the eponymous method. **Jean-Charles, chevalier de Borda** (May 4, 1733 – February 19, 1799) developed the Borda count

Images from Wikipedia Commons

An impact matrix

	Indic.	GDP	Unemp. Rate	Solid wastes	Income dispar.	Crime rate
Country						
A		25,000	0.15	0.4	9.2	40
B		45,000	0.10	0.7	13.2	52
C		20,000	0.08	0.35	5.3	80
weights		.166	.166	0.333	.166	.166

We can say that

GDP 'votes' for B>A>C (countries / options)

UR 'votes' for C>B>A

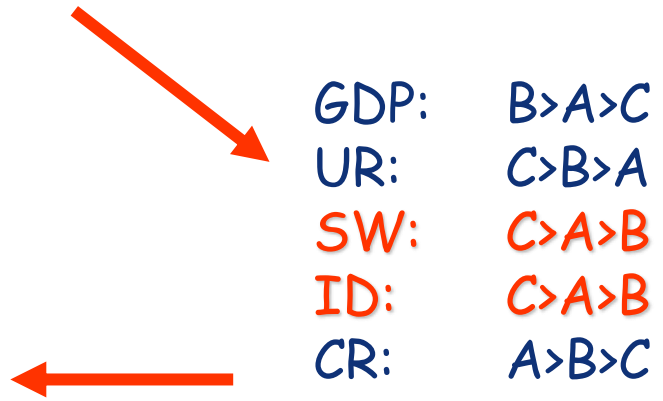
SW 'votes' for C>A>B

ID 'votes' for C>A>B

CR 'votes' for A>B>C

	Indic.	GDP	Unemp. Rate	Solid wastes	Income dispar.	Crime rate
Country						
A		25,000	0.15	0.4	9.2	40
B		45,000	0.10	0.7	13.2	52
C		20,000	0.08	0.35	5.3	80
weights		.166	.166	0.333	.166	.166

# of indicators	2	1	1	1
1st position	<i>c</i>	<i>b</i>	<i>c</i>	<i>a</i>
2nd position	<i>a</i>	<i>a</i>	<i>b</i>	<i>b</i>
3rd position	<i>b</i>	<i>c</i>	<i>a</i>	<i>c</i>



# of indicators	2	1	1	1
1st position	<i>c</i>	<i>b</i>	<i>c</i>	<i>a</i>
2nd position	<i>a</i>	<i>a</i>	<i>b</i>	<i>b</i>
3rd position	<i>b</i>	<i>c</i>	<i>a</i>	<i>c</i>



Rank	<i>a</i>	<i>b</i>	<i>c</i>
1st	1	1	3
2nd	3	2	0
3rd	1	2	2

Different ways to organize the same information: building a frequency matrix

Three countries [options/candidates] and five indicators [criteria/voters]

# of indicators	2	1	1	1
1st position	<i>c</i>	<i>b</i>	<i>c</i>	<i>a</i>
2nd position	<i>a</i>	<i>a</i>	<i>b</i>	<i>b</i>
3rd position	<i>b</i>	<i>c</i>	<i>a</i>	<i>c</i>



Rank	<i>a</i>	<i>b</i>	<i>c</i>
1st	1	1	3
2nd	3	2	0
3rd	1	2	2

For a case with three candidates Borda gives 3 minus 1 for each first rank , 2 minus 1 for each second rank and zero to the third; so

a gets $2*1+ 1*3=5$

b gets $2*1+ 1*2=4$

c gets $2*3+ 1*0=6$

But lets try Borda on a more interesting case: (from Moulin, 21 criteria 4 options, cited in Munda 2008)



21 criteria 4 alternatives

Note: $3+5+7+6=21$

# of indicators	3	5	7	6
1st position	<i>a</i>	<i>a</i>	<i>b</i>	<i>c</i>
2nd position	<i>b</i>	<i>c</i>	<i>d</i>	<i>b</i>
3rd position	<i>c</i>	<i>b</i>	<i>c</i>	<i>d</i>
4th position	<i>d</i>	<i>d</i>	<i>a</i>	<i>a</i>



Rank	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	Points
1st	8	7	6	0	3
2nd	0	9	5	7	2
3rd	0	5	10	6	1
4th	13	0	0	8	0

Borda count – Frequency matrix (Moulin, 21 criteria 4 options)

Columns add up to the
number of criteria /
voters=21

3 points if first
2 if second
1 if third
0 if last

Rank	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	Points
1st	8	7	6	0	3
2nd	0	9	5	7	2
3rd	0	5	10	6	1
4th	13	0	0	8	0

Borda score:

$$a = 8 \times 3 = 24$$

$$b = 5 + 9 \times 2 + 7 \times 3 = 44$$

$$c = 10 + 5 \times 2 + 6 \times 3 = 38$$

$$d = 6 + 7 \times 2 = 20$$

Borda solution:

$b \rightarrow c \rightarrow a \rightarrow d$

Frequency matrix
(21 criteria 4
alternatives)

Rank	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	Points
1st	8	7	6	0	3
2nd	0	9	5	7	2
3rd	0	5	10	6	1
4th	13	0	0	8	0

The Borda count was developed independently several times, (e.g. by Nicolaus Cusanus beginning XV century) but is named for **Jean-Charles de Borda**, who devised the system in 1770.

It is currently used for the election of two ethnic minority members of the National Assembly of Slovenia

(<https://www.electoral-reform.org.uk/how-do-elections-work-in-slovenia/>)

It is used throughout the world by various organisations and competitions [e.g. in academia]



Jean-Charles,
chevalier de
Borda

Borda was a mariner and a scientist. Worked on chronometers. Between 1777 and 1778, he participated in the American Revolutionary War.

The French Academy of Sciences used Borda's method to elect its members for about two decades [till Napoleon Bonaparte became president...]



Condorcet disagrees ...



Condorcet's outsourcing matrix (21 criteria 4 alternatives)

# of indicators	3	5	7	6
1st position	<i>a</i>	<i>a</i>	<i>b</i>	<i>c</i>
2nd position	<i>b</i>	<i>c</i>	<i>d</i>	<i>b</i>
3rd position	<i>c</i>	<i>b</i>	<i>c</i>	<i>d</i>
4th position	<i>d</i>	<i>d</i>	<i>a</i>	<i>a</i>

Frequency matrix

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
<i>a</i>	0	8	8	8
<i>b</i>	13	0	10	21
<i>c</i>	13	11	0	14
<i>d</i>	13	0	7	0

Outscoring matrix

b better
than a
 $7+6=13$
times

b better
than c
 $7+3=10$
times

How to move from frequency to outscoring (again)

# of indicators	3	5	7	6
1st position	<i>a</i>	<i>a</i>	<i>b</i>	<i>c</i>
2nd position	<i>b</i>	<i>c</i>	<i>d</i>	<i>b</i>
3rd position	<i>c</i>	<i>b</i>	<i>c</i>	<i>d</i>
4th position	<i>d</i>	<i>d</i>	<i>a</i>	<i>a</i>

Frequency matrix

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
<i>a</i>	0	8	8	8
<i>b</i>	13	0	10	21
<i>c</i>	13	11	0	14
<i>d</i>	13	0	7	0

Outscoring matrix

Condorcet's outscoring matrix (21 criteria 4 alternatives)

For each pair of countries a concordance index is computed by counting how many indicators/voters are in favour of each country (e.g. 13 voters prefer b to a).

Note the “constant sum property” in the outranking matrix (13+ 8=21 number of indicators/voters)

$$\begin{bmatrix} & a & b & c & d \\ a & 0 & 8 & 8 & 8 \\ b & 13 & 0 & 10 & 21 \\ c & 13 & 11 & 0 & 14 \\ d & 13 & 0 & 7 & 0 \end{bmatrix}$$

Outranking matrix

How to use Condorcet's outscoring matrix (21 criteria 4 alternatives)

Pairs with concordance index $> 50\%$ of the indicators/voters are considered: majority threshold = 11 (i.e. a number of voters $> 50\%$ of voters=21)

Thus a is never better than anyone more than 11 times, while $bPa=13$, $bPd=21$ (=always), $cPa=13$, $cPb=11$, $cPd=14$, $dPa=13$.

c is better than a,b,d so it is the winner

b is better than the remaining a,d, it is the second best

d is better than a.

→ Condorcet solution: $c \rightarrow b \rightarrow d \rightarrow a$

	a	b	c	d
a	0	8	8	8
b	13	0	10	21
c	13	11	0	14
d	13	0	7	0

Count row-wise discarding entries < 11 as there are 21 voters/criteria

Borda solution: $b \rightarrow c \rightarrow a \rightarrow d$

Condorcet solution: $c \rightarrow b \rightarrow d \rightarrow a$

Can we choose between Borda and Condorcet on some theoretical and/or practical grounds?



Exercise

Same groups as before:

Elect among yourself a rapporteur (not the same as before) using Borda count as well as Majority Rule (the one with the max votes wins)

10m

Rapporteur to report in class



Deconstructing the implicit normative framing of an indicator

What do I make of your latinorum? Sensitivity auditing of mathematical modelling

Andrea Saltelli* and
Ângela Guimarães Pereira

European Commission,
Joint Research Centre,
Institute for the Protection and Security of the Citizen,
Via E. Fermi, 2749, 21027 Ispra (VA), Italy
E-mail: andrea.saltelli@jrc.ec.europa.eu
E-mail: angela.pereira@jrc.ec.europa.eu
*Corresponding author

Jeroen P. Van der Sluijs

Copernicus Institute of Sustainable Development,
Utrecht University (NL),
Heidelberglaan 2, 3584 CS Utrecht, The Netherlands
E-mail: j.p.vandersluijs@uu.nl

Silvio Funtowicz

Centre for the Study of the Sciences and the Humanities (SVT)
Allegaten, University of Bergen (NO),
34 – Postboks 7805 5020 Bergen, Norway
E-mail: silvio.funtowicz@svt.uib.no

How about
sensitivity
auditing &
quantitative
story-telling



Contents lists available at [ScienceDirect](#)

Futures

journal homepage: www.elsevier.com/locate/futures



Original research article

What is wrong with evidence based policy, and how can it be improved?

Andrea Saltelli^{a,b,c,*}, Mario Giampietro^{a,c,d}

The issue of frames

Why quantification / framing may obfuscate an issue:

The example of GMO

Andrea Saltelli, Mario Giampietro, 2017, What is wrong with evidence based policy, and how can it be improved? Futures, DOI: <http://dx.doi.org/doi:10.1016/j.futures.2016.11.012>

Rayner, S., 2012, Uncomfortable knowledge: the social construction of ignorance in science and environmental policy discourses, *Economy and Society*, 41:1, 107–125.



GMO as a food scare

The Economist, Vermont v science, The little state that could kneecap the biotech industry, May 10th 2014

Citizens' worries (Marris, 2001, excerpts)



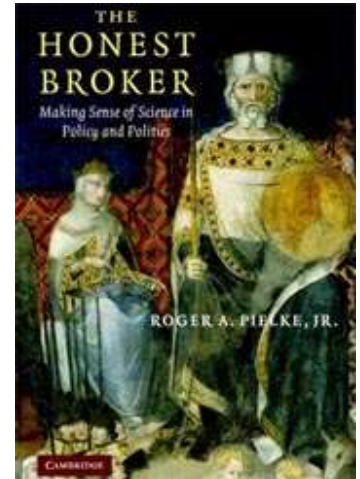
- Who decided that they should be developed and how?
- Why are we not given an effective choice about whether or not to buy and consume these products?
- Do regulatory authorities have sufficient powers and resources to effectively counter-balance large companies who wish to develop these products?

Marris, C., Wynne, B., Simmons P., and Weldon, S. 2001. Final Report of the PABE research project funded by the Commission of European Communities, Contract number: FAIR CT98-3844 (DG12 - SSMI), December 2001.

The problem is not the non-neutrality of a measure but its purported neutrality (R. Pielke's Jr 'stealth advocacy')

(more at the lesson on impact assessment)

Roger A. Pielke, Jr., 2007, *The honest broker*, Cambridge University Press



So what does quantitative story telling propose?

Instead of detailed quantification on a single frame a rough quantitative appraise of a richer set of frames

E.g. not just one convergence story

Test the various stories for:

- feasibility (compatibility with processes outside human control);
- viability (compatibility with processes under human control, in relation to both the economic and technical dimensions); and
- desirability (compatibility with a multitude of normative considerations relevant to a plurality of actors).

Implications for CI's

Awareness of the imperfections and non-neutrality of measures

Investigate properties and assumptions

Use for social discovery, (deliberative) extended participation; quality as fitness for purpose

Don't commit to a single view

Too much is being read in the OECD–PISA data

IJCED
19,1

Do PISA data justify PISA-based education policy?

Luisa Araujo

*Department of Human Capital and Employment,
European Commission Joint Research Centre Ispra Sector, Ispra, Italy*

Andrea Saltelli

*University of Bergen, Bergen, Norway and
Universitat Autònoma de Barcelona, Barcelona, Spain, and*

Sylke V. Schnepf

*Competence Centre on Microeconomic Evaluation,
European Commission Joint Research Centre Ispra Sector, Ispra, Italy*



International Journal of
Comparative Education and
Development
Vol. 19 No. 1, 2017
pp. 20-34
© Emerald Publishing Limited
2396-7404
DOI 10.1108/IJCED-12-2016-0023

20

Received 14 December 2016
Revised 17 February 2017
Accepted 24 February 2017



International PISA tests show how evidence-based policy can go wrong

June 12, 2017 3:55pm AEST

Taking issue with:

“If every EU Member State achieved an improvement of 25 points in its PISA score [...] the GDP of the whole EU would increase by between 4% and 6% by 2090; such an 6% increase would correspond to 35 trillion Euro” (Woessmann, 2014)

Advocating for a more collegial construction of the measure

Woessmann, L. (2014), “The economic case for education”, EENEE Analytical Report 20, European Expert Network on Economics of Education (EENEE), Institute and University of Munich, Munich

See also ‘OECD and Pisa tests are damaging’, The Guardian, 6 May 2014
<https://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics>

The Ecological Footprint; top in advocacy, bottom in quality

Ecological Indicators 46 (2014) 610–621

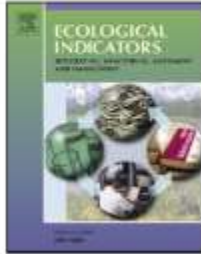


ELSEVIER

Contents lists available at [ScienceDirect](#)

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind



Footprints to nowhere

Mario Giampietro^{a,c}, Andrea Saltelli^{b,*}

^a Institute of Environmental Science and Technology (ICTA), Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

^b Institute for the Protection and Security of the Citizen (IPSC), The European Commission, Joint Research Centre, TP 361, 21027 Ispra, VA, Italy

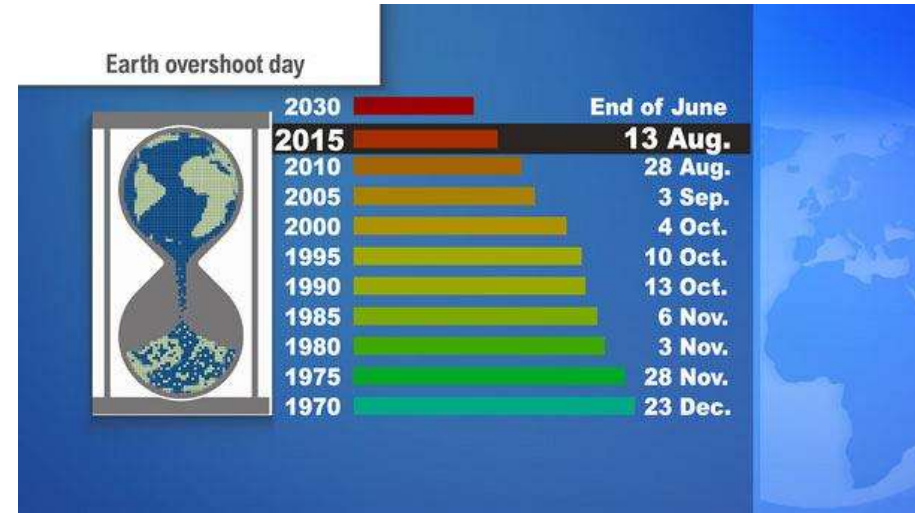
^c Catalan Institution for Research and Advanced Studies (ICREA), Passeig Lluis Companys, 23, 08010 Barcelona, Spain



The Ecological Footprint suggests compressing sustainability to a single metric (acres of equivalent land). Volatile assumptions are taken, spurious precision is generated, and a meaningless ‘interpretant’ is proposed



Interpretant: Paradoxical policy implications – e.g. intensive non sustainable agricultural practices promoted



Alessandro Galli, Mario Giampietro , Steve Goldfinger , et al., 2016, **Questioning the ecological footprint** , *Ecological Indicators*, 69, 224–232.

Giampietro, M., and Saltelli, A., 2014, **Footprints to nowhere**, *Ecological Indicators*, 46, 610–621.

Goldfinger, S., Wackernagel, M., Galli, A., Lazarus, E., Lin, D., 2014, **Footprint facts and fallacies: A response to Giampietro and Saltelli (2014) “Footprints to Nowhere”**, 46, 622–632.

Giampietro, M., and Saltelli, A., 2014, **Footworking in Circles**, *Ecological Indicators*, 46 (2014) 260–263.

Conclusions: CI – instructions for use

Awareness of the imperfections and non-neutrality of measures

Beware damage; mind the interpretant

Investigate properties and assumptions (uncertainty and sensitivity analysis, sensitivity auditing)

Use for social discovery; deliberative extended participation; quality as fitness for purpose (interpretant)

Reading material

Becker, W. et al. (2017) 'Weights and Importance in Composite Indicators: Mind the Gap', in Roger Ghanem, David Higdon, H. O. (ed.) Handbook of Uncertainty Quantification. Springer.

http://www.andreasaltelli.eu/file/repository/Full_Copy_CI_Handbook_2017.pdf

OECD-JRC Handbook on Constructing Composite Indicators: Methodology and User Guide, https://www.oecd-ilibrary.org/economics/handbook-on-constructing-composite-indicators-methodology-and-user-guide_9789264043466-en

Paul-Marie Boulanger, Sustainable development indicators: a scientific challenge, a democratic issue, S.A.P.I.EN.S, Vol.1 / n°1, <https://journals.openedition.org/sapiens/166>

Become a deconstructor

Your assignment





“What follows is a hypothetical executive summary from an imagined Food and Agriculture Organization (FAO) report on the state of the world’s food systems, written from the perspective of the 2050s”

<https://www.thesolutionsjournal.com/article/pathways-leading-sustainable-healthy-global-food-system/>

Executive Summary: FAO State of World Agriculture in 2050 Draft Report

“[...]this FAO report presents evidence that the international food system of the second half of the 21st century is more sustainable than the food system of the late 20th or early 21st centuries.



[...] today more people are being fed on less land and agriculture is requiring fewer inputs”

Executive Summary: FAO State of World Agriculture in 2050 Draft Report

“[...] despite there being 10 billion people on the planet, today agriculture requires 438 million hectares* less land than it did in 2015, yet produces more adequate nutrition for all.”

*Authors' estimate

This [438 Mha] figure was arrived at by assuming that:

- Agriculture shifts away from over production of cereals, oils, and sugars, but increases fruit and vegetables;
- Agricultural yields increase $\sim 1\%/y$ between now and 2050.
- Protein consumption shifts from 86% animals and 14% plants to 50% animal and 50% plant.

“Please contact the authors for references etc. pertaining to these calculations”



Solution

Our study:

- Gain in number of hectares: three significant digits (438 millions)?
- Balancing hectares growth and population growth (our computation) results in no change in food per capita at planetary scale.

Our study:

- Neglect of diminishing returns and ecosystem stress (fertilizers, pesticides)
- More adults (higher caloric intake) in 2050 population
- Can one educate citizens globally? The case of tobacco

In conclusion the

“mismatch between what the world needed for everyone to enjoy a nutritious diet and what the world was actually producing”

is the substitution of a political problem with a technical one

Our analysis here:

Food ethics (2017) 1:173–179
DOI 10.1007/s41055-017-0020-6



DISCUSSION PAPER

Problematic Quantifications: a Critical Appraisal of Scenario Making for a Global ‘Sustainable’ Food Production

Andrea Saltelli^{1,2,3}  • Samuele Lo Piano¹

Accepted: 4 August 2017 / Published online: 15 August 2017
© Springer International Publishing AG 2017



Thank you