Ethics of quantification

Andrea Saltelli
Centre for the Study of the Sciences and the Humanities (SVT), University of Bergen (UIB), and Open Evidence Research, Open University of Catalonia

Birthday extravaganza seminar at SVT, August 26, 2019

Where to find this talk: www.andreasaltelli.eu

Centre for the Study of the Sciences and the Humanities

Photo: 123RF

Ethics of quantification

2019 Symposium of the UIB Senter for vitenskapsteori, December 5 and 6, Bergen, opening talk of Theodor Porter

The paper:

What issues for an ethics of quantification?

What recipes would be offered by an ethics of quantification?

What issues for an ethics of quantification?

-The issue of trust.
-A defence against abuse
-To prevent consequentialism in scientific quantification
-To moderate excesses of optimism about the merits of quantification
-For the non-neutrality of the techniques; for the non-separability of facts and values
-For the need to contextualize any quantification
-To deter quantification hubris

What recipes would be offered by an ethics of quantification?

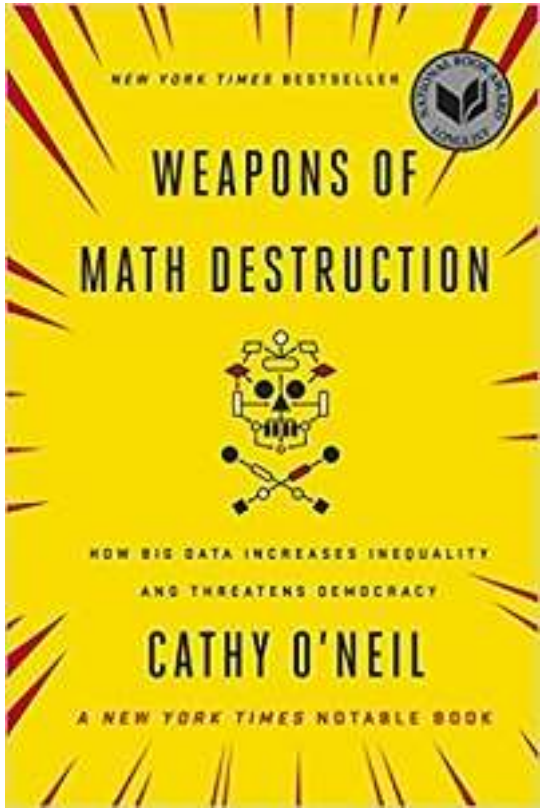
-A license not-to-quantify
-Taming hubris: memento Figure 1.
-Make use of the existing disciplinary arrangements
-Make quantifications interpretable, conveyable in plain English and context specific; use existing pedigrees
-NUSAP
-Sensitivity auditing

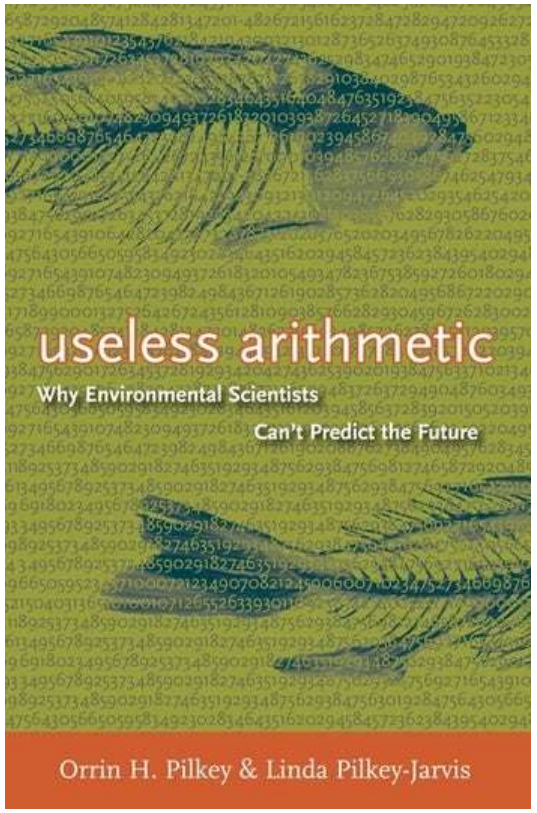
Blurring lines:

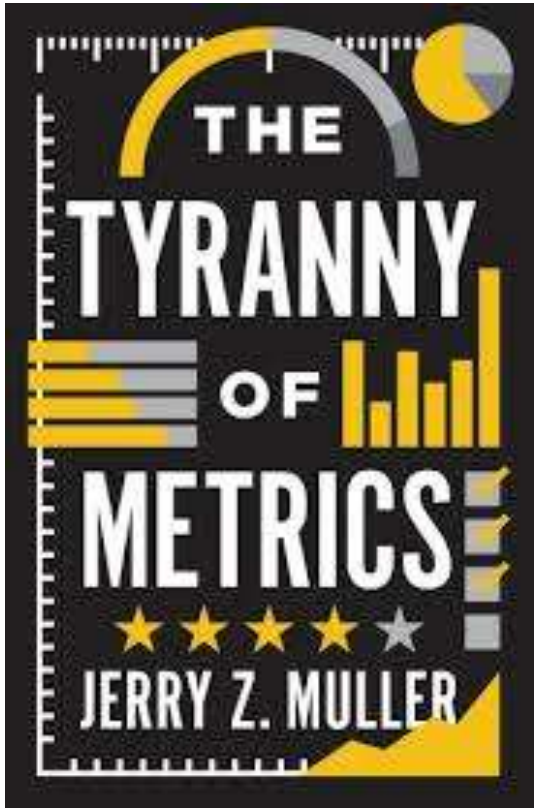
"what qualities are specific to rankings, or indicators, or models, or algorithms?"

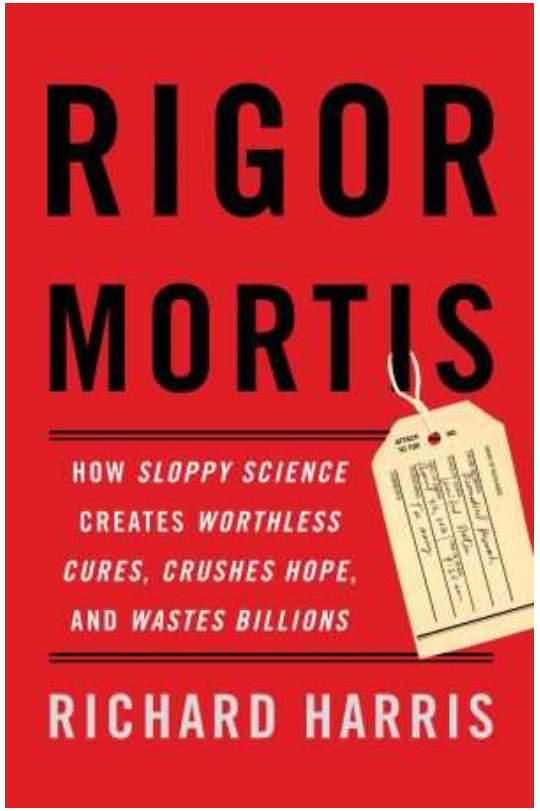
E. Popp Berman and D. Hirschman, The Sociology of Quantification: Where Are We Now?, Contemp. Sociol., vol. in press, 2017.

Algorithms, models, metrics, statistics





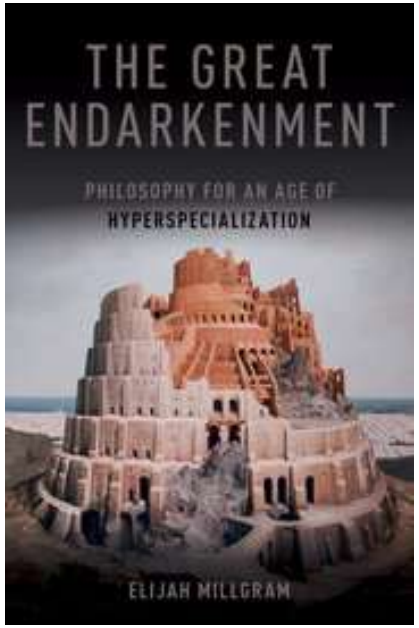




Common root causes?

Back to Elijah Millgram

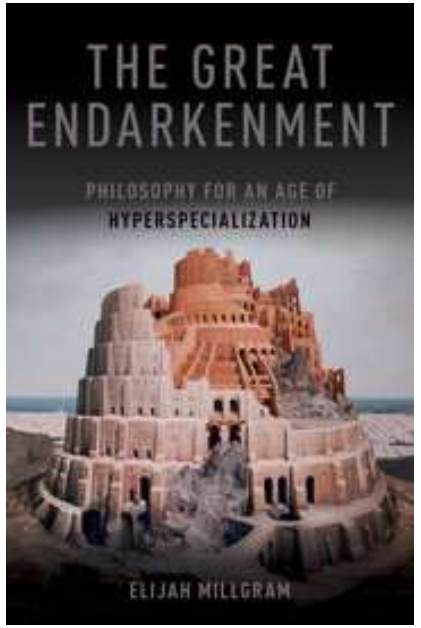
The Great Endarkenment.
Philosophy for an Age of Hyperspecialization
By Elijah Millgram



Describes a world in which all knowledge and products are the result of some form of extremely specialized expertise, and in which expertise is itself highly circumscribed, since experts depend in turn on other experts whose knowledge claims and styles of argumentation cannot be exported from one discipline to the next.

Back to Elijah Millgram

The Great Endarkenment.
Philosophy for an Age of Hyperspecialization
By Elijah Millgram



This is the world of “serial hyperspecializers” (p. 26), where the experts are “logical aliens” (p. 32)

One of the theses of Millgram is that Enlightenment's project of 'thinking for oneself' instead of deferring to authorities – produced a new class of experts (named scientists in the mid XIX century) – who become the hyper-specializers & undid the project of thinking for oneself

E. Millgram The Great Endarkenment, p. 29

Abandon the dream of a “procedural utopia”, a machinery to take the right decision based on a set of logical rules and methods

E. Millgram The Great Endarkenment, p. 23

This dream started with Condorcet's Mathématique sociale; Bentham's utilitarianism;

Today's 'decisionism' (G. Majone) – the idea that decisions can always systematically arrived at given a modicum of computation

E. Millgram The Great Endarkenment, p. 23

The critique of Andrew Stirling:

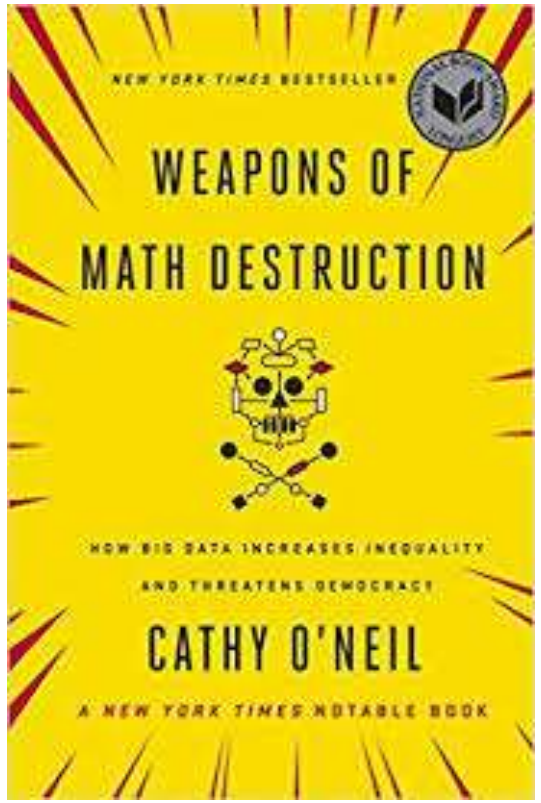
Andrew Stirling

"[…] rhetoric clamour [surrounds] 'expected utility', 'decision theory', 'life cycle assessment', 'ecosystem services' 'sound scientific decisions' and 'evidence-based policy'

[…] Each technique routinely delivers its answers with formidable levels of precision. Yet the resulting impression of accuracy is deeply misplaced"

Alarm for Weapons of Math Destruction

Cathy O'Neil

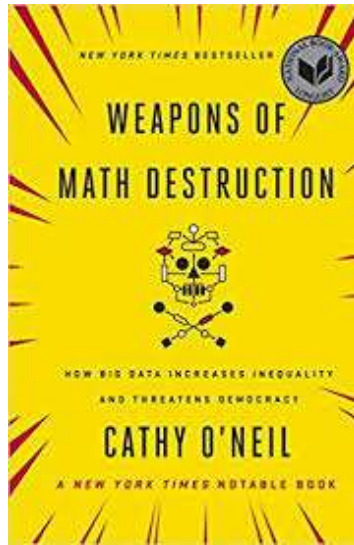


O'Neil, C. (2016). Weapons of math destruction : how big data increases inequality and threatens democracy. Random House Publishing Group.

Opacity (also because of trade secrecy) of algorithms used to decide on recruiting, carriers (including of researchers), prison sentencing, paroling, custody of minors, political campaigns…

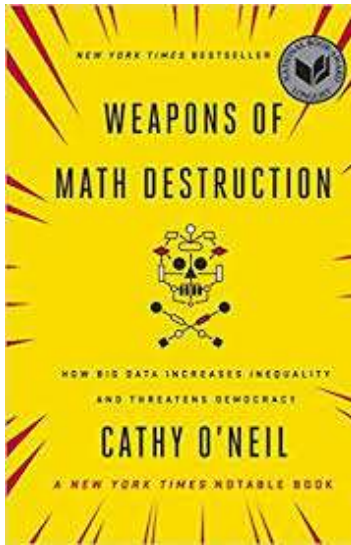
O'Neil, C. (2016). Weapons of math destruction : how big data increases inequality and threatens democracy. Random House Publishing Group.

Brauneis, R., & Goodman, E. P. (2018). Algorithmic Transparency for the Smart City. Yale Journal of Law & Technology, 20, 103–176. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3012499



Opacity coupled with opportunity for scale and damage and with non-appealability make them an instrument of oppression & inequality

Cathy O'Neil Google talk https://www.youtube.com/watch?v=TQHs8SA1qpk



From metrics fixation to the blues of statistics, from algorithmic Far West to dubious mathematical modelling, our relation with quantification needs attention

A. Saltelli, “Should statistics rescue mathematical modelling?,” arXiv, vol. arXiv:1712, no. 06457, 2018

‘Decisionism’ is mainstream

Cass Sunstein, winner of the 2018 Holberg Prize

“In a series of books (The Cost Benefit State, 2002, Risk and Reason, 2002, and The Laws of Fear, 2004), Sunstein shows the ways in which cost benefit analysis can discipline regulatory agencies”

https://www.holbergprisen.no/en/holberg-prize/prize-winners/cass-r-sunstein

Can technocracy be saved? An interview with Cass Sunstein.

Obama's regulation czar makes the case that "the issues that most divide us are fundamentally about facts rather than values."

By Dylan Matthews | @dylanmatt | dylan@vox.com | Oct 22, 2018, 9:00am EDT

https://www.vox.com/future-perfect/2018/10/22/18001014/cass-sunstein-cost-benefit-analysis-technocracy-liberalism

“Often, immersion in the facts often makes value disagreements feel much less relevant”
(C. Sunstein)

https://www.vox.com/future-perfect/2018/10/22/18001014/cass-sunstein-cost-benefit-analysis-technocracy-liberalism

The Sameness of Cass Sunstein

His books keep pushing the same technocratic fixes. But today's most pressing questions cannot be depoliticized.

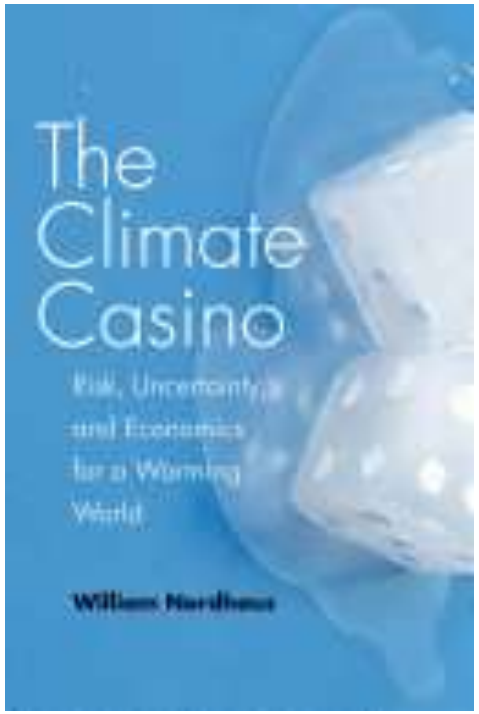
By **AARON TIMMS** | June 20, 2019

A critique of Sunstein's faith in 'nudge' and cost benefit analysis

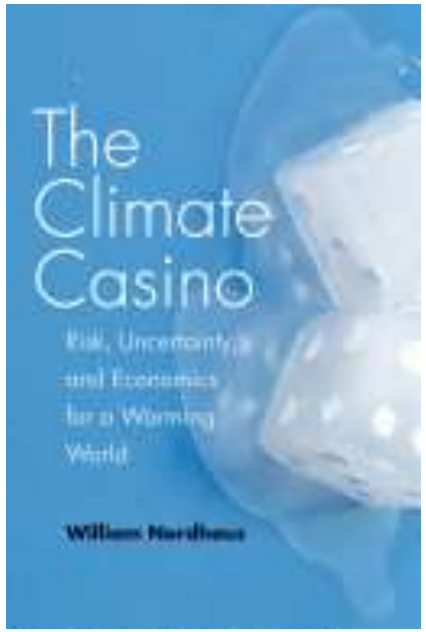
https://newrepublic.com/article/154236/sameness-cass-sunstein

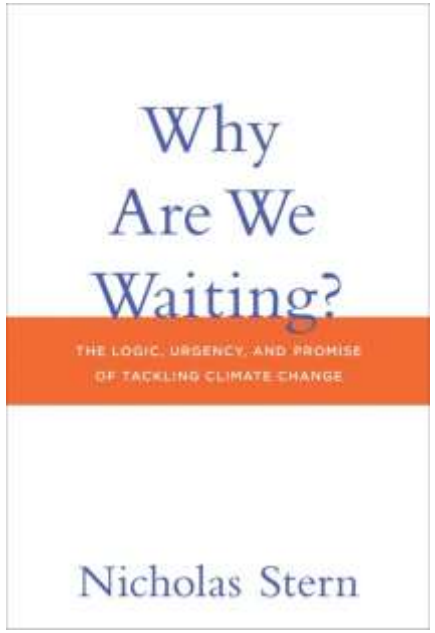
One of the winner of Nobel prize for economics 2018 is Willem Nordhaus, for his work on the economics of climate change.

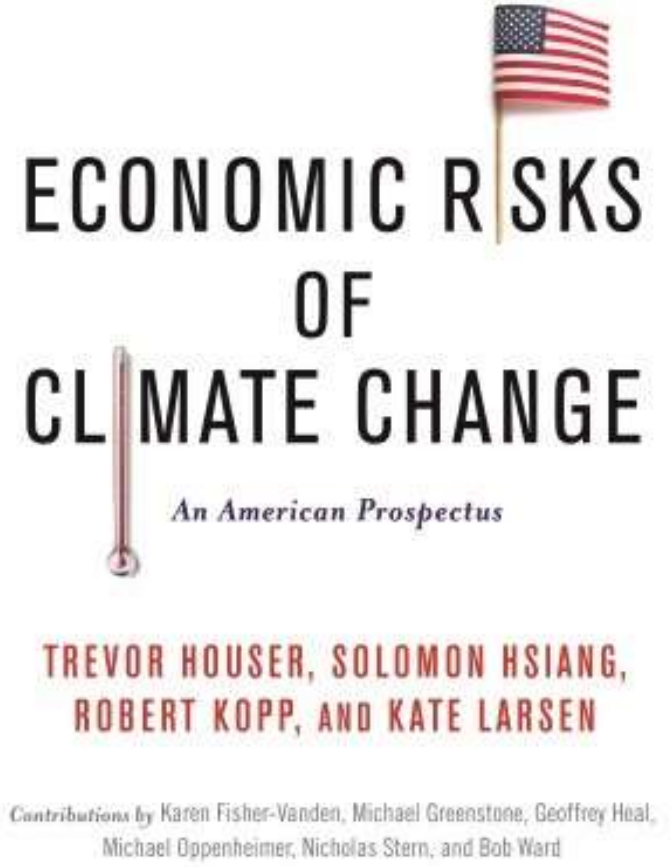
Cost benefit analysis to the year 2100?



Are these licit quantifications?

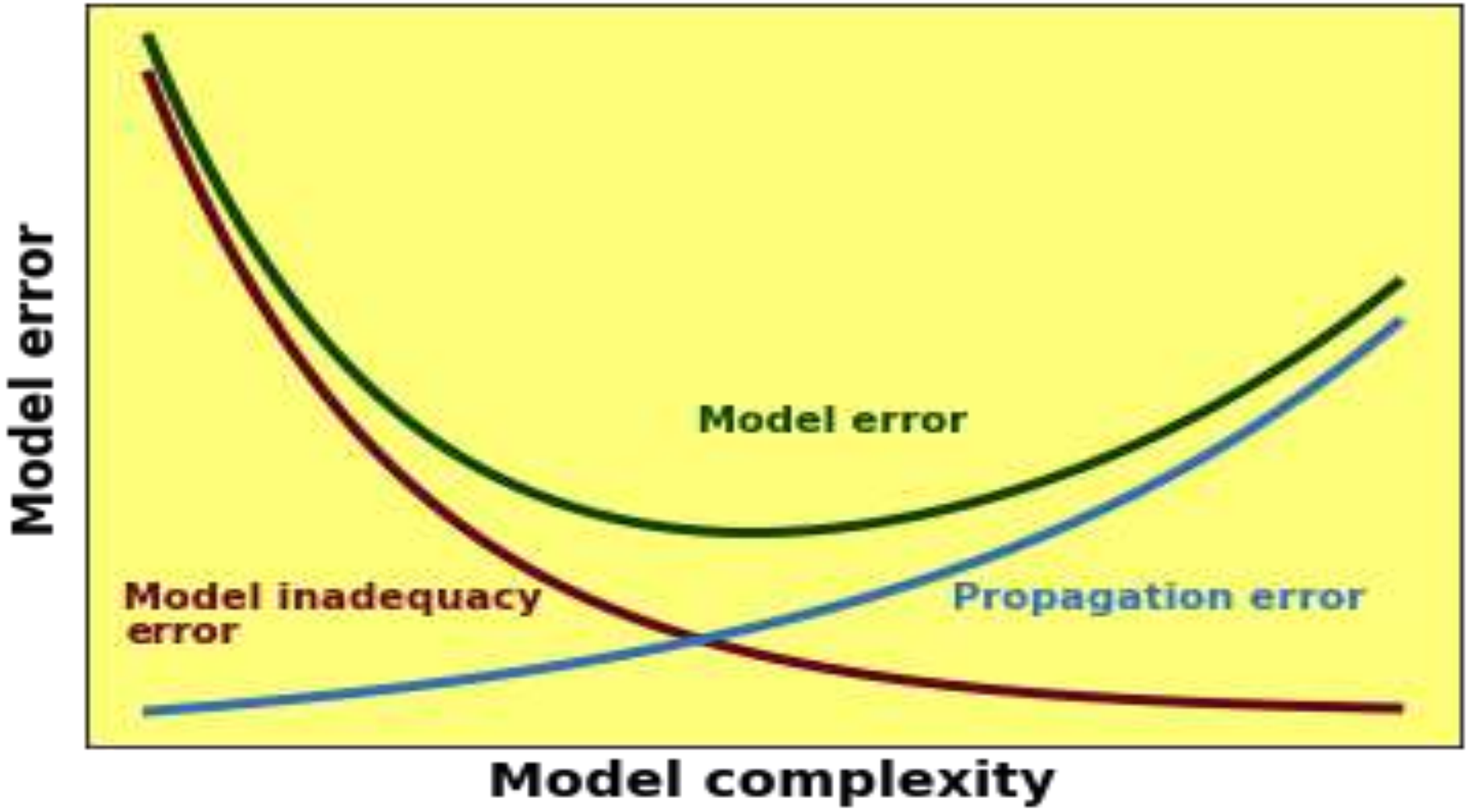






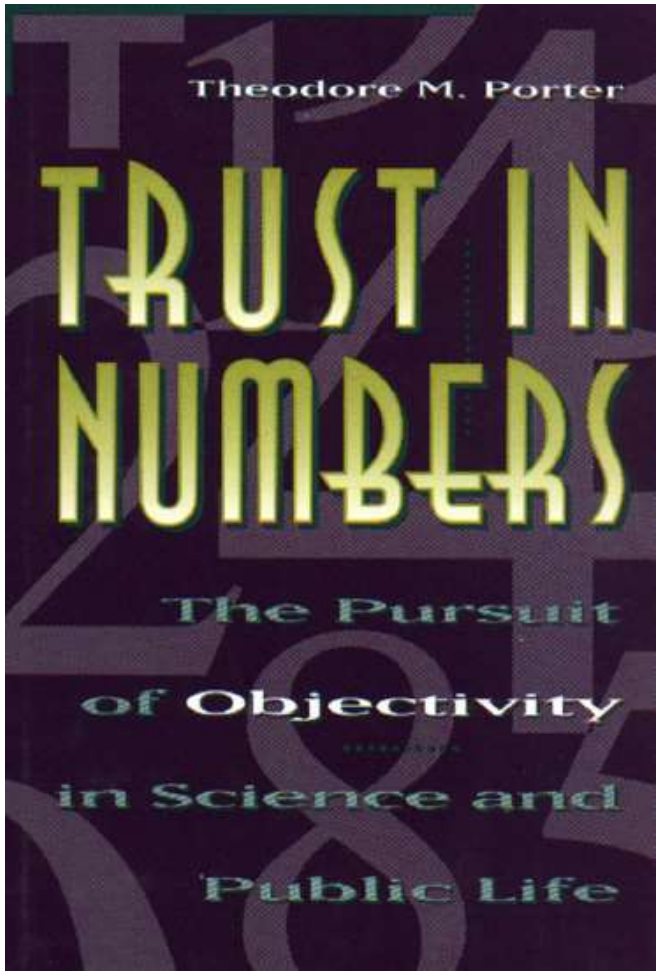
Saltelli, A., Stark, P.B., Becker, W., and Stano, P. , 2015, Climate Models as Economic Guides. Scientific Challenge or Quixotic Quest? Issues in Science and Technology (IST), Volume XXXI Issue 3, Spring 2015, https://issues.org/climate-models-as-economic-guides-scientific-challenge-or-quixotic-quest/

Uncertainty quantification and sensitivity analysis to calibrate complexity



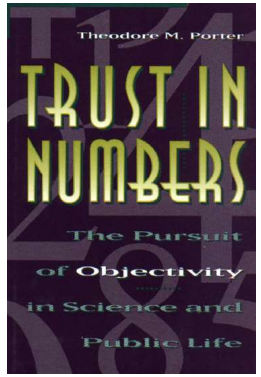
Numbers and trust

Theodor
M. Porter



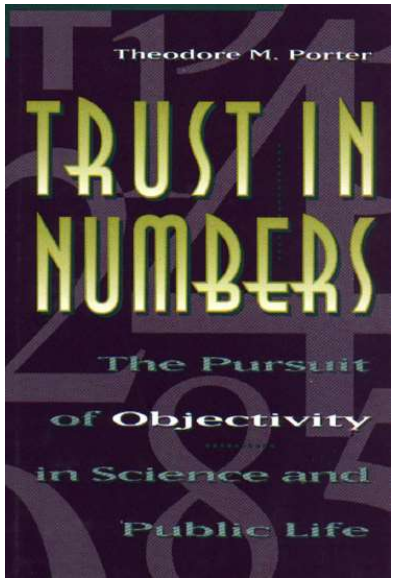
Theodore M. Porter, Trust in Numbers,
The Pursuit of Objectivity in Science and Public Life, Princeton 1995

p. 8: “The appeal of numbers is especially compelling to bureaucratic officials who lack the mandate of a popular election, or divine right.



Arbitrariness and bias are the most usual grounds upon which such officials are criticized.

A decision made by the numbers (or by explicit rules of some other sort) has at least the appearance of being fair and impersonal.”



p. 8: “Scientific objectivity thus provides an answer to a moral demand for impartiality and fairness.

Quantification is a way of making decisions without seeming to decide.

Objectivity lends authority to officials who have very little of their own.”

Trust, authority and styles of quantification: two different stories

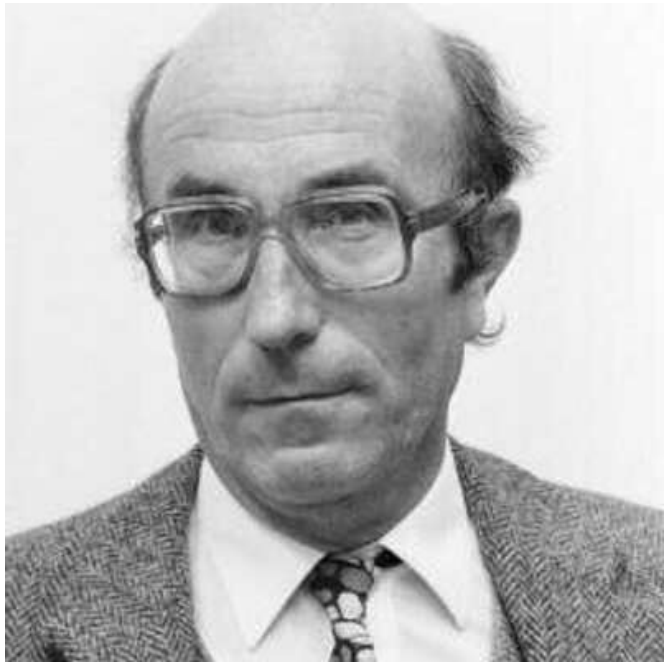
ECOLE NATIONALE
DES PONTS ET CHAUSSEES

US Army Corps of Engineers®

Porter's story: Quantification needs judgment which in turn needs trust …without trust quantification becomes mechanical, a system, and 'systems can be played'.

‘System trust’, is social system theory:

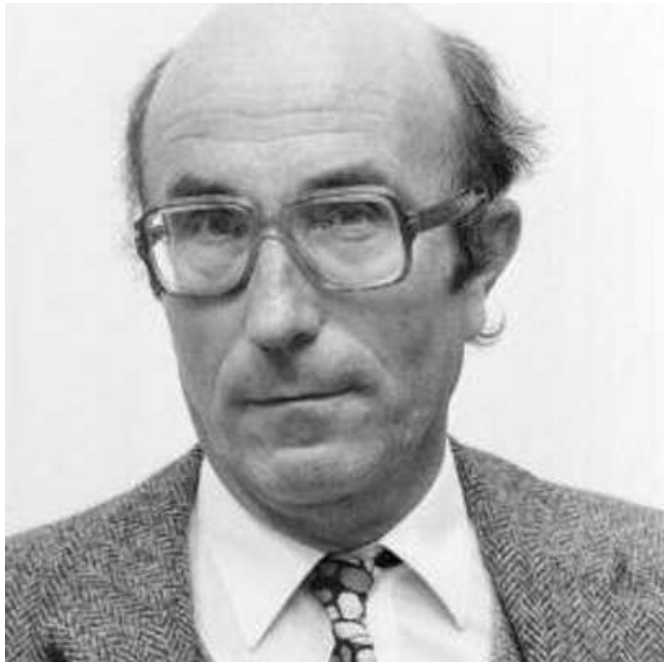
“The reduction of complexity [made possible by generalized media of communication as money, power and truth] **assumes trust** on the part of those who are expecting such reduction and of those who are supposed to accept it once it is accomplished”



Niklas Luhmann

N. Luhmann, Trust and Power. Polity Press, 2017.

"[System trust thus permits] the bank to lend more money than it possess, the state to issue more commands than it can enforce using the police, that more information is divulged in professional advice than could be backed up empirically or logically".



Niklas Luhmann

N. Luhmann, Trust and Power. Polity Press, 2017.

‘the essential fiduciary status’ of science= Trust in science is necessary for the general society to continue to support it, materially and with recruits. And mutual trust within science is necessary for its systems of quality assurance to function

Jerome R. Ravetz

Charles Goodhart

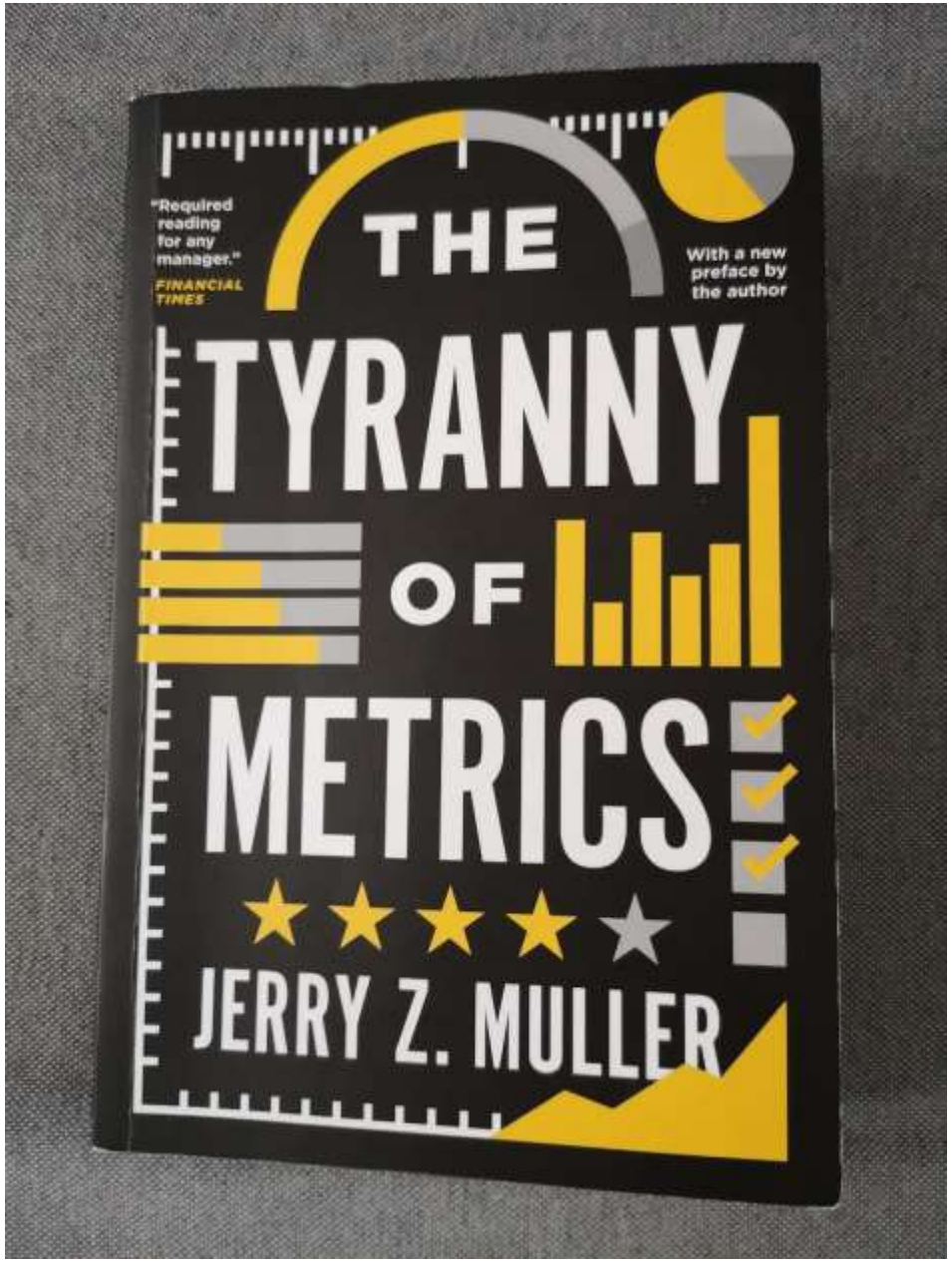
p. 44 “Any … measures necessarily involve a loss of information … [and distorts behavior]” (Porter, 1995)

This is what we normally call Goodhart’s law, from Charles Goodhart. "When a measure becomes a target, it ceases to be a good measure."

Also known as Campbell's law (1976);
https://en.wikipedia.org/wiki/Goodhart%27s_law

For Ravetz (1971, pp. 295–296), when the goals of a task are complex, sophisticated, or subtle, then crude systems of measurements can be played exactly by those persons possessing the skills to execute the tasks properly, who thus manage to achieve their own goals to the detriment of those assigned.

Ravetz, J.R., 1971, Scientific Knowledge and Its Social Problems, 1996 Edition, Transaction Publishers. See plenty of examples in Muller, J.Z., 2018, The Tyranny of Metrics, Princeton.



More reading

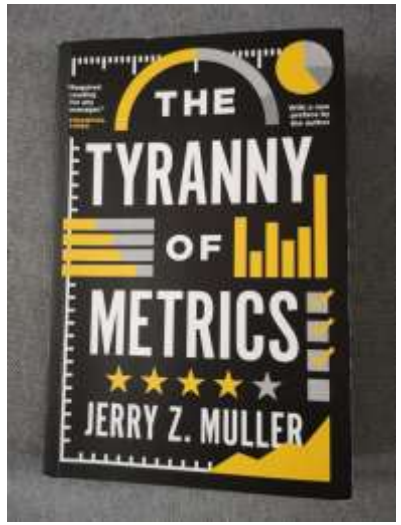
J. Z. Muller, The tyranny of metrics. Princeton University Press , 2018.

Metric fixation, or the irresistible pressure to measure performance

Gaming of metrics (recall Goodhart law)

“The calculative is the enemy of the imaginative”

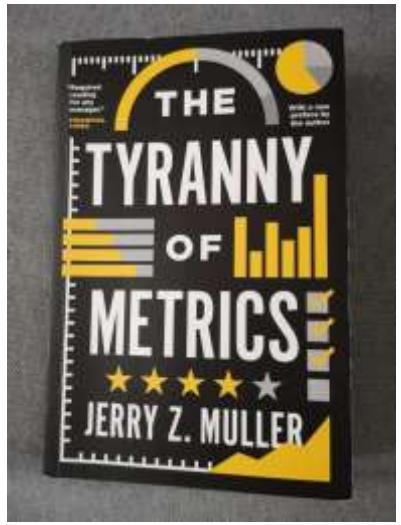
A wealth of case studies from education to war to medicine to foreign aid..



Critiques of metrics

From the left: metric fixation promotes deskilling

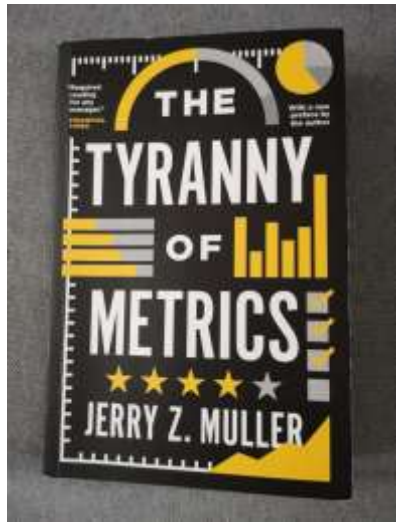
From the right (Friedrich Hayek): metric fixation reproduces features of the soviet system



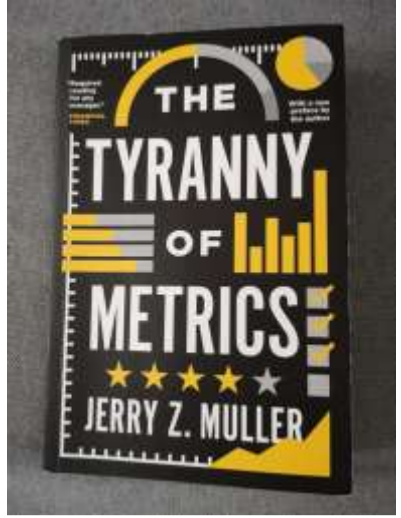
Critiques of metrics

An epistemological critique: metrics privilege abstract and formulaic knowledge against practical and tacit knowledge

(Greek concept of metis)



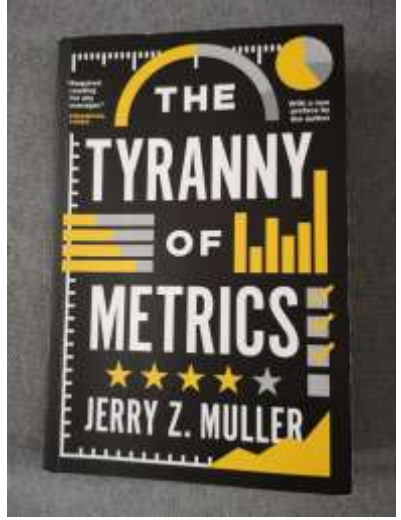
Unintended consequences: a litany



- 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

A concluding remark

Considering all of the above keep in mind at every step that “the best use of metrics may be not to use it at all”



Theodor Porter:

"The evasion of goals and corruption of measures tends to make these numbers "funny" in the sense of becoming dishonest, while the mismatch between boring, technical appearances and cunning backstage manipulations supplies dark humor"

T. M. Porter, "Funny Numbers," *Culture Unbound*, vol. 4, pp. 585–598, 2012

The numbers of neoliberalism

How CEOs profited from the ambiguities and manipulability. “These men did not allow their enterprises to fail until they failed catastrophically”

T. M. Porter, “Funny Numbers,” *Culture Unbound*, vol. 4, pp. 585–598, 2012

“[CEOs] had the power to keep the numbers boring, maintaining a screen in front of this theater of the absurd…”

Tin description (a result of standardization) allow tin prescriptions, a strategy of impersonal regulation, deploying statistics as insurance against casuistry

T. M. Porter, “Funny Numbers,” *Culture Unbound*, vol. 4, pp. 585–598, 2012

Thus onstage we see the boring numbers of thin prescription, which ensure trust and the containment of subjectivity

Offstage we see the resulting intense struggle about how the quantification should be made

T. M. Porter, "Funny Numbers," *Culture Unbound*, vol. 4, pp. 585–598, 2012

E.g. an immediate impact of thin prescriptions in education is “to encourage the reconstruction of school curricula to match the content of the tests, and sometimes to make the temptation to cheat almost irresistible” (➔ J.Z. Muller; ➔ OECD-PISA example)

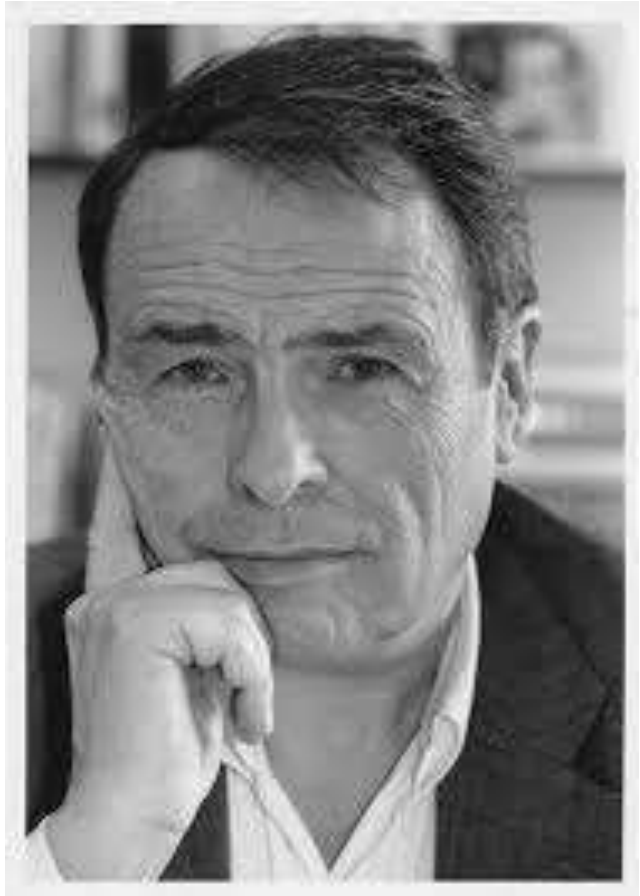
T. M. Porter, “Funny Numbers,” *Culture Unbound*, vol. 4, pp. 585–598, 2012

Do we need a movement of resistance?

I. Bruno, E. Didier, and J. Prévieux, Stat-activisme. Comment lutter avec des nombres. Paris: Zones, La Découverte, 2014

How to be a "statactiviste"?

1. Deconstruct existing metrics, including using irony (Pierre Bourdieu, *Les héritiers*).



La sociologie, ça doit être rigolo

(Sociology must be fun)

How to be a "statactiviste"?

2. Gaming metrics (statistical judo) – use Goodhart's law to your advantage – or make the ruse public.

- Police statistics in NY

How to be a "statactiviste"?

3. Bring to the surface what is hidden / unsaid/ excluded – new social classes, marginalization, minorities:

- ‘Creative class’ or ‘precarious intellectuals’?

How to be a "statactiviste"?

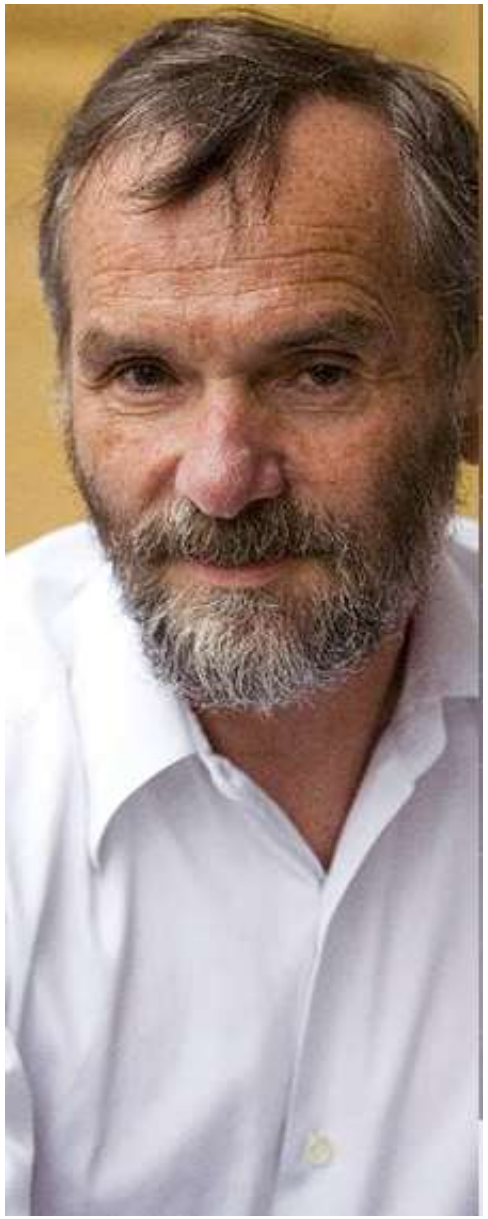
4. Measure something different.

- Suicides at France Telecom;
- BIP 40, a new French measure of poverty/inequality

Important:

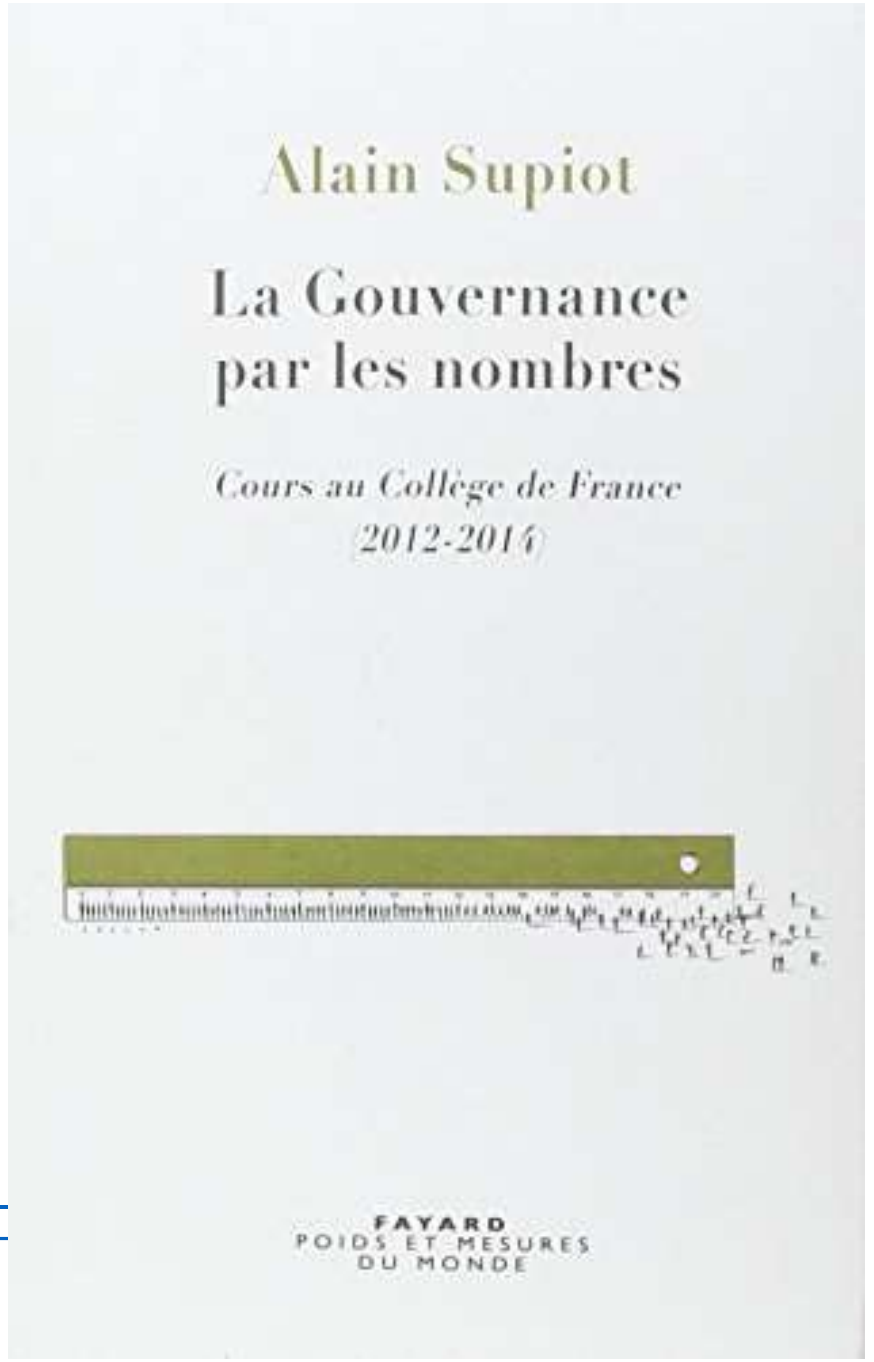
"Quantification should not be abandoned to the advantage of exalting qualities, singularities, and the incommensurable. Such an abandon would be a tactical error"

Alain Supiot

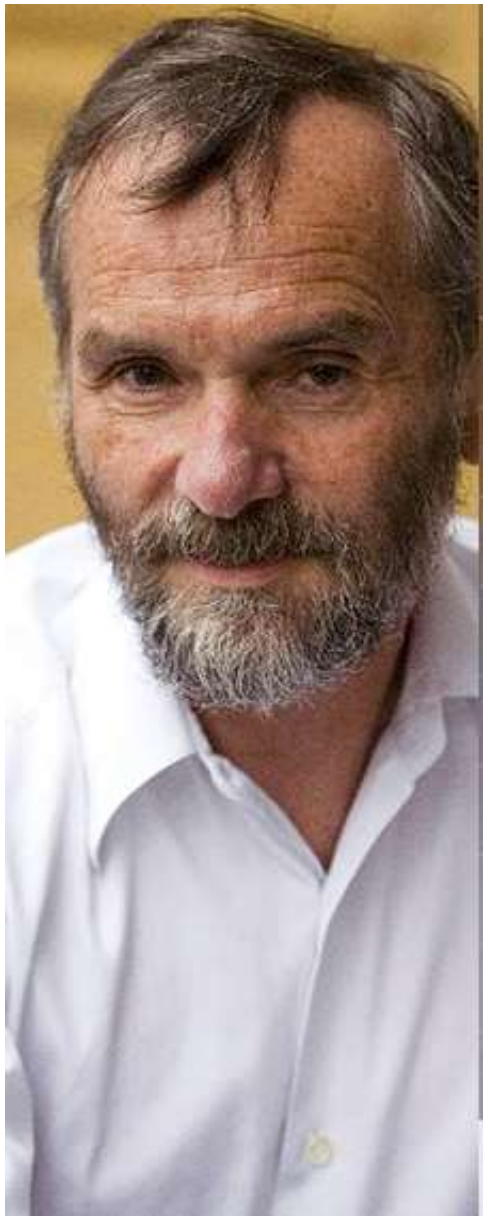


An indictment of the Total Market and the normative uses of economic quantification

https://www.college-de-france.fr/site/en-alain-supiot/Governance-by-Numbers-Introduction.htm



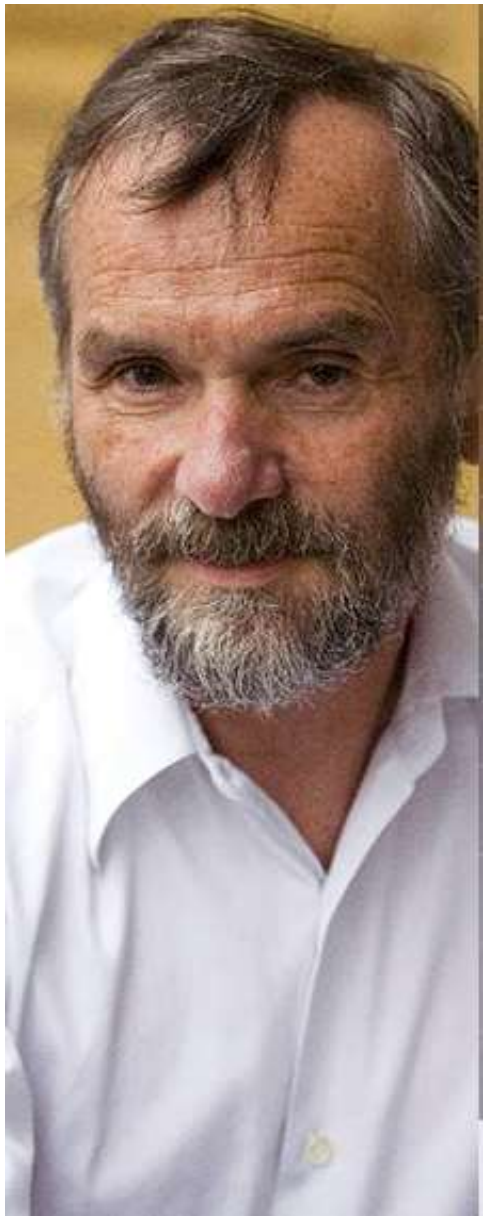
Alain Supiot



…we have entered the era of the cybernetic imaginary, which revives the West's age-old dream of grounding social harmony in calculations.

Repudiating the goal of governing by just laws, this new discourse advocates in its stead the attainment of measurable objectives efficiently

Alain Supiot



… This leaves no option open to populations or countries than to ride roughshod over social legislation, and pledge allegiance to those stronger than they are

The End

@andreasaltelli

Caeteris are never paribus

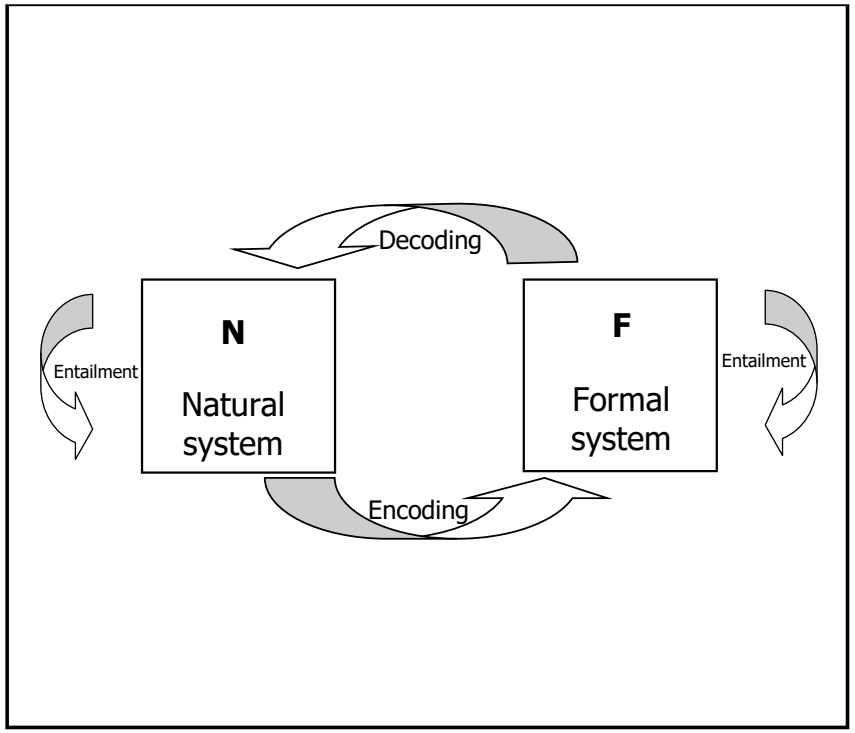
On mathematical modelling

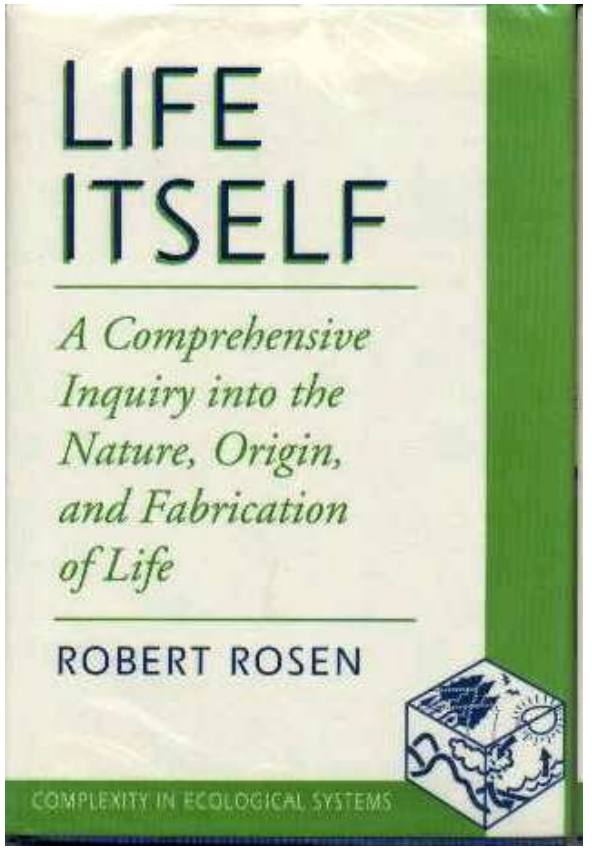
An example

"[…] The process of constructing and validating [value-at risk] models is time consuming and detail oriented; normally even the people who produced the model will not remember many of the assumptions incorporated into it, short of redoing their work, which means that the client cannot simply ask then what went into it."

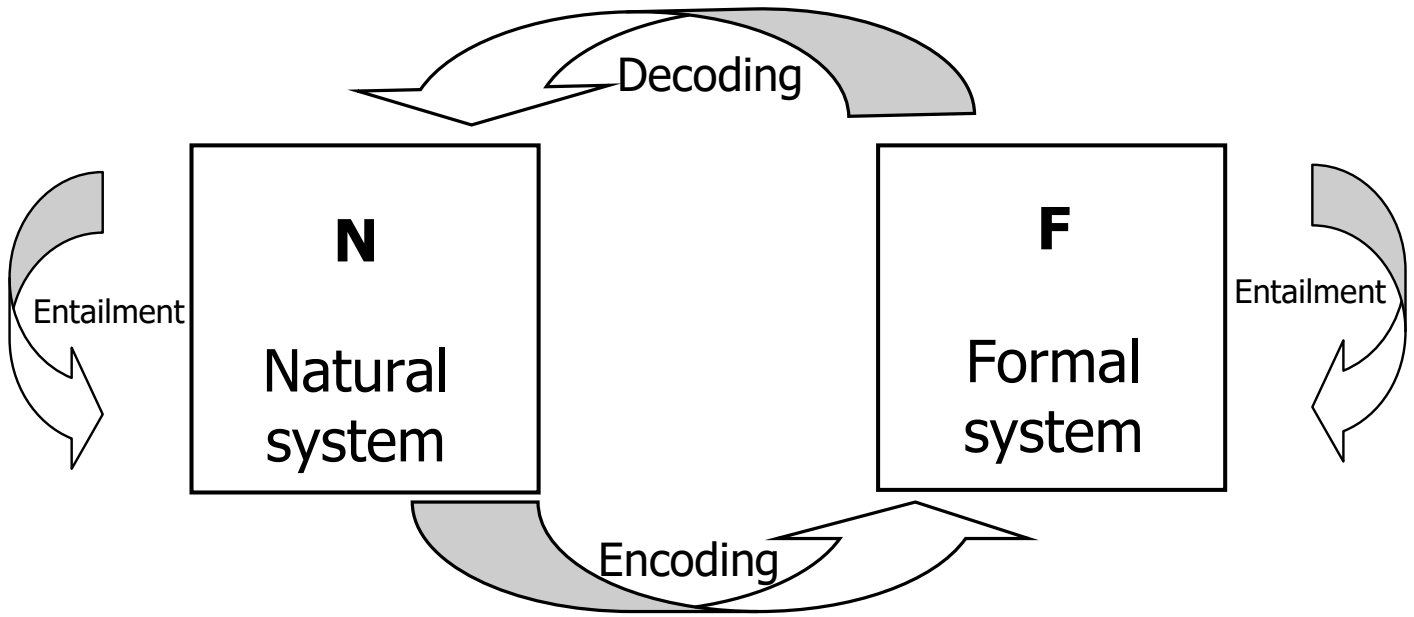
E. Millgram The Great Endarkenment, p. 29

Modelling as a craft rather than as a science for Robert Rosen





R. Rosen, Life Itself: A Comprehensive Inquiry Into the Nature, Origin, and Fabrication of Life. Columbia University Press, 1991.



What is a model ?

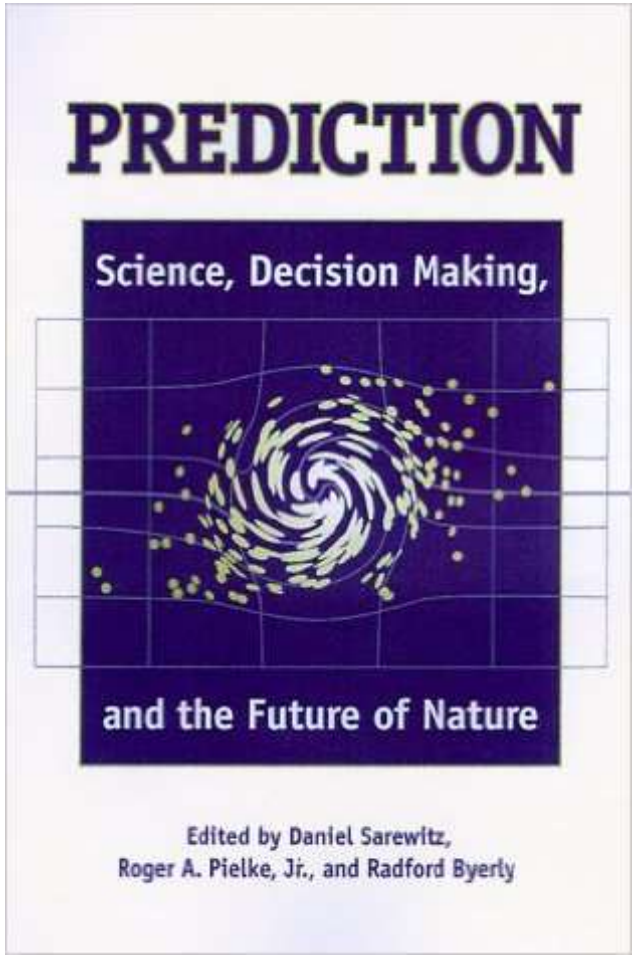
Robert Rosen

“models are most useful when they are used to challenge existing formulations, rather than to validate or verify them”

Naomi Oreskes

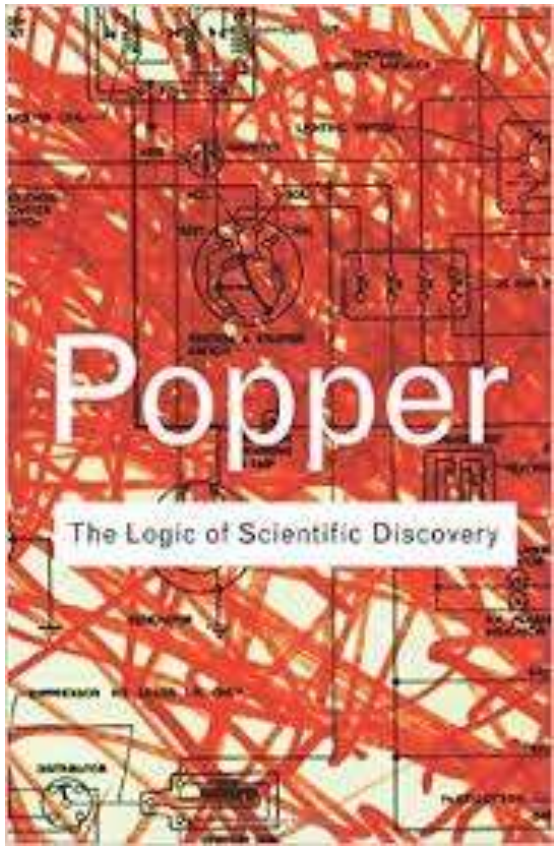
N. Oreskes, K. Shrader-Frechette, and K. Belitz, “Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences,” Science, 263, no. 5147, 1994.

Models are not physical laws



Oreskes, N., 2000, Why predict? Historical perspectives on prediction in Earth Science, in Prediction, Science, Decision Making and the future of Nature, Sarewitz et al., Eds., Island Press, Washington DC

“[…] to be of value in theory testing, the predictions involved must be capable of refuting the theory that generated them”
(N. Oreskes)



“In many cases, these temporal predictions **are treated with the same respect** that the hypothetic-deductive model of science accords to logical predictions. But this respect is largely misplaced”

“[…] models are complex amalgam of theoretical and phenomenological laws (and the governing equations and algorithms that represent them), empirical input parameters, and a model conceptualization […] When a model generates a prediction, of what precisely is the prediction a test? The laws? The input data? The conceptualization? Any part (or several parts) of the model might be in error, and there is no simple way to determine which one it is”

Model-based knowing is conditional

When models need as input information which we don't have

John Kay

J. A. Kay, "Knowing when we don't know," 2012, https://www.ifs.org.uk/docs/john_kay_feb2012.pdf

WebTAG: Annual Percentage Change in Car Occupancy (% pa) up to 2036

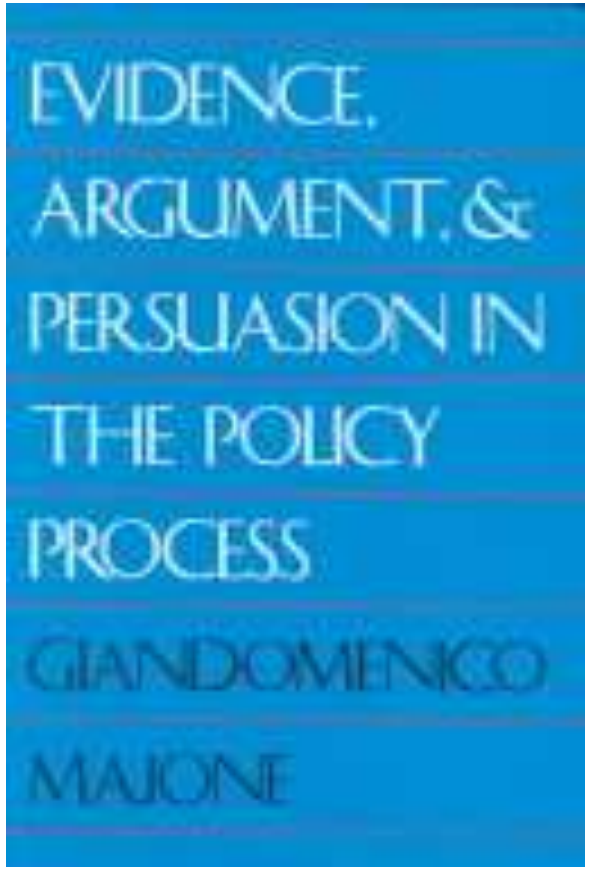
Journey Purpose	Weekday					Weekend	All Week
	7am-10am	10am-4pm	4pm-7pm	7pm-7am	Weekday Average		
Work	-0.48	-0.4	-0.62	-0.5	-0.44	-0.48	-0.45
Non - Work (commuting and other)	-0.67	-0.65	-0.53	-0.47	-0.59	-0.52	-0.56

Definitions

Uncertainty analysis: Focuses on just quantifying the uncertainty in model output

Sensitivity analysis: The study of the relative importance of different input factors on the model output

Why Sensitivity analysis?



"Are the results from a particular model more sensitive to changes in the model and the methods used to estimate its parameters, or to changes in the data?"

4. SENSITIVITY AND UNCERTAINTY ANALYSES

Page 391

Six steps for a global SA:

1. Select <u>one</u> output of interest;
2. Participatory step: discuss which input may matter;
3. Participatory step (extended peer review): define distributions;
4. Sample from the distributions;
5. Run (=evaluate) the model for the sampled values;
6. Obtain in this way bot the uncertainty of the prediction and the relative importance of variables.

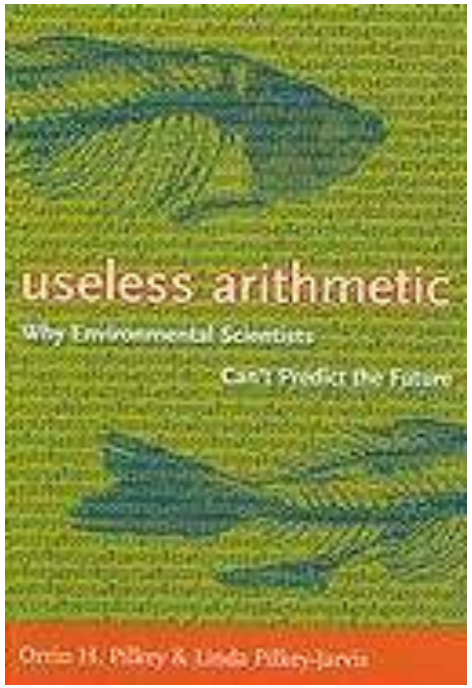
Is something wrong with this statement (p. 384 of EC guidelines)

The influence of the key variables

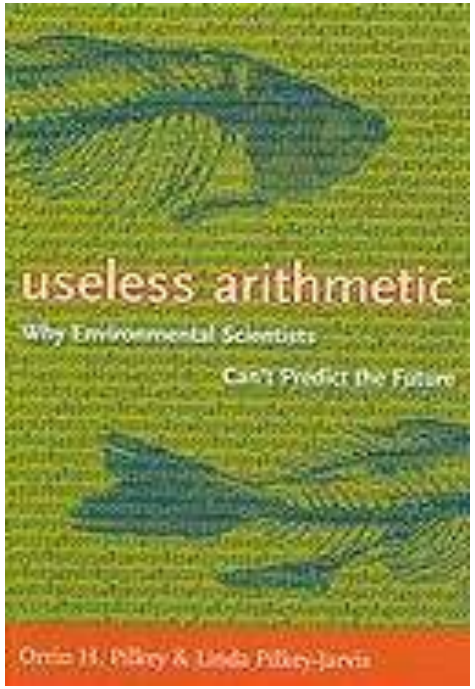
should be investigated by a sensitivity analysis.

Limits of sensitivity analysis

Orrin H.
Pilkey



Useless Arithmetic: Why Environmental Scientists Can't Predict the Future by Orrin H. Pilkey and Linda Pilkey-Jarvis, Columbia University Press, 2009.

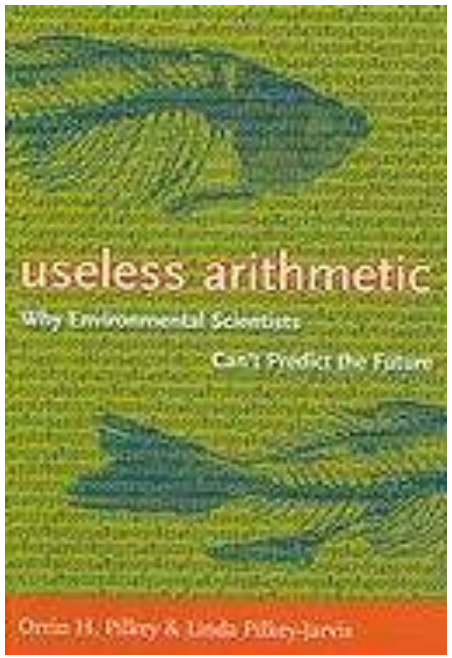


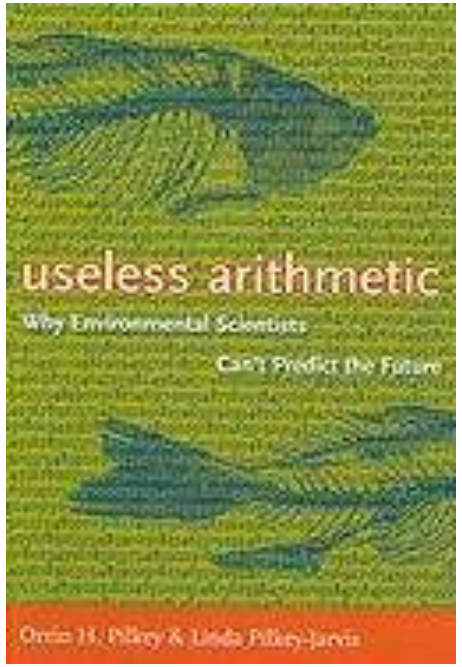
<<It is important, however, to recognize that the sensitivity of the parameter in the equation is what is being determined, not the sensitivity of the parameter in nature.

[…] If the model is wrong or if it is a poor representation of reality, determining the sensitivity of an individual parameter in the model is a meaningless pursuit.>>

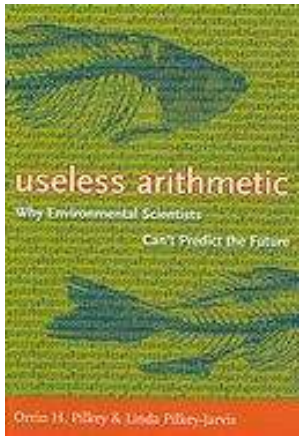
One of the examples discussed concerns the Yucca Mountain repository for radioactive waste. TSPA model (for total system performance assessment) for safety analysis.

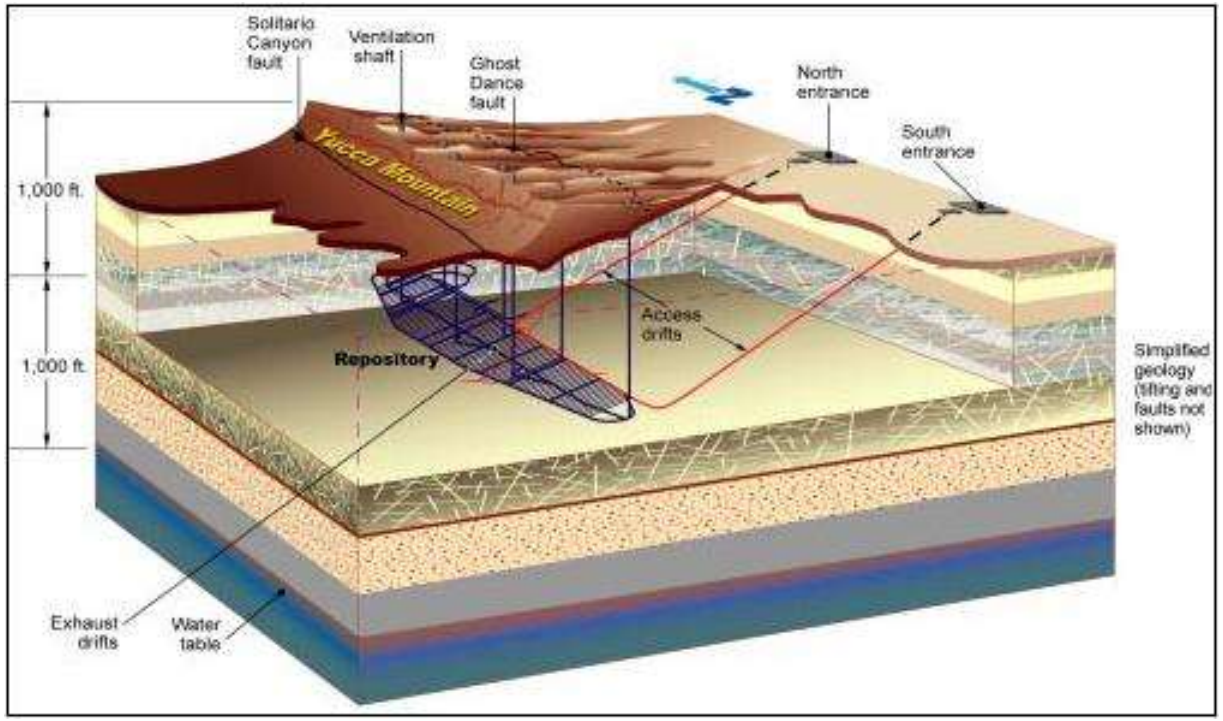
TSPA is Composed of 286 sub-models.





TSPA (like any other model) relies on assumptions → one is the low permeability of the geological formation → long time for the water to percolate from surface to disposal.





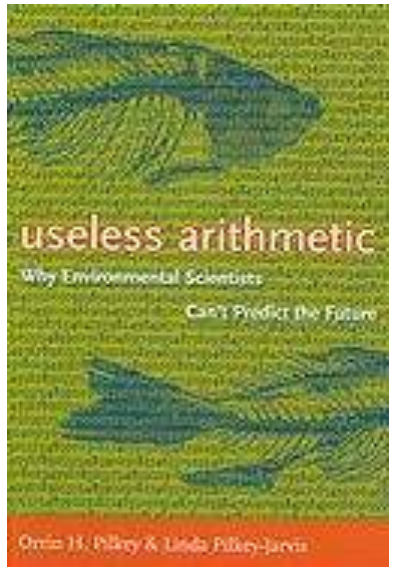
The confidence of the stakeholders in TSPA was not helped when evidence was produced which could lead to an upward revision of 4 orders of magnitude of this parameter

(the ^{36}Cl story)

Type III error in sensitivity: Examples:

In the case of TSPA (Yucca mountain) a range of 0.02 to 1 millimetre per year was used for percolation of flux rate.

→… SA useless if it is instead ~ 3,000 millimetres per year.



Robert K. Merton

"Scientific mathematical modelling should involve constant efforts to falsify the model"

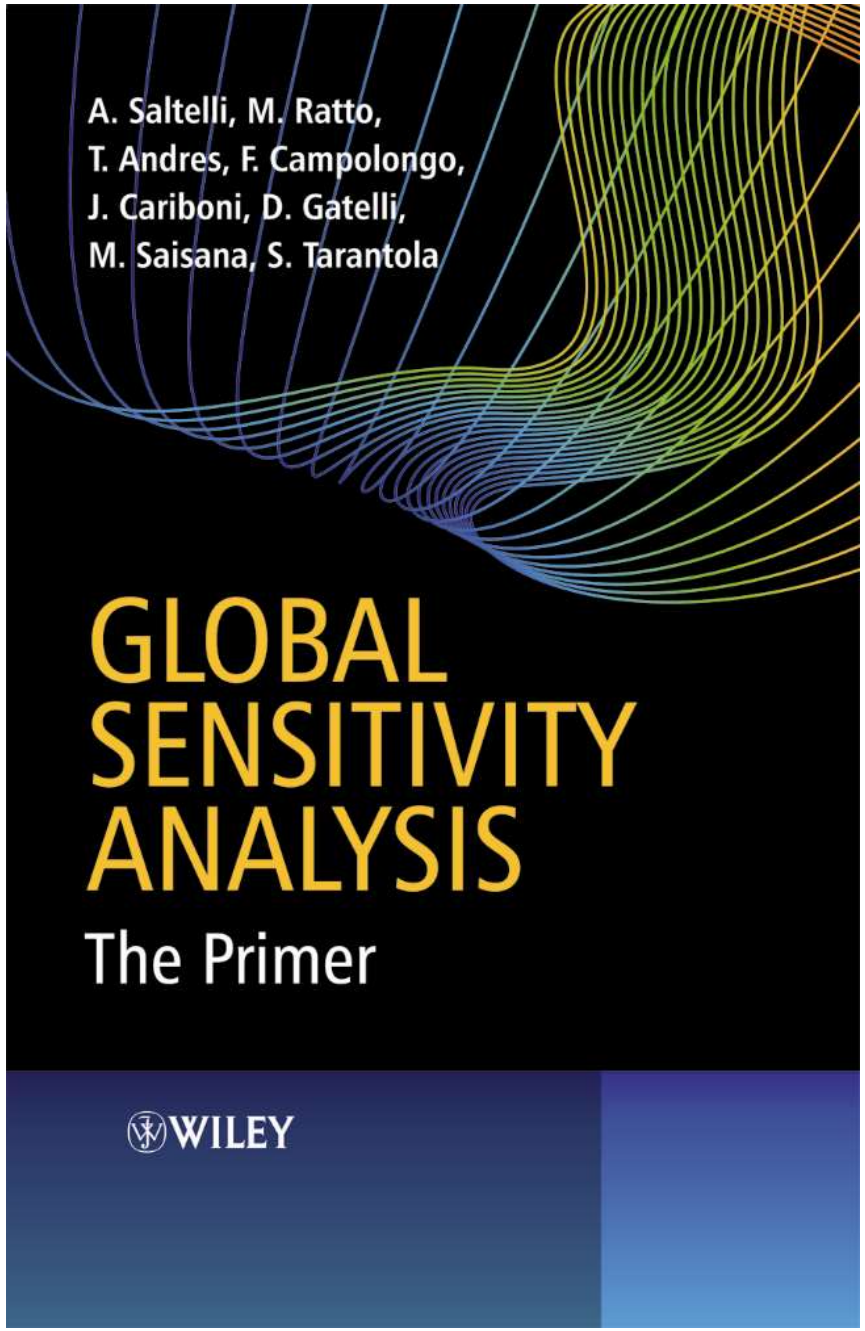
Ref. ➔ Robert K. Merton's 'Organized skepticism '

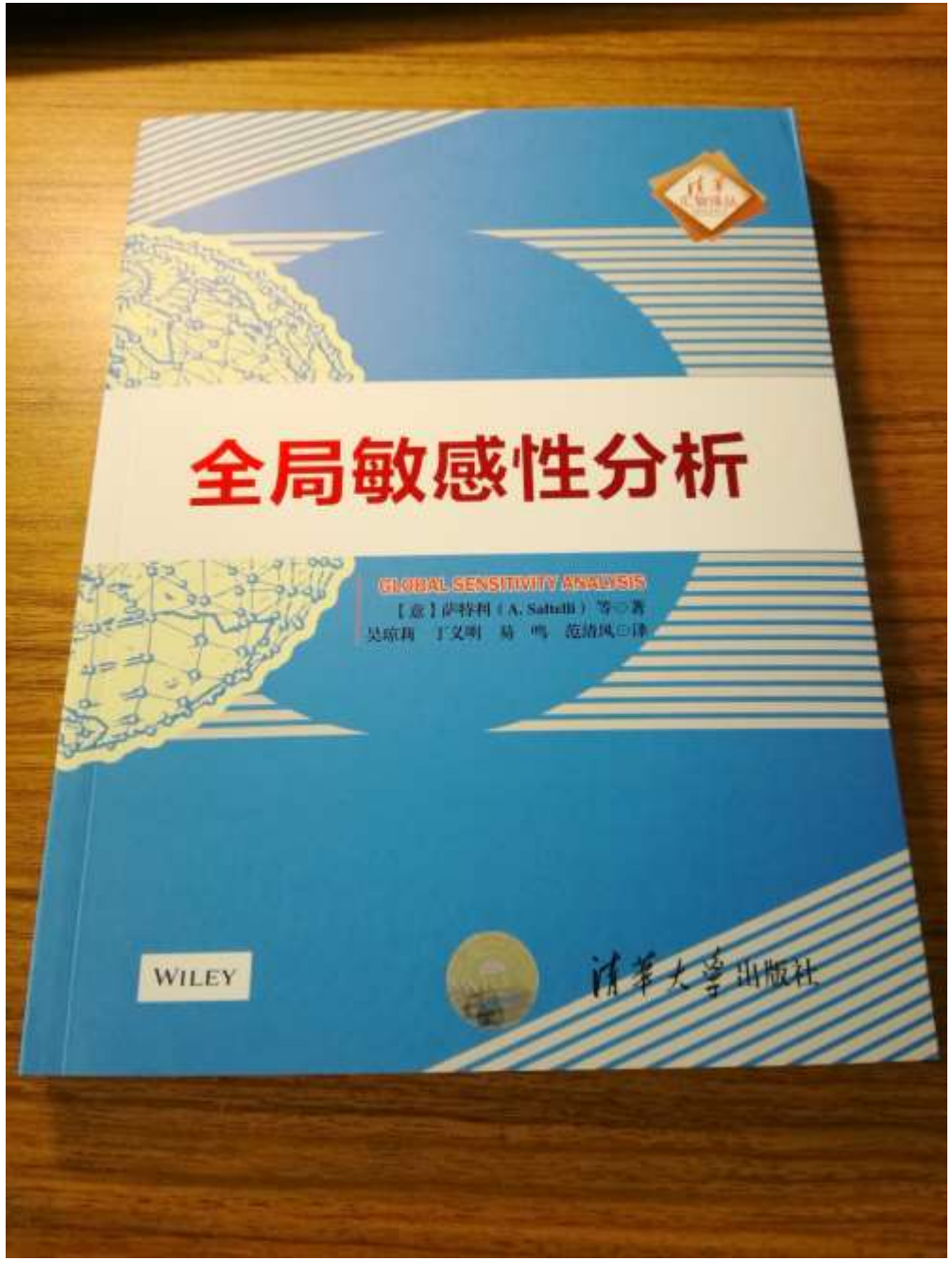
Communalism – the common ownership of scientific discoveries, according to which scientists give up intellectual property rights in exchange for recognition and esteem (Merton actually used the term Communism, but had this notion of communalism in mind, not Marxism);

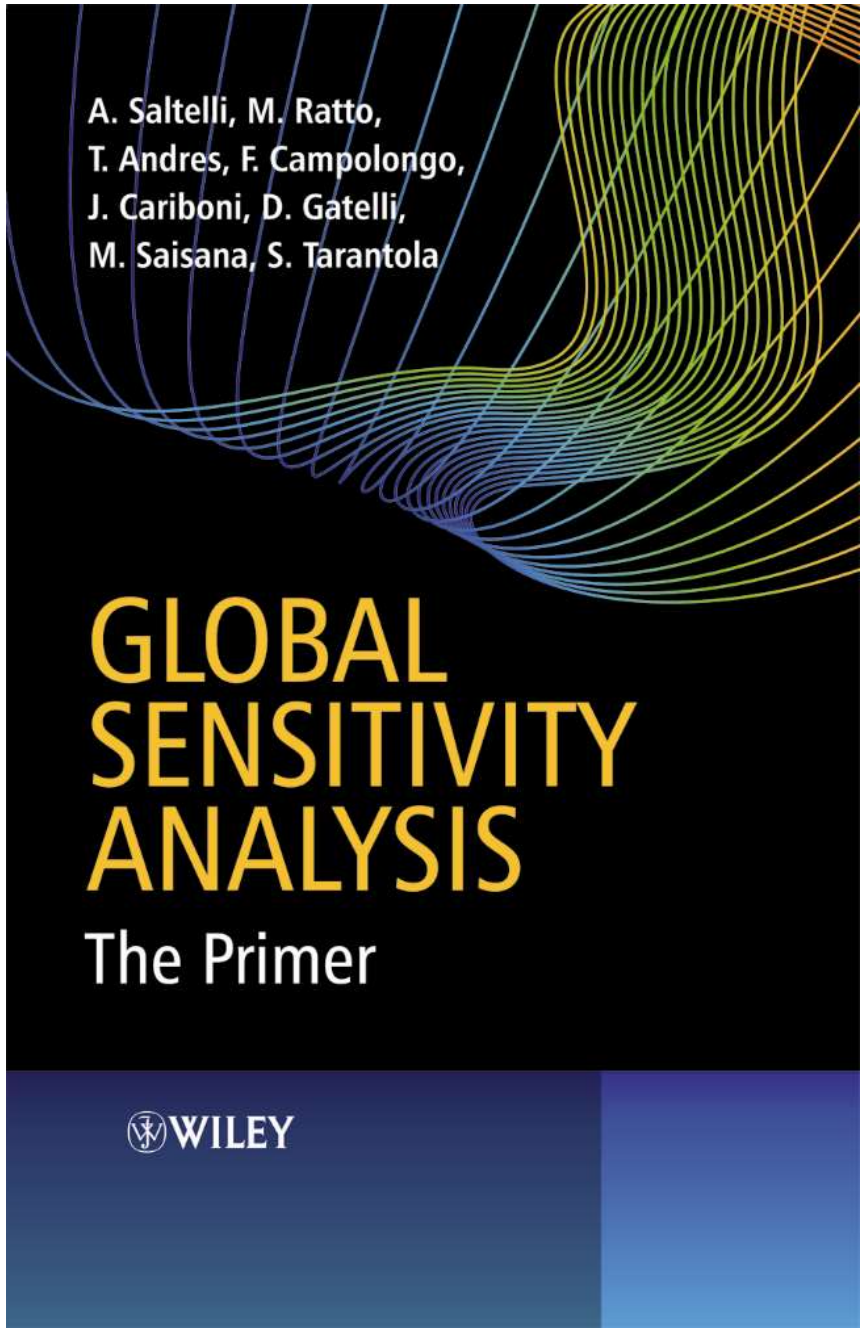
Universalism – according to which claims to truth are evaluated in terms of universal or impersonal criteria, and not on the basis of race, class, gender, religion, or nationality;

Disinterestedness – according to which scientists are rewarded for acting in ways that outwardly appear to be selfless;

Organized Skepticism – all ideas must be tested and are subject to rigorous, structured community scrutiny.


A. Saltelli, M. Ratto,
T. Andres, F. Campolongo,
J. Cariboni, D. Gatelli,
M. Saisana, S. Tarantola
GLOBAL
SENSITIVITY
ANALYSIS
The Primer
WILEY


全局敏感性分析
GLOBAL SENSITIVITY ANALYSIS
WILEY



Available for free at

http://www.andreasaltelli.eu

Secrets of sensitivity analysis

Why should one ever run a model just once?

EC impact assessment guidelines: sensitivity analysis & auditing

http://ec.europa.eu/smart-regulation/guidelines/docs/br_toolbox_en.pdf

First secret: The most important question is the question.

Or: sensitivity analysis is not “run” on a model but on a model once applied to a question

Second secret: Sensitivity analysis should not be used to hide assumptions [it often is]

Third secret: If sensitivity analysis shows that a question cannot be answered by the model one should find another question or model

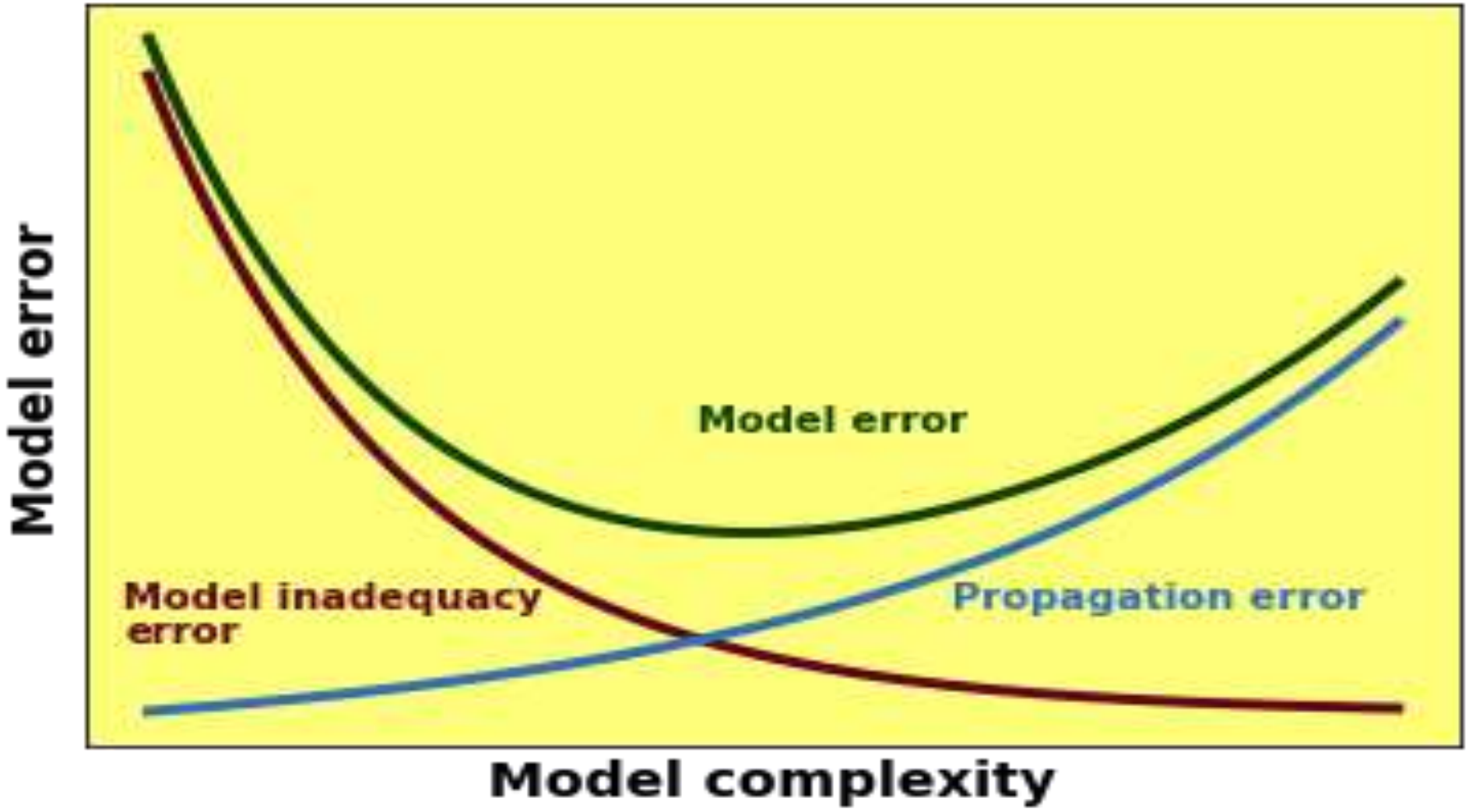
[Often the love for one's own model prevails]

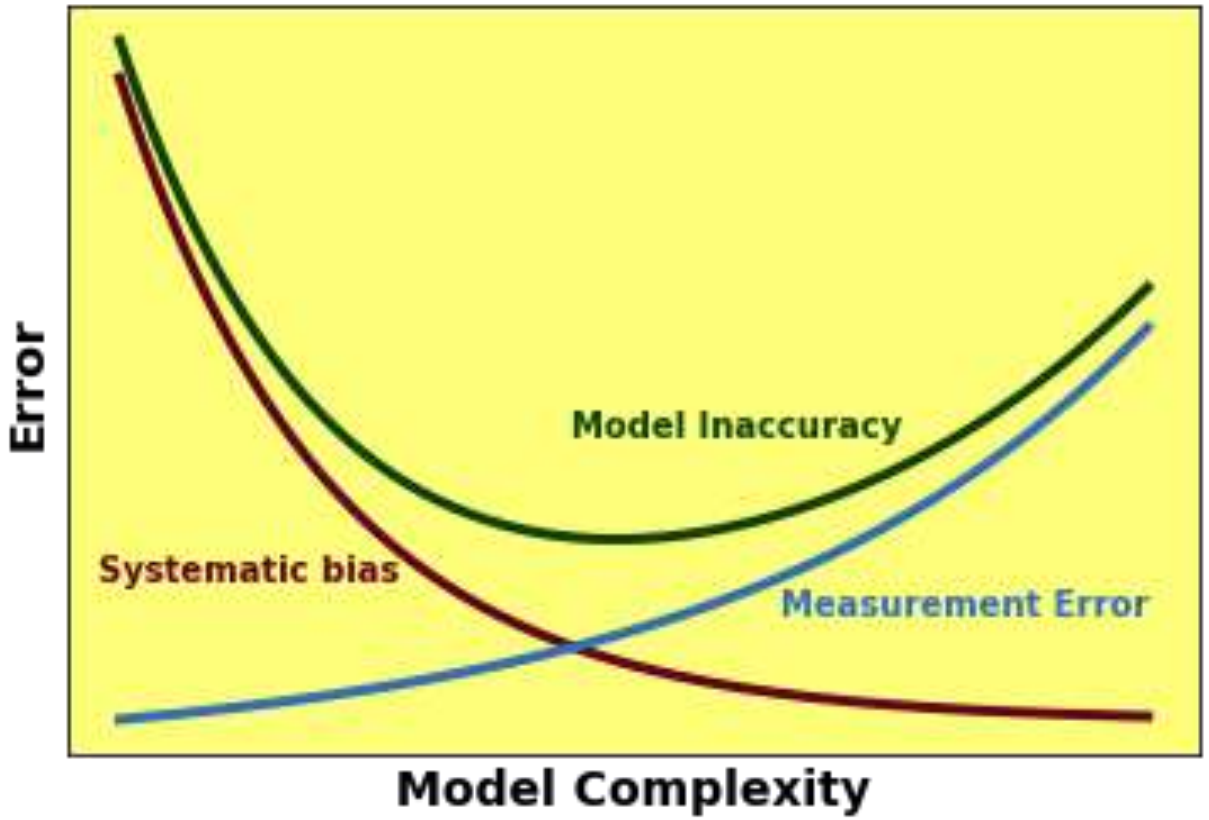
Fourth (badly kept) secret:

There is always one more bug!

=Lubarsky's Law of Cybernetic Entomology

Fifth secret: use SA to calibrate complexity





Presented as ‘Conjecture by O’Neill’

In M. G. Turner and R. H. Gardner, “Introduction to Models” in Landscape Ecology in Theory and Practice, New York, NY: Springer New York, 2015, pp. 63–95.

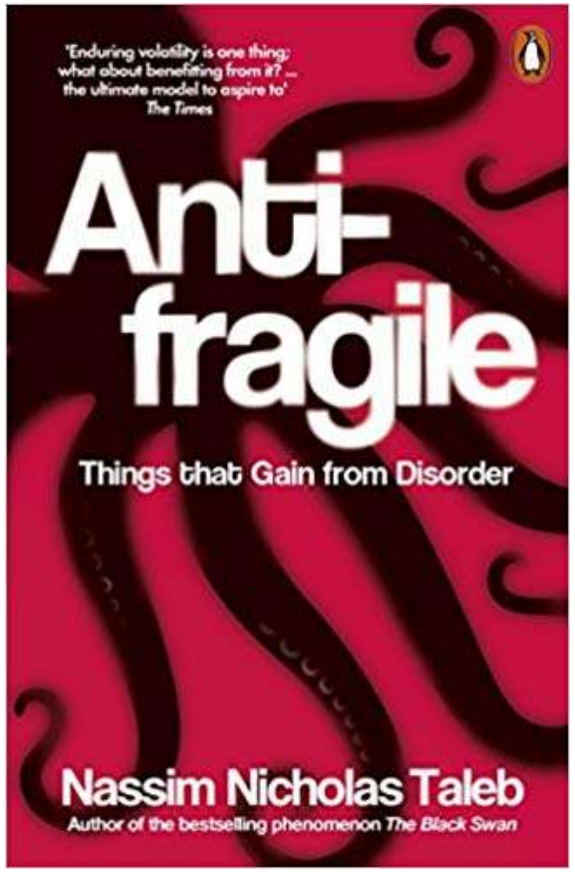
Lofti Aliasker Zadeh

Also known as Zadeh's principle of incompatibility, whereby as complexity increases "precision and significance (or relevance) become almost mutually exclusive characteristics"

L. Zadeh, "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," IEEE Trans. Syst. Man. Cybern., vol. 3, no. 1, pp. 28–44, 1973.

Sixth secret:

With SA it is easier to disprove than to prove; use SA ‘via negativa’:



Doing the right thing

or

Avoiding something wrong?

And of course please don’t run a sensitivity analysis where each factors has a 5% uncertainty

Why?

“[in climate modelling] it looks very little like our idealized image of science, in which pure theory is tested with pure data. [impossible to] eliminate the model-dependency of data or the data-ladenness of models”

Paul N. Edwards, 1999, Global climate science, uncertainty and politics: Data-laden models, model-filtered data.

“[For] philosophers Frederick Suppe and Stephen Norton the blurry model/data relationship pervades all science”

Paul N. Edwards, 1999, Global climate science, uncertainty and politics:

Data-laden models, model-filtered data.

More than a technical uncertainty and sensitivity analysis?

1. Uncertainty and sensitivity analysis (never execute the model once)

2. Sensitivity auditing and quantitative storytelling (investigate frames and motivations)

Saltelli, A., Guimarães Pereira, Â., Van der Sluijs, J.P. and Funtowicz, S., 2013, 'What do I make of your latinorum? Sensitivity auditing of mathematical modelling', Int. J. Foresight and Innovation Policy, (9), 2/3/4, 213–234.

Saltelli, A., Does Modelling need a reformation? Ideas for a new grammar of modelling, available at https://arxiv.org/abs/1712.06457

3. Replace ‘model to predict and control the future’ with ‘model to help mapping ignorance about the future’ …

… in the process exploiting and making explicit the metaphors embedded in the model

J. R. Ravetz, “Models as metaphors,” in Public participation in sustainability science : a handbook, and W. A. B. Kasemir, J. Jäger, C. Jaeger, Gardner Matthew T., Clark William C., Ed. Cambridge University Press, 2003, available at http://www.nusap.net/download.php?op=getit&lid=11

Padilla et al. call for a more structured, generalized and standardized approach to verification

Jakeman et al. call for a 10 points participatory checklist including NUSAP and J. R. Ravetz's process based approach

For NUSAP: Funtowicz, S.O., Ravetz, J.R., 1990. Uncertainty and Quality in Science and Policy. Kluwer, Dordrecht.

J. R. Ravetz, "Integrated Environmental Assessment Forum, developing guidelines for 'good practice', Project ULYSSES", 1997, http://www.jvds.nl/ulysses/eWP97-1.pdf

Sensitivity auditing

EC impact assessment guidelines: what do they say about sensitivity auditing ?

http://ec.europa.eu/smart-regulation/guidelines/docs/br_toolbox_en.pdf

p. 392

… where there is a major disagreement among stakeholders about the nature of the problem, … then sensitivity auditing is more suitable but sensitivity analysis is still advisable as one of the steps of sensitivity auditing.

Andrea Saltelli, Ksenia Aleksankina, William Becker, Pamela Fennell, Federico Ferretti, Niels Holst, Sushan Li, Qiongli Wu, Why so many published sensitivity analyses are false: a systematic review of sensitivity analysis practices, Environmental Modelling and Software, Volume 114, April 2019, Pages 29-39.

Sensitivity auditing, [...] is a wider consideration of the effect of all types of uncertainty, including structural assumptions embedded in the model, and subjective decisions taken in the framing of the problem.
[...]
The ultimate aim is to communicate openly and honestly the extent to which particular models can be used to support policy decisions and what their limitations are.

p. 393

“In general sensitivity auditing stresses the idea of honestly communicating the extent to which model results can be trusted, taking into account as much as possible all forms of potential uncertainty, and to anticipate criticism by third parties.”

The rules of sensitivity auditing

Rule 1: Check against rhetorical use of mathematical modelling;

Rule 2: Adopt an “assumption hunting” attitude; focus on unearthing possibly implicit assumptions;

Rule 3: Check if uncertainty been instrumentally inflated or deflated.

The rules of sensitivity auditing

Rule 4: Find sensitive assumptions before these find you; do your SA before publishing;

Rule 5: Aim for transparency; Show all the data;

Rule 6: Do the right sums, not just the sums right; the analysis should not solve the wrong problem;

Rule 7: Perform a proper global sensitivity analysis.

The rules of sensitivity auditing ca be used as columns for NUSAP pedigree matrix

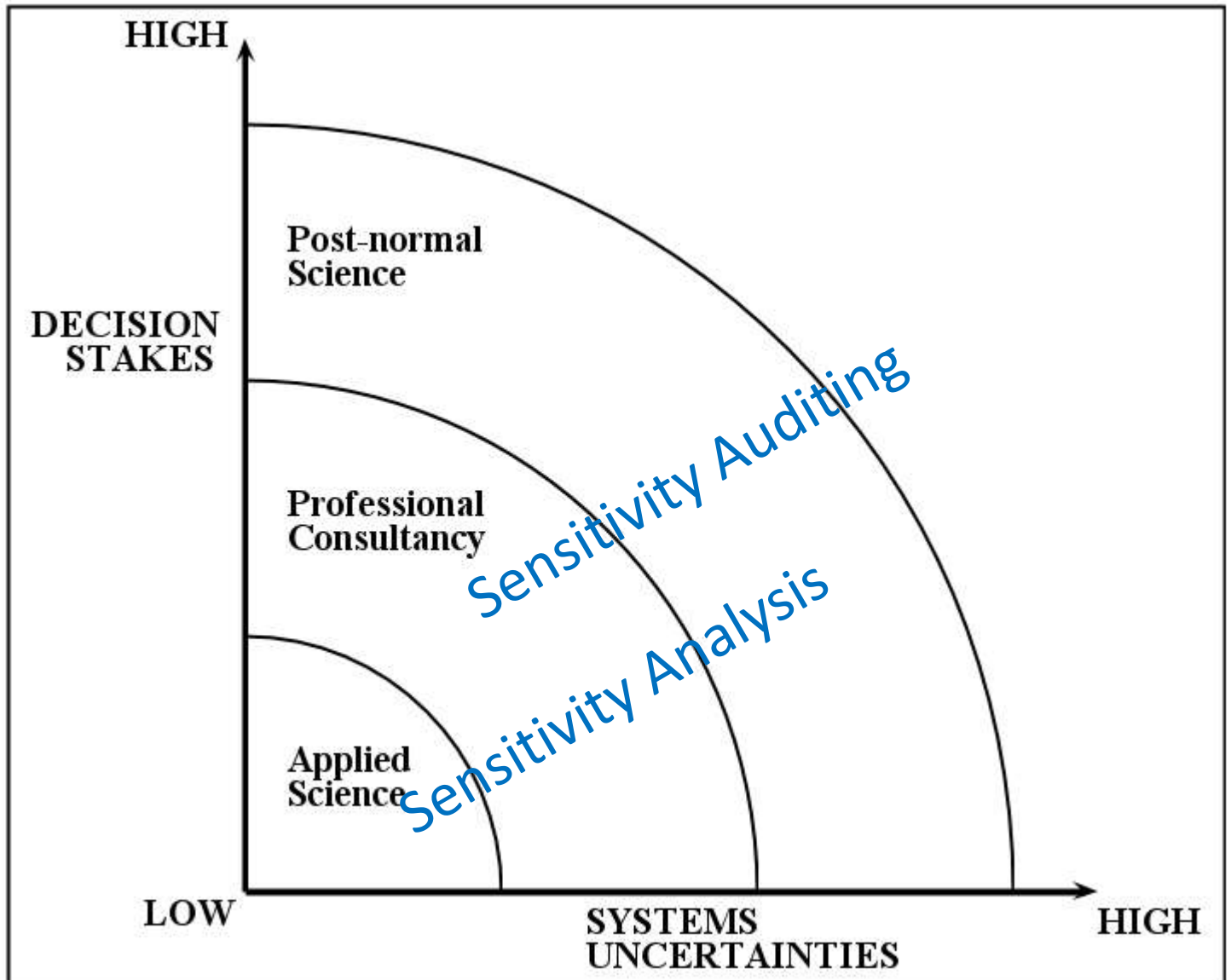
Universiteit Utrecht

Example Pedigree matrix parameter strength

Code	Proxy	Empirical	Theoretical basis	Method	Validation
4	Exact measure	Large sample direct mmts	Well established theory	Best available practice	Compared with indep. mmts of same variable
3	Good fit or measure	Small sample direct mmts	Accepted theory partial in nature	Reliable method commonly accepted	Compared with indep. mmts of closely related variable
2	Well correlated	Modeled/derived data	Partial theory limited consensus on reliability	Acceptable method limited consensus on reliability	Compared with mmts not independent
1	Weak correlation	Educated guesses / rule of thumb est	Preliminary theory	Preliminary methods unknown reliability	Weak / indirect validation
0	Not clearly related	Crude speculation	Crude speculation	No discernible rigour	No validation

Jeroen van der Sluijs

http://www.nusap.net/


HIGH
DECISION STAKES
Post-normal Science
Professional Consultancy
Applied Science
Sensitivity Auditing
Sensitivity Analysis
LOW
SYSTEMS UNCERTAINTIES
HIGH

Some examples:

Sensitivity analysis: the case of the Stern review

Global Environmental Change 20 (2010) 298–302

Contents lists available at ScienceDirect

Global Environmental Change

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

Sensitivity analysis didn't help. A practitioner's critique of the Stern review

Andrea Saltelli *, Beatrice D'Hombres

Joint Research Centre, Institute for the Protection and Security of the Citizen, Ispra, Italy

The case of Stern's Review – Technical Annex to postscript

Nicholas Stern, London School of Economics

William Nordhaus, University of Yale

Stern, N., Stern Review on the Economics of Climate Change. UK Government Economic Service, London, www.sternreview.org.uk.

Nordhaus W., Critical Assumptions in the Stern Review on Climate Change, SCIENCE, 317, 201-202, (2007).

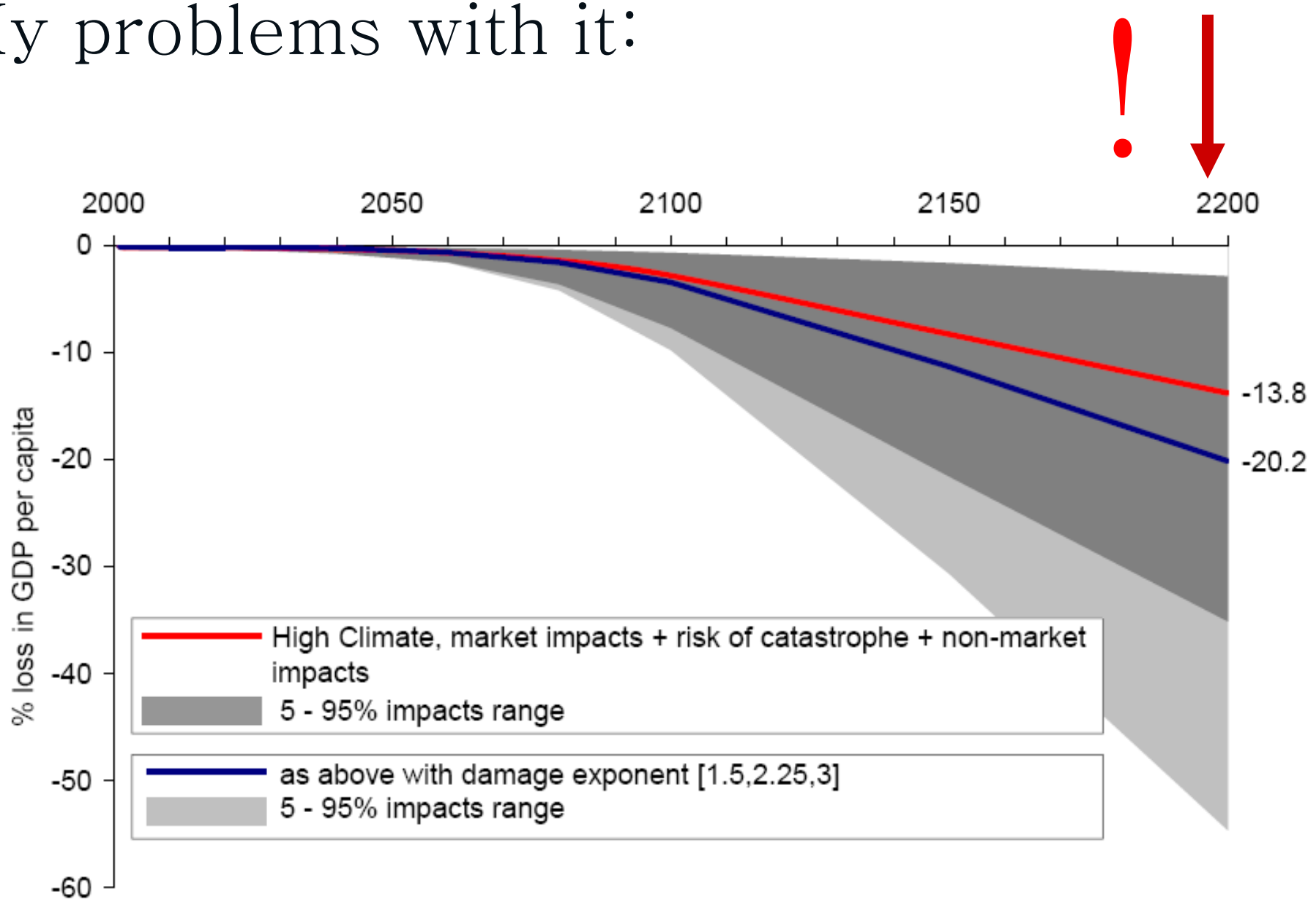
The Stern – Nordhaus exchange on *SCIENCE*

1) Nordhaus falsifies Stern based on 'wrong' range of discount rate

2) Stern's complements its review with a postscript: a sensitivity analysis of the cost benefit analysis

3) Stern infers: My analysis shows robustness'

My problems with it:

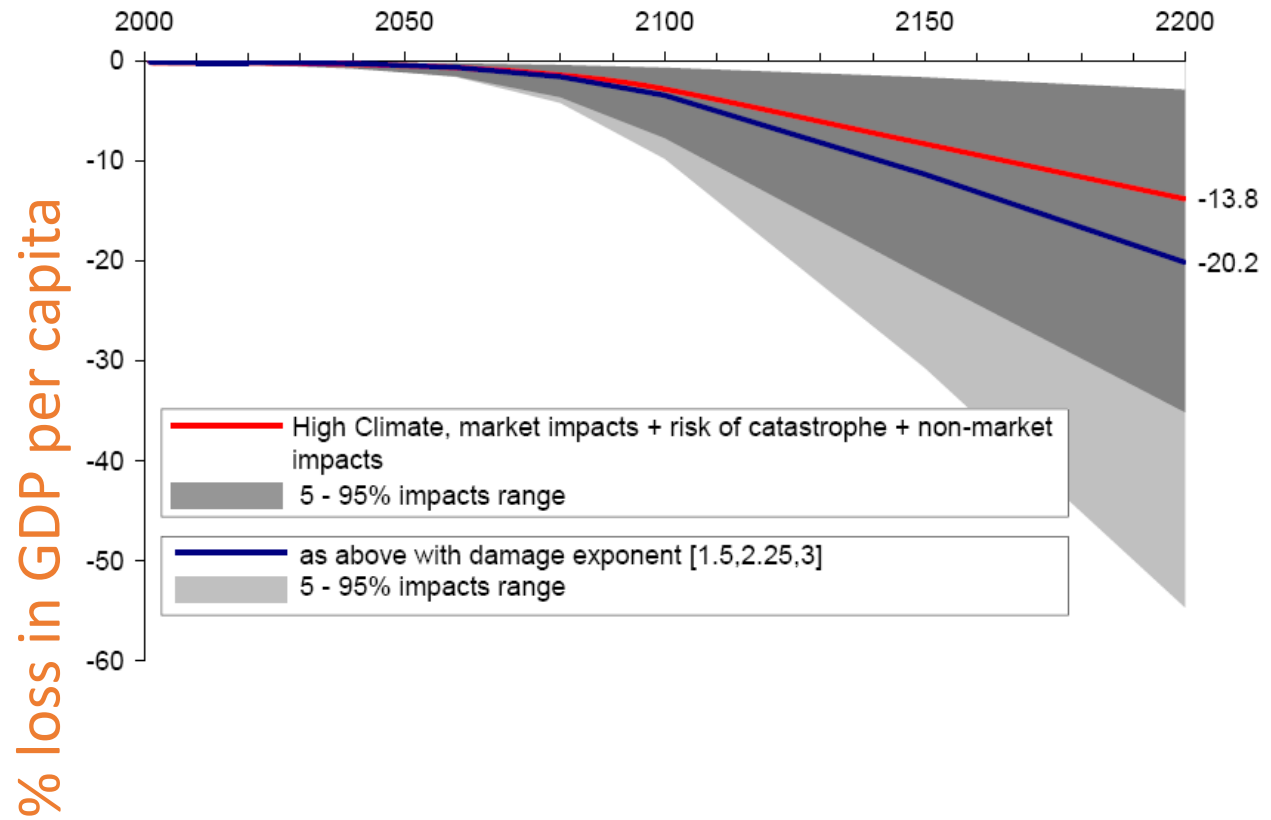


… but foremost Stern says:

changing assumptions → important effect

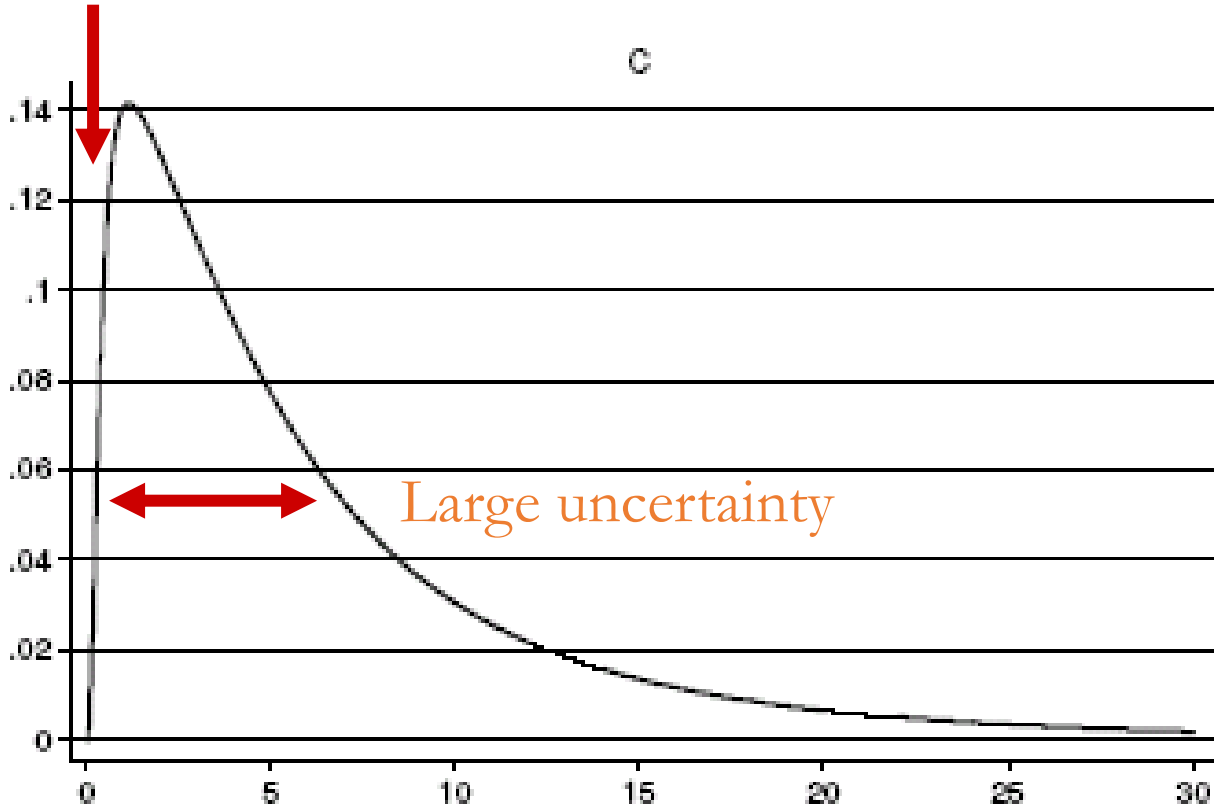
when instead he should admit that:

changing assumptions → all changes a lot

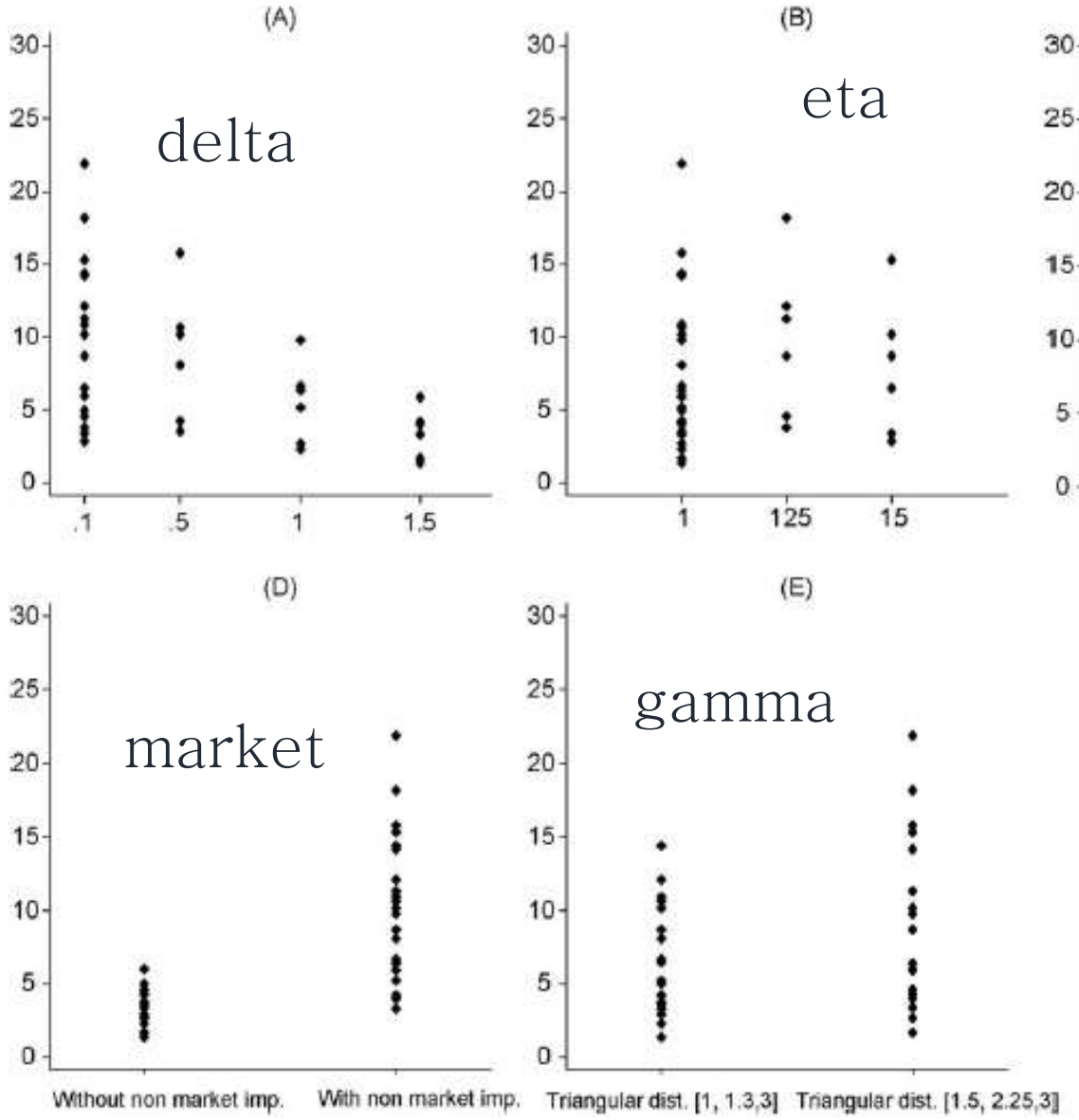


How was it done? A reverse engineering of the analysis

Missing points

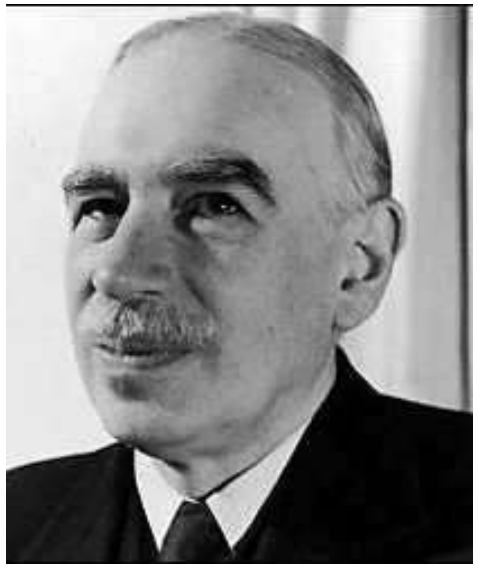


% loss in GDP per capita



Sensitivity analysis, also by reverse engineering

Same criticism applies to Nordhaus – both authors frame the debate around numbers which are …



… precisely wrong

From: Saltelli, A., D'Hombres, 2010, Sensitivity analysis didn't help. A practitioner's critique of the Stern review, *GLOBAL ENVIRONMENTAL CHANGE*, 20, 298-302.

The End

@andreasaltelli

Some examples:

Sensitivity auditing: the OECD PISA study

Do PISA data justify PISA-based education policy?

International Journal of Comparative Education and Development
Vol. 19 No. 1, 2017
pp. 1-17

2396-7404
DOI 10.1108/IJCED-12-2016-0023

THE CONVERSATION

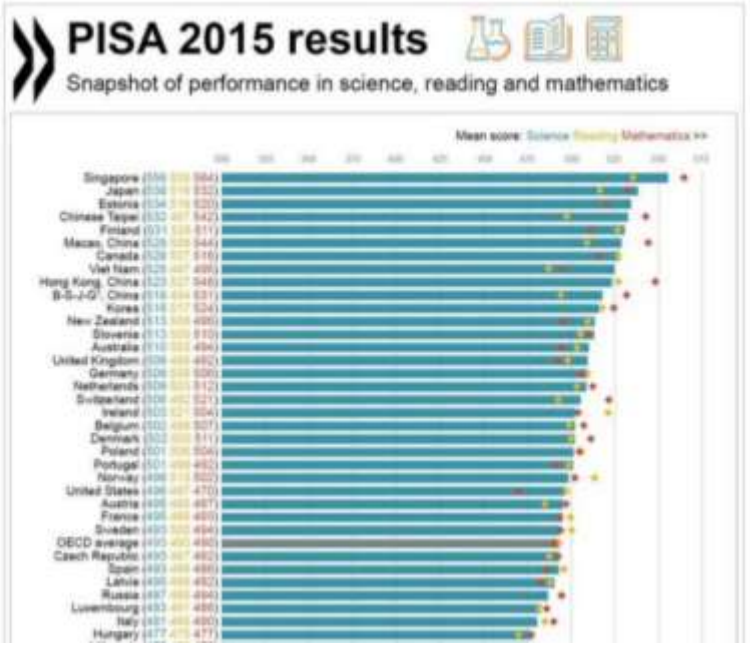
Academic rigour, journalistic flair

Arts + Culture | Business + Economy | Cities | Education | Environment + Energy | FactCheck | Health + Medicine | Politics + Society | Science + Technology

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

June 12, 2017 3.55pm AEST

Chemistry class at the Dong Tien Secondary School, Thai Nguyen Province, Vietnam. Asian Development Bank/flickr, CC BY-SA



A condensed version of the article

With PISA the OECD gained the centre-stage in the international arena on education policies, which led to important controversies

http://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics

theguardian

OECD and Pisa tests are damaging education worldwide - academics

In this letter to Dr Andreas Schleicher, director of the OECD's Programme for International Student Assessment, academics from around the world express deep concern about the impact of Pisa tests and call for a halt to the next round of testing

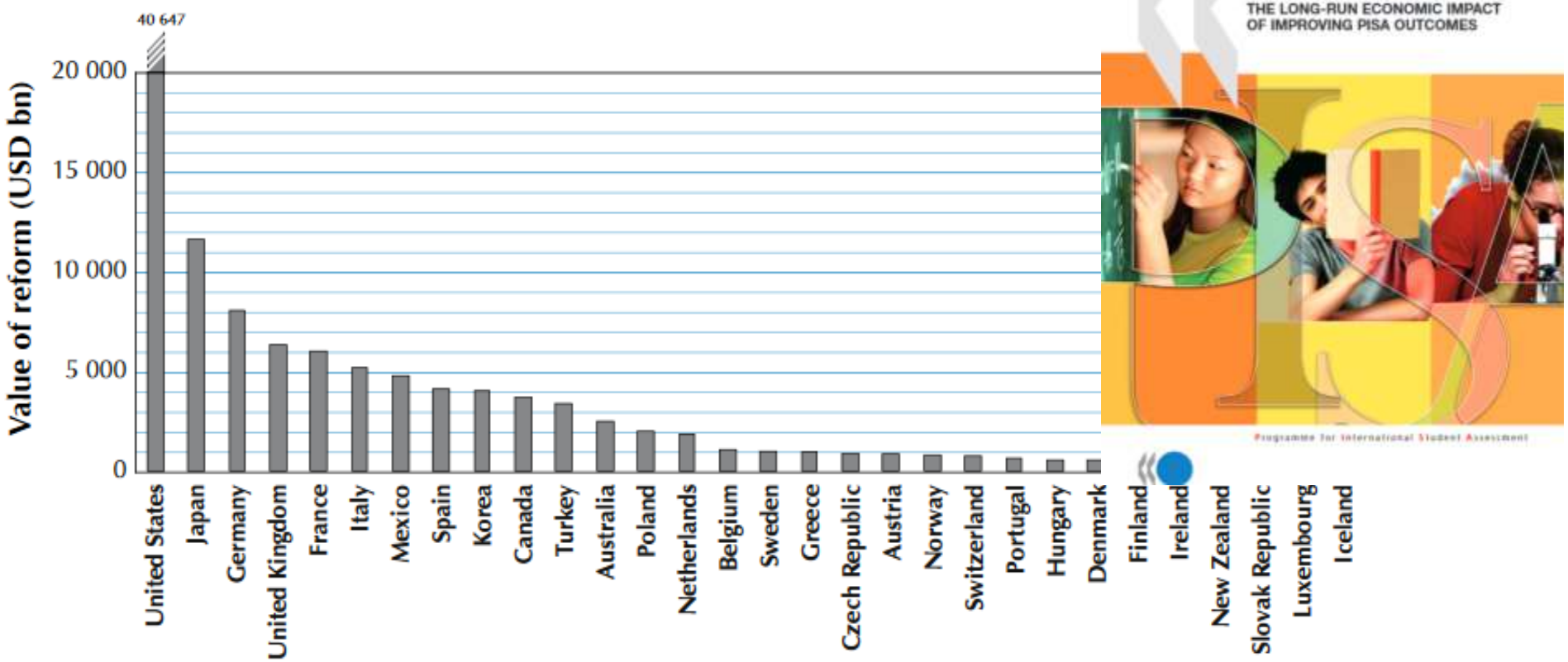
School children in Sichuan province in China. Academics say the OECD should develop alternatives to league tables and find more meaningful ways of reporting assessment, taking account of different cultures. Photograph: James Zeng Huang/Corbis Sygma

Critical remarks by the 80 signatories of the letter:

- Flattening of curricula (exclusion of subjects)
- Short-termism (teaching to the test)
- Promoting “life skills to function in knowledge societies”
- Stressing the student
- … ➔ Stop the test!
- A more participatory run of the study would be advisable

Figure 1

Present value of Scenario I (improve student performance in each country by 25 points on the PISA scale) in billion USD (PPP)



Note: Discounted value of future increases in GDP until 2090 due to reforms that improve student performance in each

http://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/thehighcostofloweducationalperformance.htm

PISA's daring quantifications:

"If every EU Member State achieved an improvement of 25 points in its PISA score (which is what for example Germany and Poland achieved over the last decade), 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, L. (2014), "The economic case for education", EENEE Analytical Report 20, European Expert Network on Economics of Education (EENEE), Institute and University of Munich.

Our study identifies both technical and normative issues:

1) Non response bias (what students are excluded; PISA non-response for England: the bias turned out to be twice the size of the OECD declared standard error in 2003.

2) Non open data, which makes SA impossible

Our study identifies both technical and normative issues:

3) Flattening curricula (do all countries wish to prosper by becoming knowledge societies?)
4) Power implications: power in the use of evidence. OECD (unelected officers and scholars) becoming a global super-ministry of education

Practicum

Grade a set of questions using a Likert scale

Likert scale

5. Strongly agree
4. Agree
3. Neutral
2. Disagree
1. Strongly disagree

A. Our duty is to provide objective numbers to policy makers. A cost benefit analysis is useful to make sure that taxpayer money is well spent.

B. Given proper statistical tools it is always possible to arrive at a number quantifying our present state of knowledge.

C. Numbers should be objective and not the result of ‘stealth advocacy’.

D. Numbers can convey a misleading impression of accuracy and precision.

E. The analyst should strive to highlight the difference between risk and uncertainty.

F. The analyst should strive to identify different values underpinning different framing of the issue.

The End

@andreasaltelli