

Course: Numbers for policy: Practical problems in quantification: approaches

Universitat Oberta de Catalunya, Edifici B3, Parc Mediterrani
de la Tecnologia, Avinguda Carl Friedrich Gauss, 5, 08860,
Castelldefels (Barcelona), November 18–21

Andrea Saltelli

Centre for the Study of the Sciences and the Humanities,
University of Bergen, and Open Evidence Research, Open
University of Catalonia



Something general about
mathematical modelling

Caeteris are
never paribus

The case of DSGE, dynamic stochastic general equilibrium models

Philip Mirowski



Philip Mirowski, 2013, Never let a serious crisis go wasted, Verso Books.

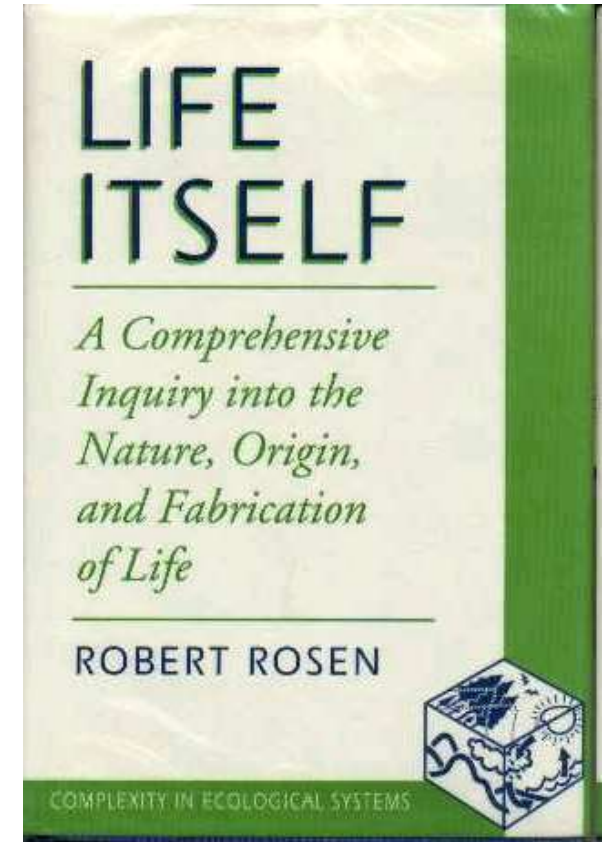
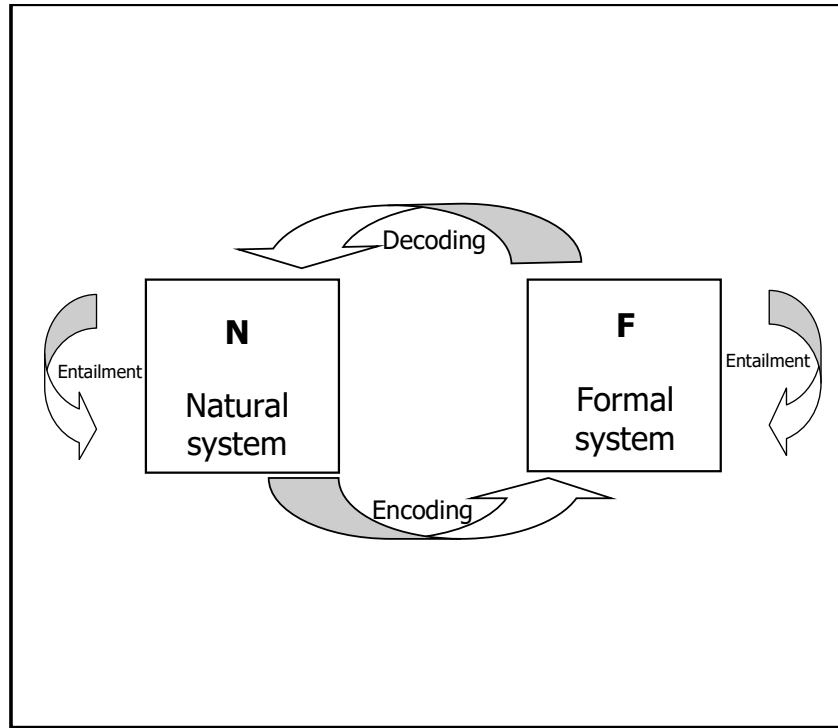
Models have
little memory

“[...] 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.”

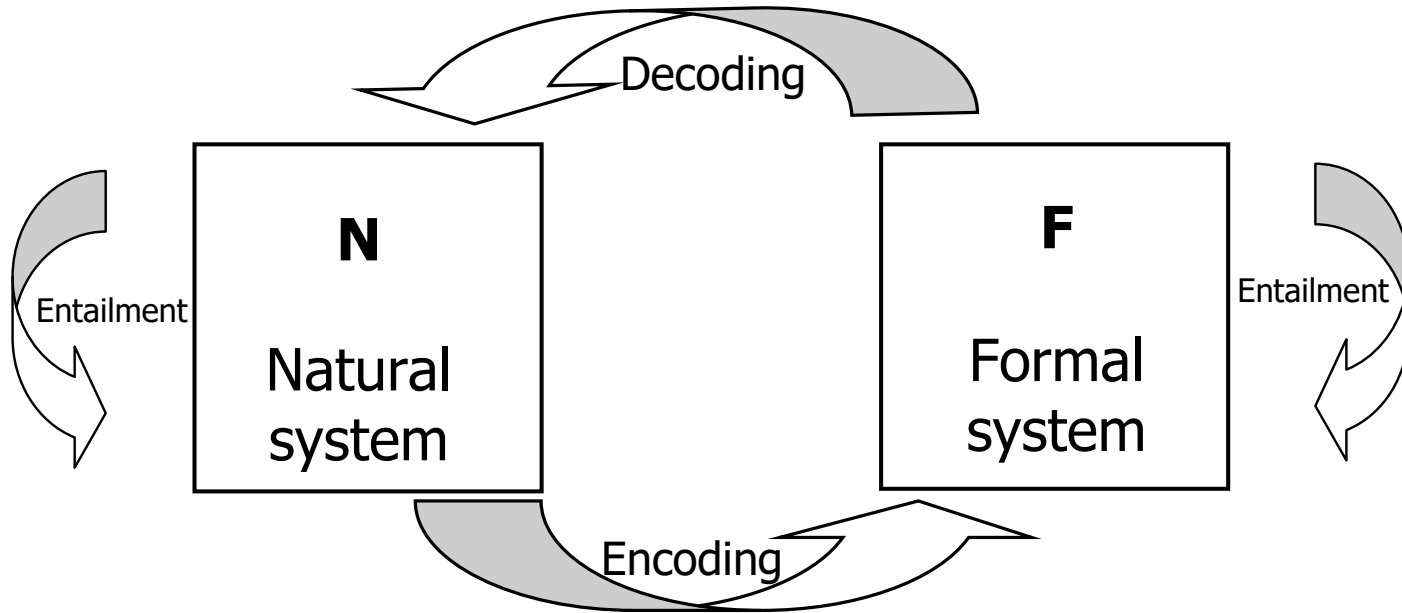
E. Millgram The Great Endarkenment, p. 29

Modelling is a
craft more than
a science

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

Can models be
falsified?

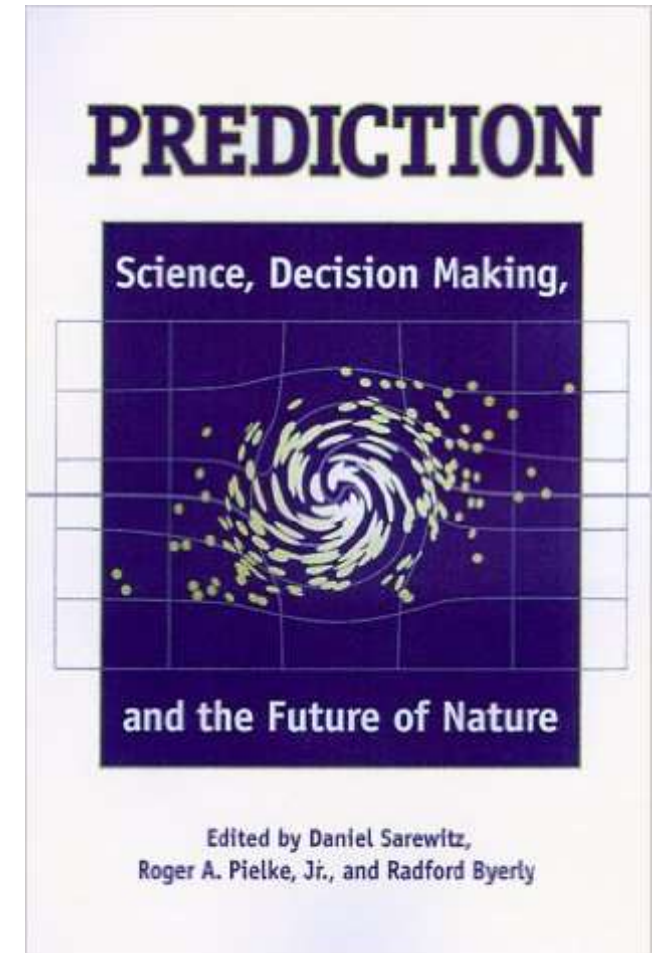
“models are most useful when they are used to challenge existing formulations, rather than to validate or verify them”



Naomi
Oreskes

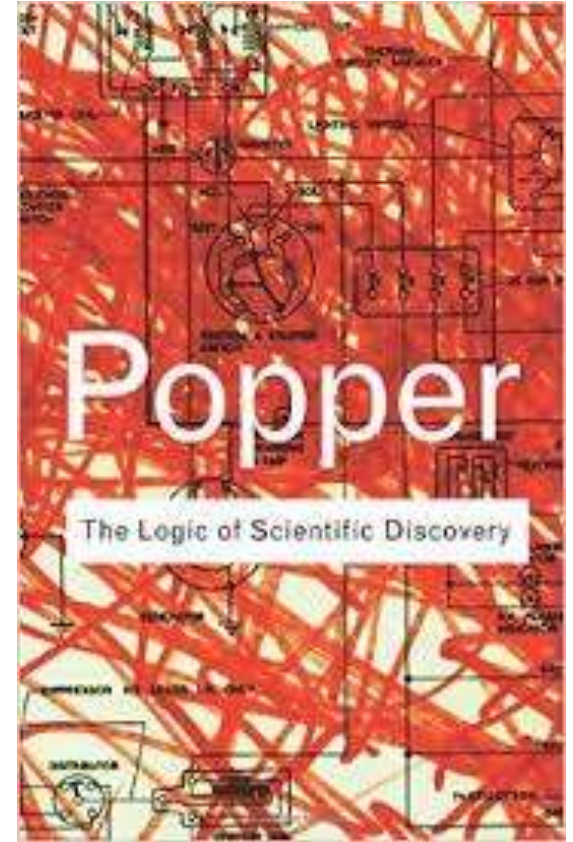
N. Oreskes, K. Shrader-Frechette, and K. Belitz, “Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences,” *Science*, 263, no. 5147, 1994.

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

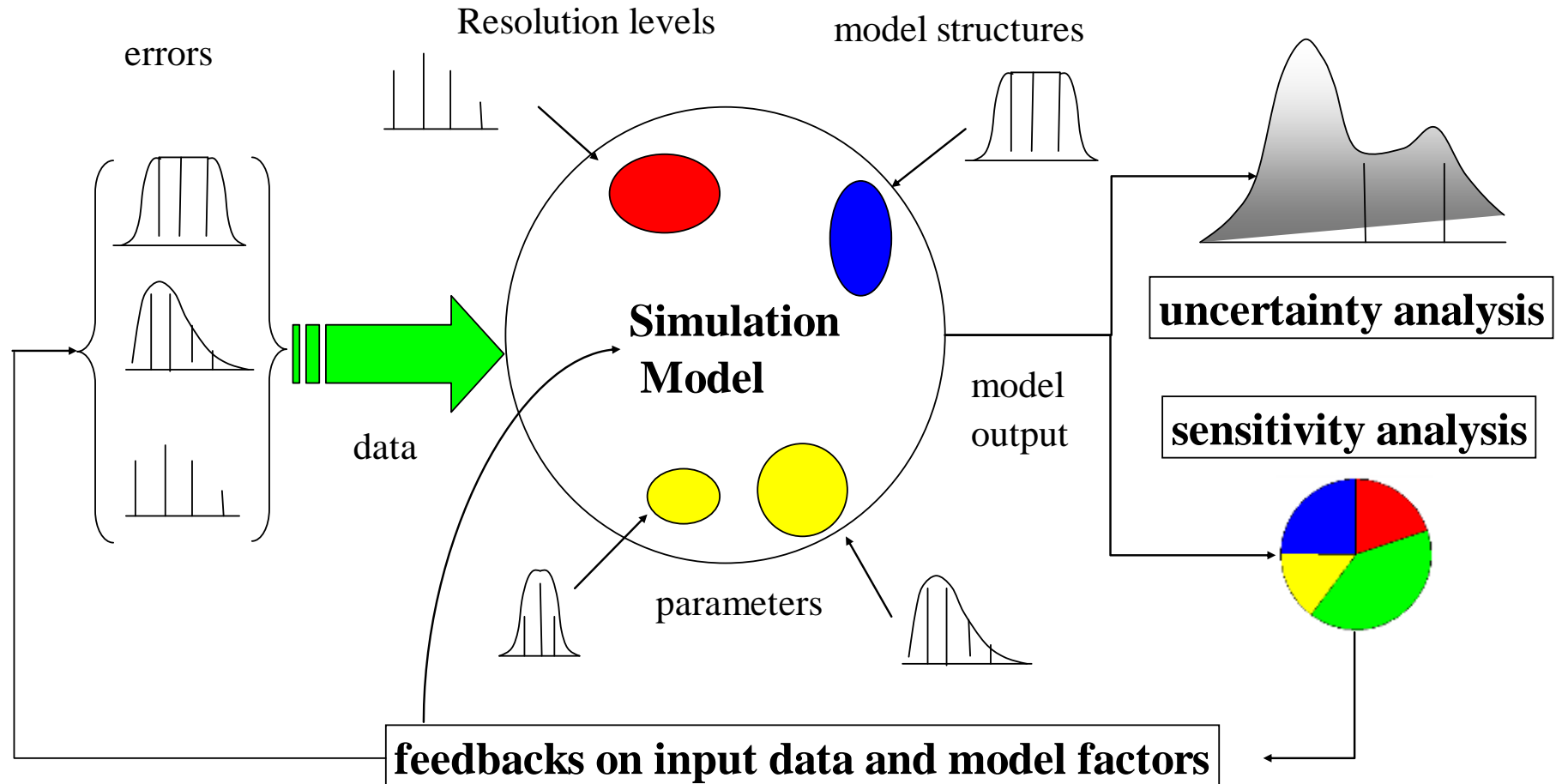
Journey Purpose	Weekday					Weekend	All Week
	7am-10am	10am-4pm	4pm-7pm	7pm-7am	Weekday Average		
Work	-0.48	-0.4	-0.62	-0.5	-0.44	-0.48	-0.45
Non - Work (commuting and other)	-0.67	-0.65	-0.53	-0.47	-0.59	-0.52	-0.56

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

An engineer's vision of UA, SA

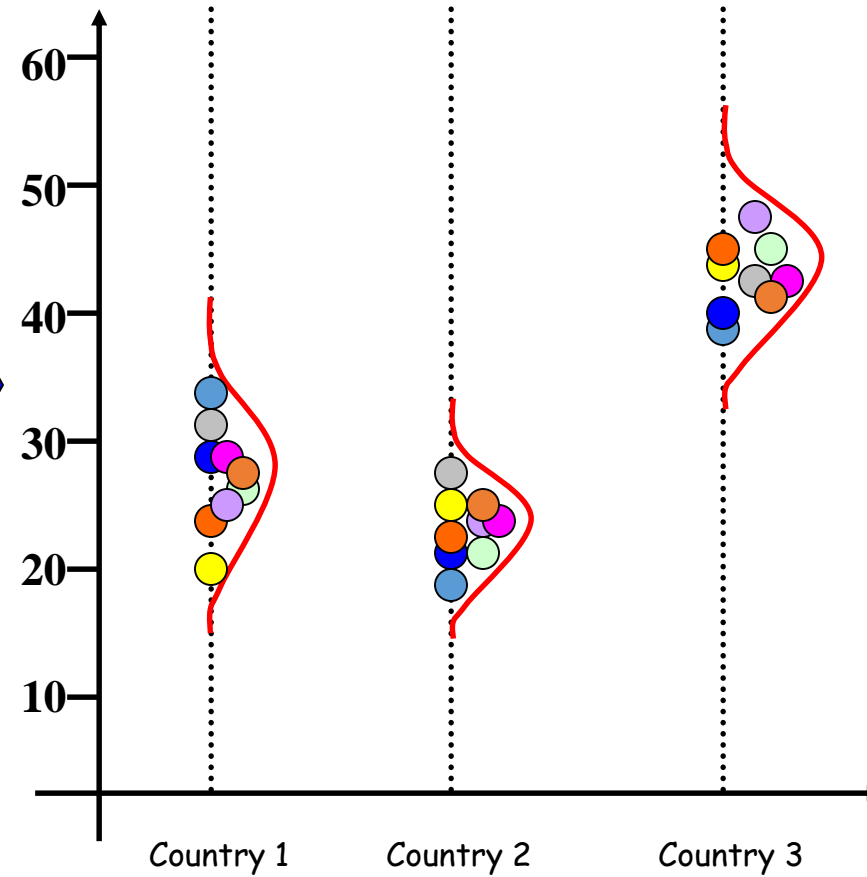
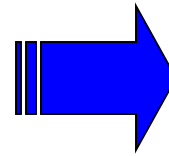
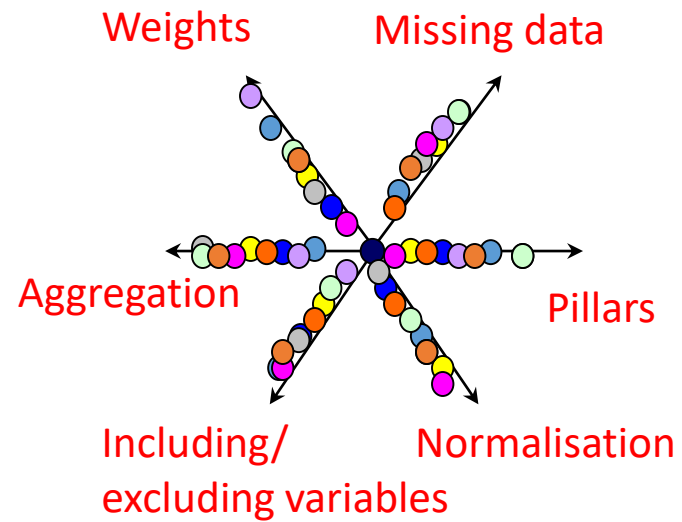


One can sample more than just factors:

- modelling assumptions,
- alternative data sets,
- resolution levels,
- scenarios ...

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



Why Sensitivity analysis?

It is in the guidelines!

European Commission, 2015

Office for the Management and Budget, 2006


Environmental Protection Agency, 2009

EPA, 2009, March. Guidance on the Development, Evaluation, and Application of Environmental Models. Technical Report EPA/100/K-09/003. Office of the Science Advisor, Council for Regulatory Environmental Modeling, <http://nepis.epa.gov/Exe/ZyPDF.cgi?Dockkey=P1003E4R.PDF>, Last accessed December 2015.

EUROPEAN COMMISSION, Better regulation toolbox, appendix to the Better Regulation Guidelines, Strasbourg, 19.5.2015, SWD(2015) 111 final, COM(2015) 215 final, http://ec.europa.eu/smart-regulation/guidelines/docs/swd_br_guidelines_en.pdf.

OMB, Proposed risk assessment bulletin, Technical report, The Office of Management and Budget's – Office of Information and Regulatory Affairs (OIRA), January 2006, https://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/proposed_risk_assessment_bulletin_010906.pdf, pp. 16–17, accessed December 2015.

Press | Archives | Sitemap | About this site | Legal notice | Contact | Search | English (en) ▼



Better Regulation

European Commission > Better Regulation > Guidelines

Home

REFIT

Stakeholder consultations

Roadmaps / Inception Impact Assessments

Impact Assessment

Evaluation

Regulatory Scrutiny Board

Guidelines

■ Better Regulation Guidelines

■ Better Regulation "Toolbox"

Key documents

Better Regulation Guidelines

These guidelines explain what Better Regulation is and how it should be applied in the day to day practices when preparing new initiatives and proposals or managing existing policies and legislation.

They cover the whole policy cycle, from policy preparation and adoption to implementation and application, to evaluation and revision of EU law. For each of these phases there are a number of Better Regulation principles, objectives, tools and procedures to make sure that the EU has the best regulation possible. These relate to planning, impact assessment, stakeholder consultation, implementation and evaluation.

The [Better Regulation Guidelines](#) are structured into chapters which cover each of the instruments of the law-making process. The corresponding [toolbox](#) gives more detailed and technical information.

Better Regulation Guidelines are based on the outcomes of public consultation exercises carried out in 2013 and 2014.

- [Public consultation on the revision of the Commission's Impact Assessment Guidelines](#)
- [Stakeholder Consultation Guidelines](#)
- [Consultation on the draft Commission Evaluation Policy Guidelines](#)

Share

Search

Stay connected

Facebook Twitter YouTube

Blog

Latest documents

- [19/05/2015 - Better Regulation Package](#)

Help us improve

Find what you wanted?

Yes ☐ No ☐

What were you looking for?

Any suggestions?

Send

Last update: 11/08/2015 | [Legal notice](#) | [Cookies](#) | [Contact](#) | [Search](#) | [Top](#)

<http://ec.europa.eu/smart-regulation/>

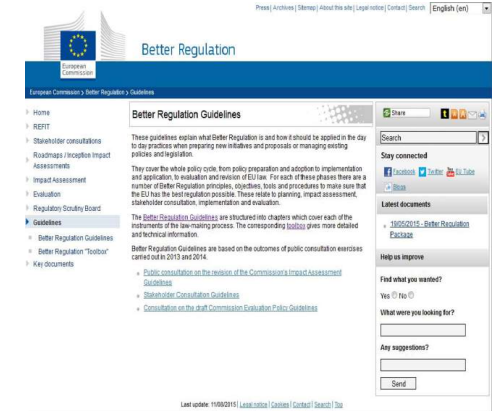
Source: IA Toolbox,
p. 391

4. SENSITIVITY AND UNCERTAINTY ANALYSES

Page 391

Six steps for a global SA:

1. Select one 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 both 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.



The screenshot shows the 'Better Regulation' section of the European Commission's website. The header includes the European Commission logo and the text 'Better Regulation'. Below the header, there is a navigation menu with links to Home, REFIT, Stakeholder consultations, Roadmaps / Inception Impact Assessments, Impact Assessment, Evaluation, Regulatory Scrutiny Board, and Guidelines. The 'Guidelines' link is highlighted. The main content area is titled 'Better Regulation Guidelines' and contains a search bar, social media links (Facebook, Twitter, YouTube, LinkedIn), and a list of latest documents, including '19/05/2015 - Better Regulation Package'. The footer includes a 'Help us improve' section with a feedback form and a 'Last update' notice.

Why Sensitivity analysis?

It can answer interesting
questions

EVIDENCE,
ARGUMENT, &
PERSUASION IN
THE POLICY
PROCESS

GIANDOMENICO
MAJONE

"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?"

Why sensitivity analysis?

It can detect garbage in garbage out
(GIGO)

Funtowicz & Ravetz's GIGO (Garbage In, Garbage Out) Science “where uncertainties in inputs must be suppressed least outputs become indeterminate”



Leamer's “Conclusions are judged to be sturdy only if the neighborhood of assumptions is wide enough to be credible and the corresponding interval of inferences is narrow enough to be useful”

S. Funtowicz and J. R. Ravetz, *Uncertainty and Quality in Science for Policy*. Dordrecht: Kluwer, 1990; E. E. Leamer, “Sensitivity Analyses Would Help,” *Am. Econ. Rev.*, vol. 75, no. 3, pp. 308–313, 1985.

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



William Nordhaus,
University of Yale
Nobel 'Economics'
2018



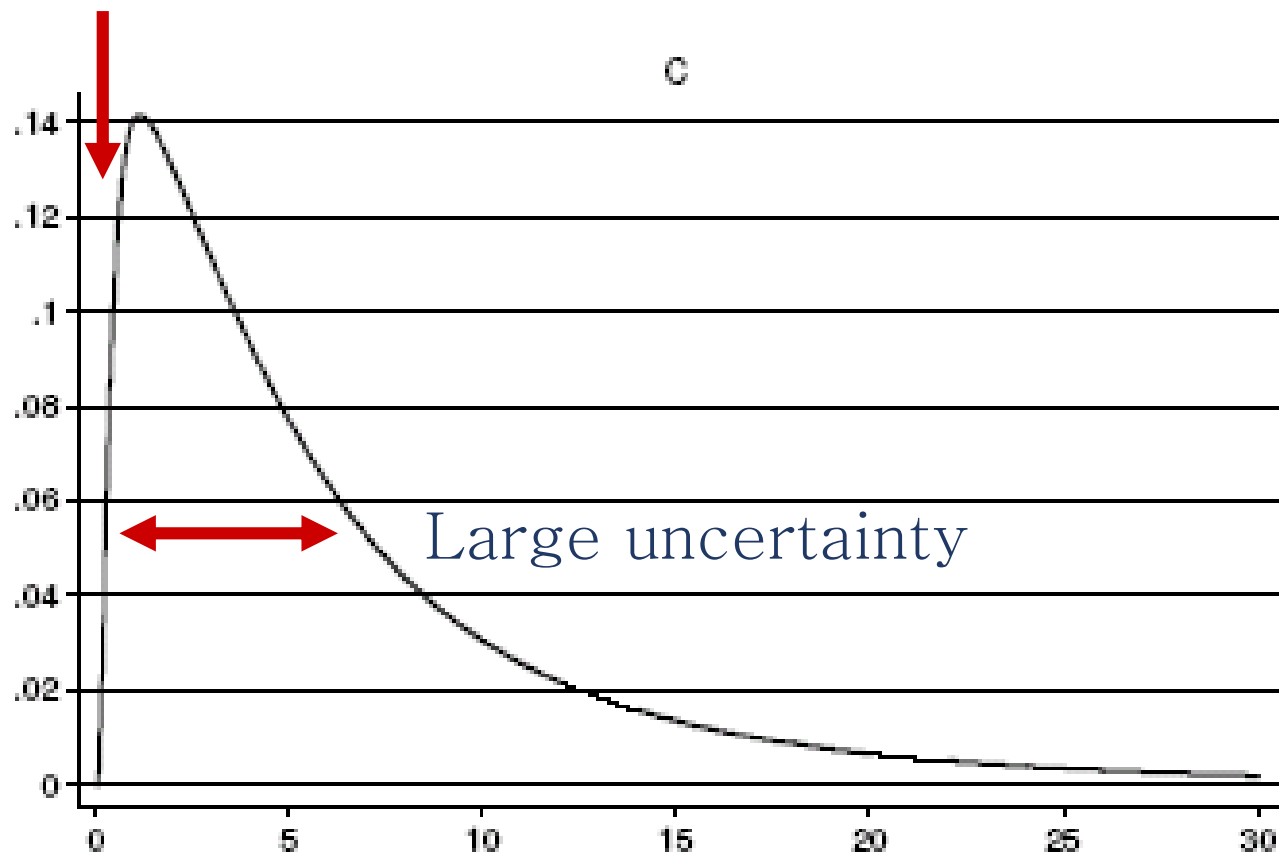
Nicholas Stern, London School
of Economics

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).

How was it done? A reverse engineering of the analysis

Missing points



% loss in GDP per capita

Why sensitivity analysis?

It allows interesting discoveries

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

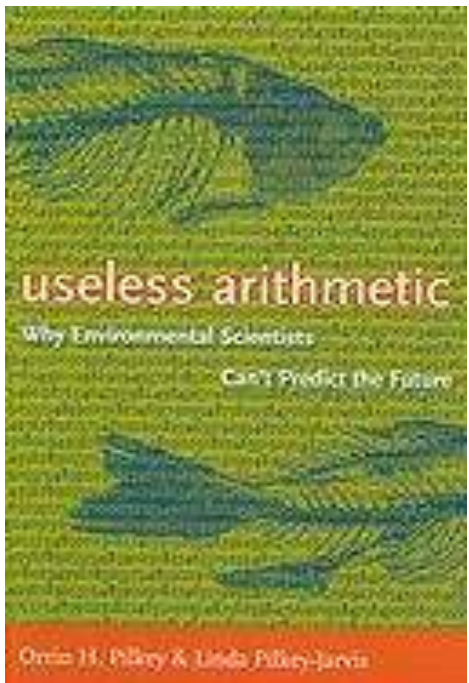
European Commission, Ispra, Italy

Why sensitivity
analysis

Limits of sensitivity analysis

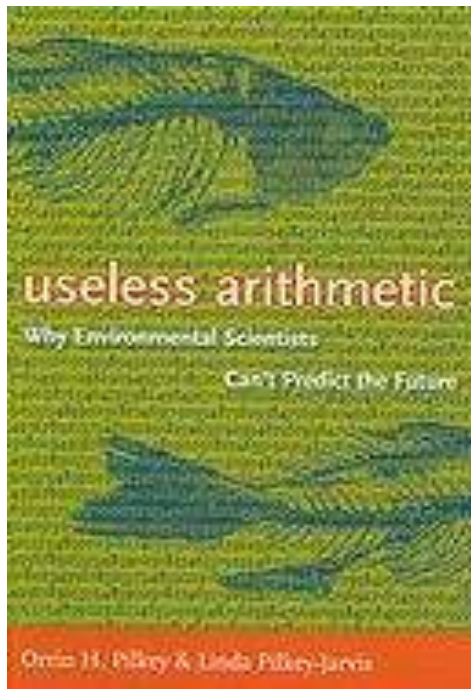


Orrin H.
Pilkey



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

The map is not the
territory

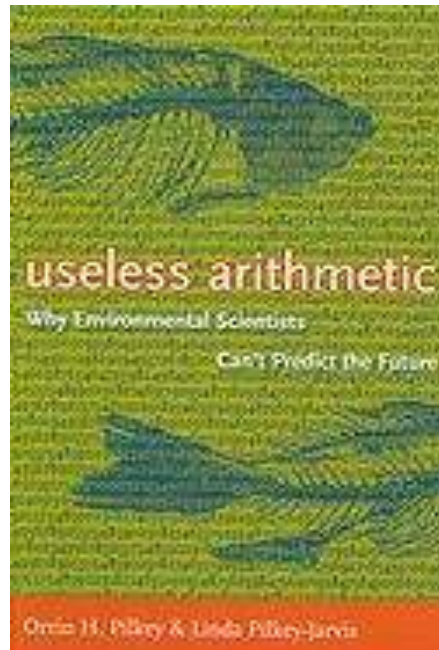


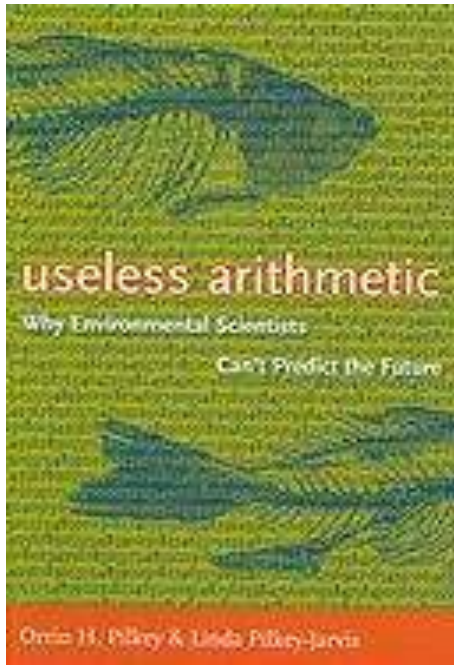
<<It is important, however, to recognize that the sensitivity of the parameter in the equation is what is being determined, not the sensitivity of the parameter in nature.

[...] 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.

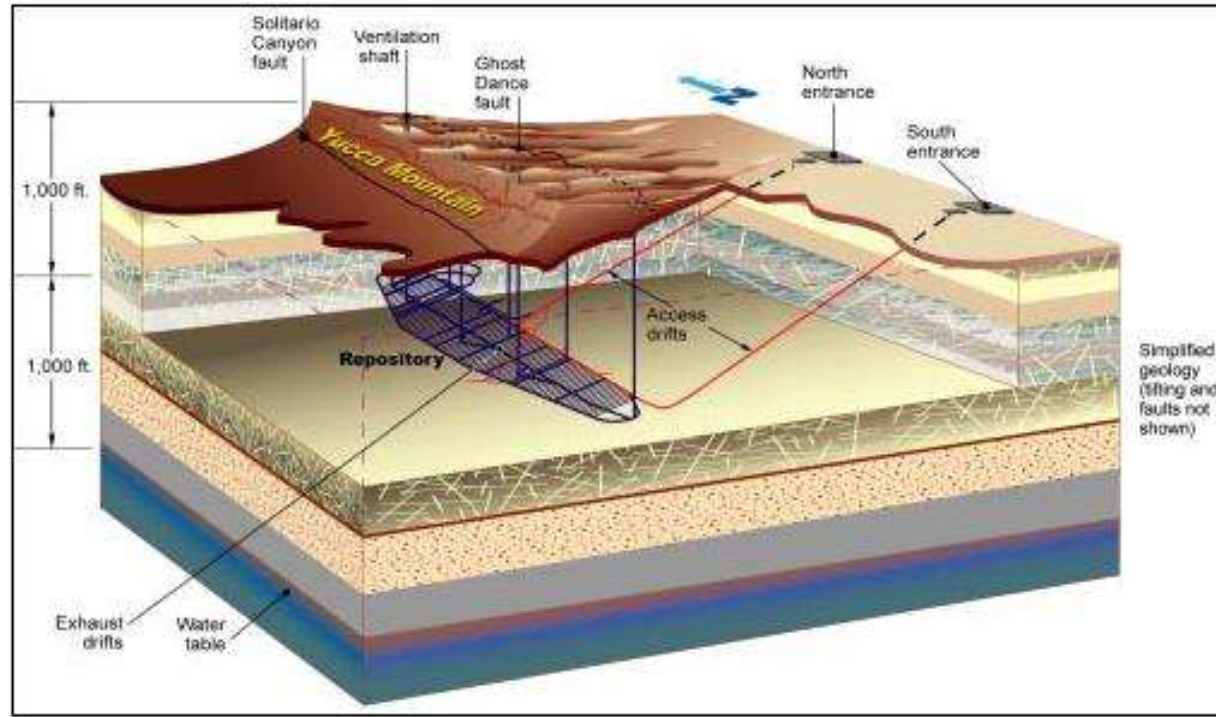
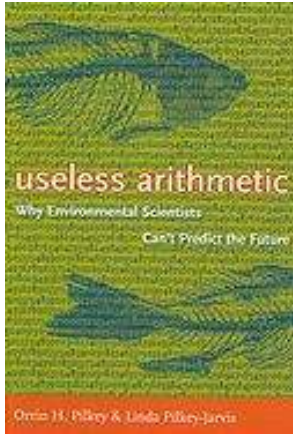
TSPA is Composed of 286 sub-models.





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.



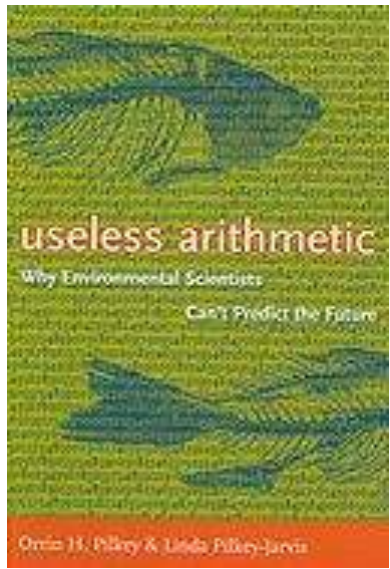


The confidence of the stakeholders in TSPA was not helped when evidence was produced which could lead to an upward revision of 4 orders of magnitude of this parameter
(the ^{36}Cl story)

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.



“Scientific mathematical modelling should involve constant efforts to falsify the model”

➔ Organized skepticism



Organized **S**kepticism – all ideas must be tested and are subject to rigorous, structured community scrutiny

Robert K. Merton
(1910–2003)

Communalism – the common ownership of scientific knowledge according to which scientists give up intellectual property in exchange for recognition and esteem (Merton actually used the term Communism, but had this notion of communalism in the context of Marxism);

Universalism – the use of the same terms of unqualified truth, regardless of race, class, or religion;

CUDOS

ruth and
not of

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

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



Robert K.
Merton

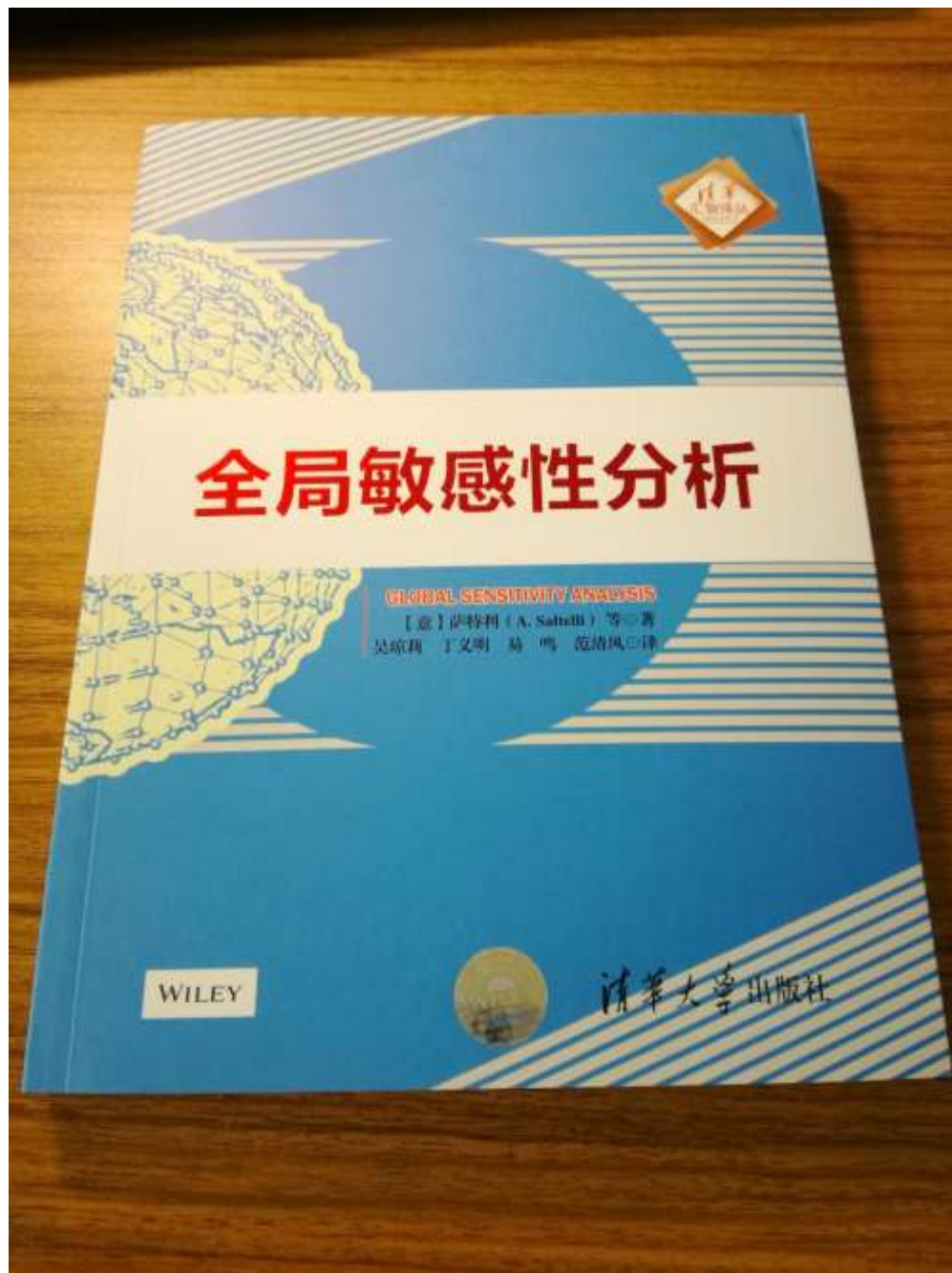
Where to study
sensitivity analysis?

A. Saltelli, M. Ratto,
T. Andres, F. Campolongo,
J. Cariboni, D. Gatelli,
M. Saisana, S. Tarantola

GLOBAL SENSITIVITY ANALYSIS

The Primer

 WILEY



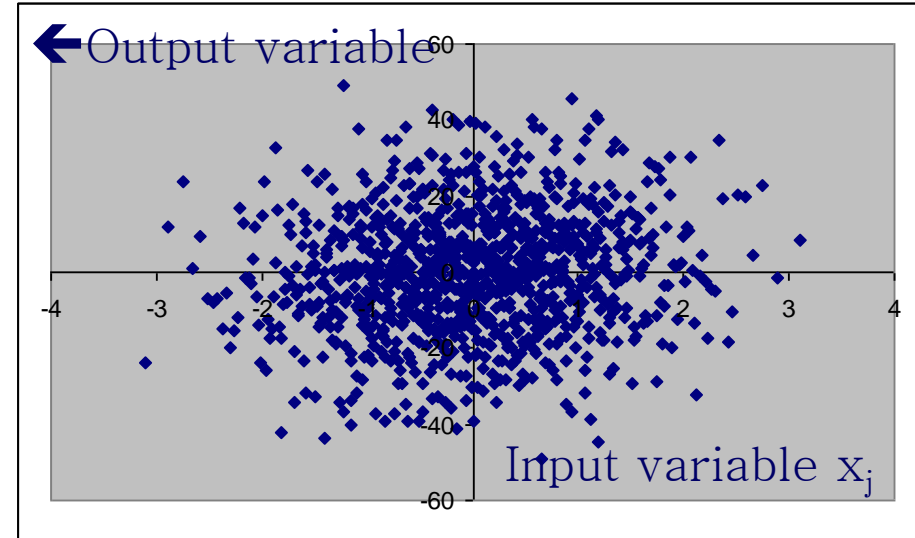
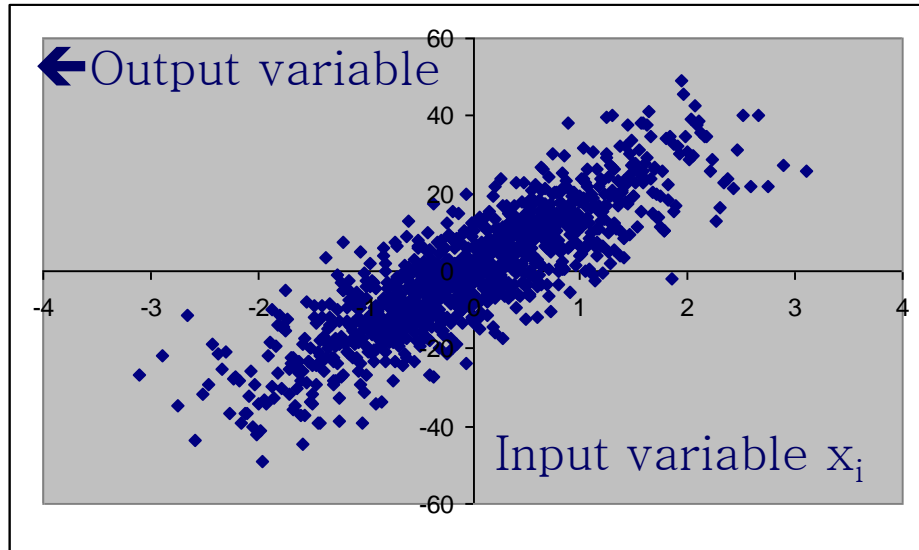


A. Saltelli, M. Ratto,
T. Andres, F. Campolongo,
J. Cariboni, D. Gatelli,
M. Saisana, S. Tarantola

Available for free at

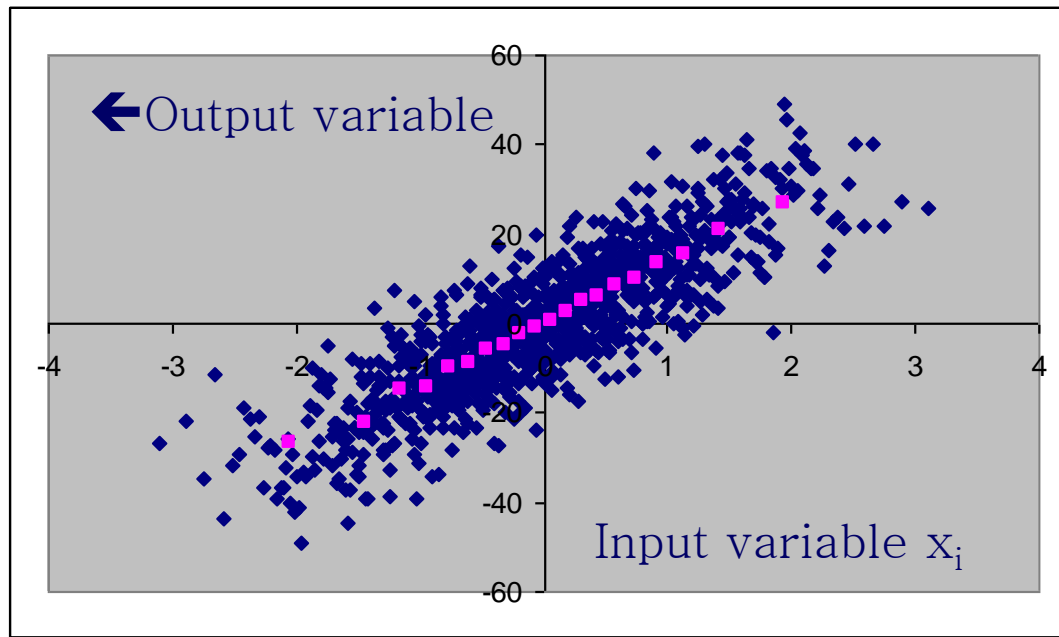
<http://www.andreasaltelli.eu>

How is it done in
practice?



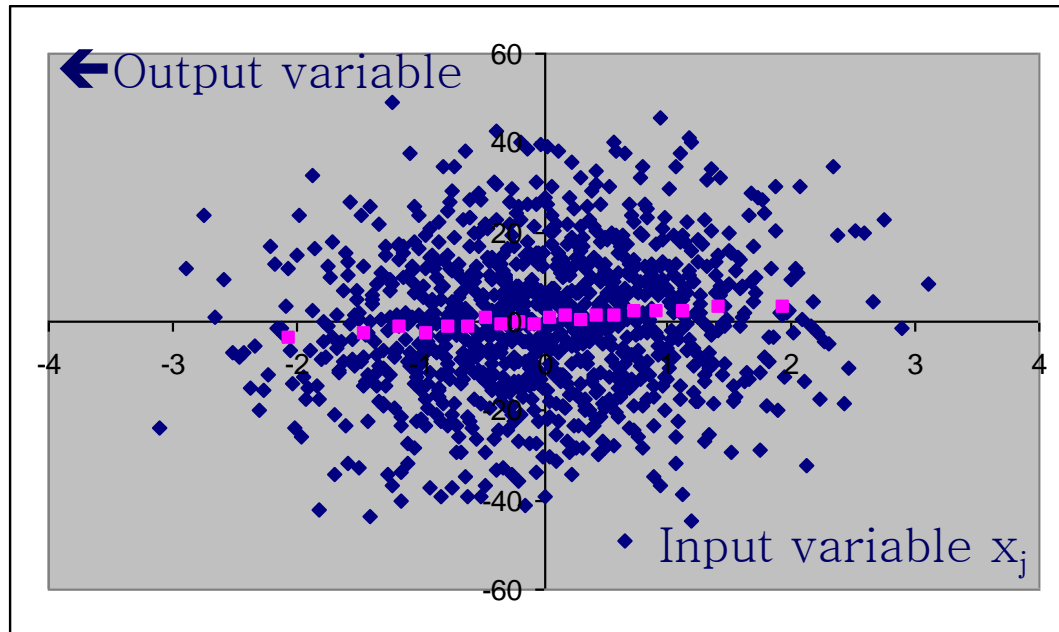
Plotting the output as a function of two different input factors

Which factor is more important?

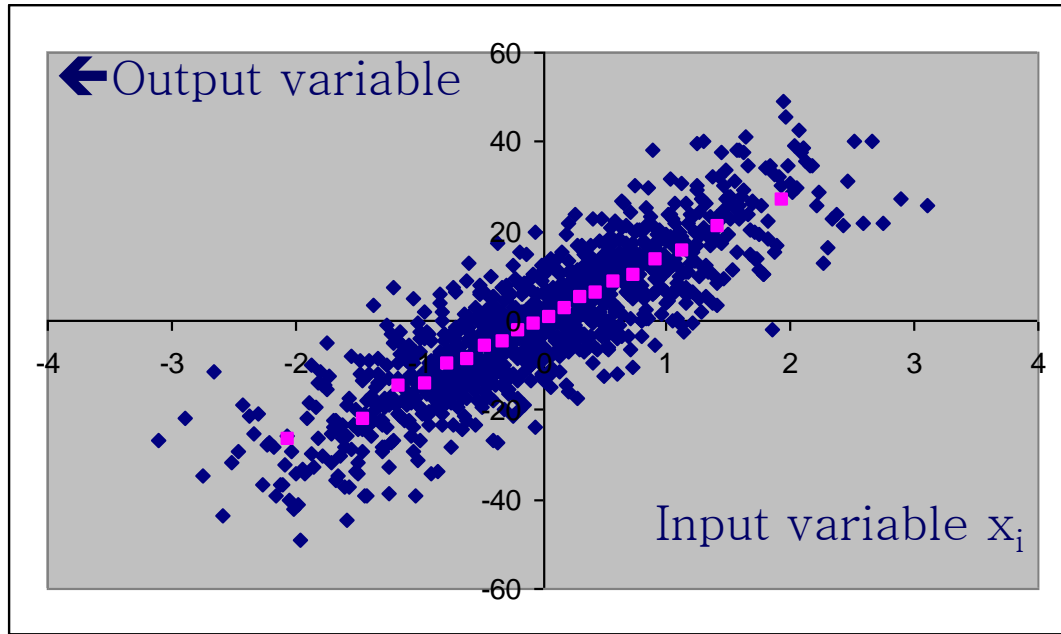


~1,000 blue points

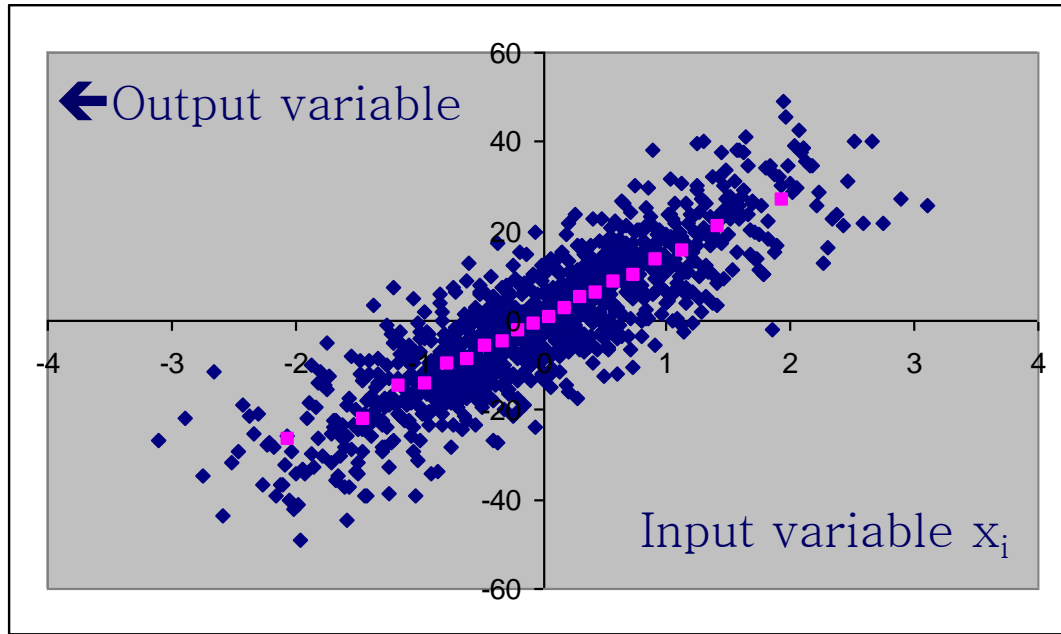
Divide them in 20 bins of ~ 50 points



Compute the bin's average (pink dots)

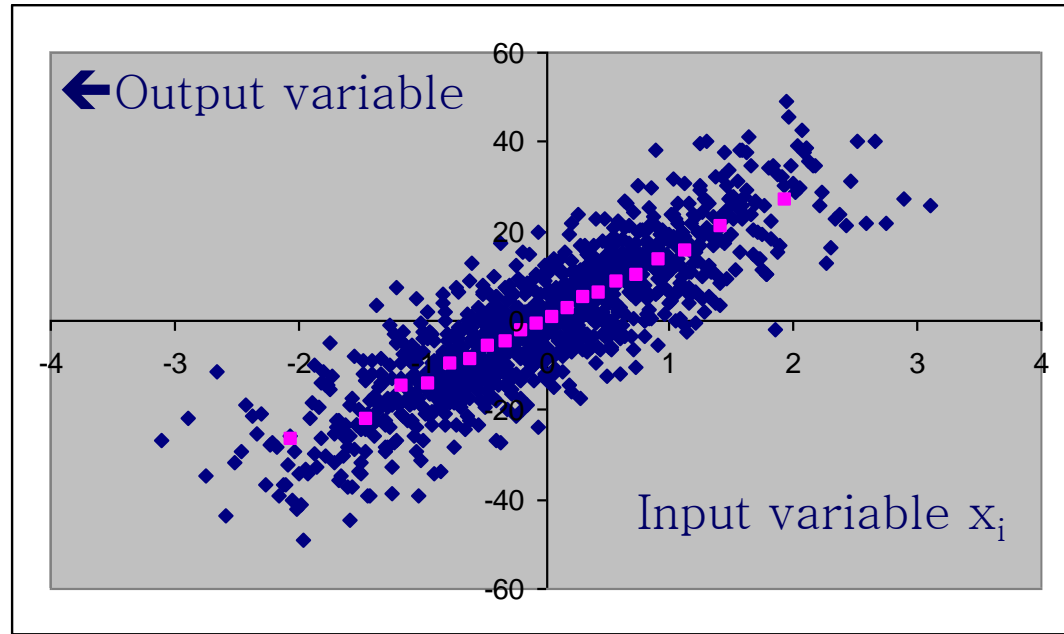


Each pink point is $\sim E_{\mathbf{x}_{\sim i}}(Y|X_i)$

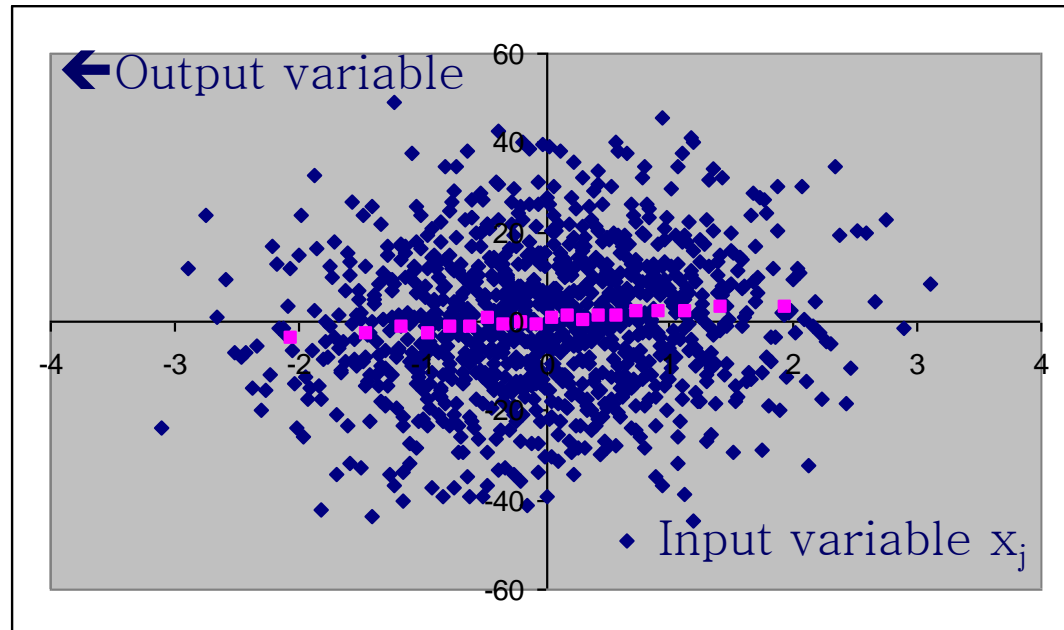


Take the variance of
the pink points one
obtains a sensitivity
measure

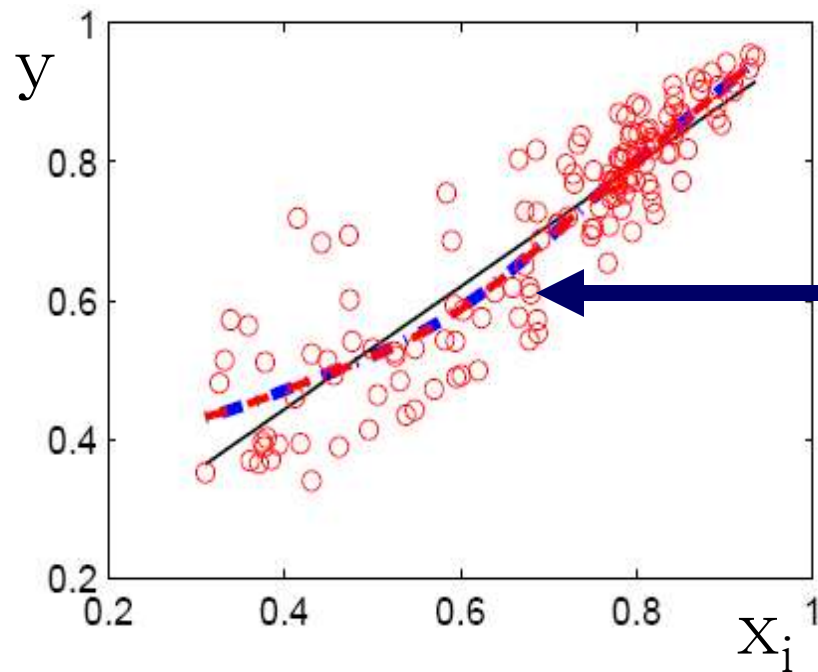
$$V_{X_i} \left(E_{\mathbf{X}_{\sim i}} (Y | X_i) \right)$$



Which factor
has the highest
 $V_{X_i} \left(E_{\mathbf{x}_{\sim i}} (Y | X_i) \right) ?$



$$S_i \equiv \frac{V\left(E\left(Y|X_i\right)\right)}{V_Y}$$



Smoothed curve:

$$\mathbf{E}_{\mathbf{x} \sim i} (y \mid x_i)$$

First order
sensitivity index:

$$\frac{V_{x_i} (\mathbf{E}_{\mathbf{x} \sim i} (y \mid x_i))}{V(y)}$$

Pearson's correlation
ratio

Smoothed curve

$$S_i \equiv \eta_i^2 := \frac{V_{x_i} (\mathbf{E}_{\mathbf{x}_{\sim i}} (y \mid x_i))}{V(y)}$$

First order sensitivity index

Unconditional
variance

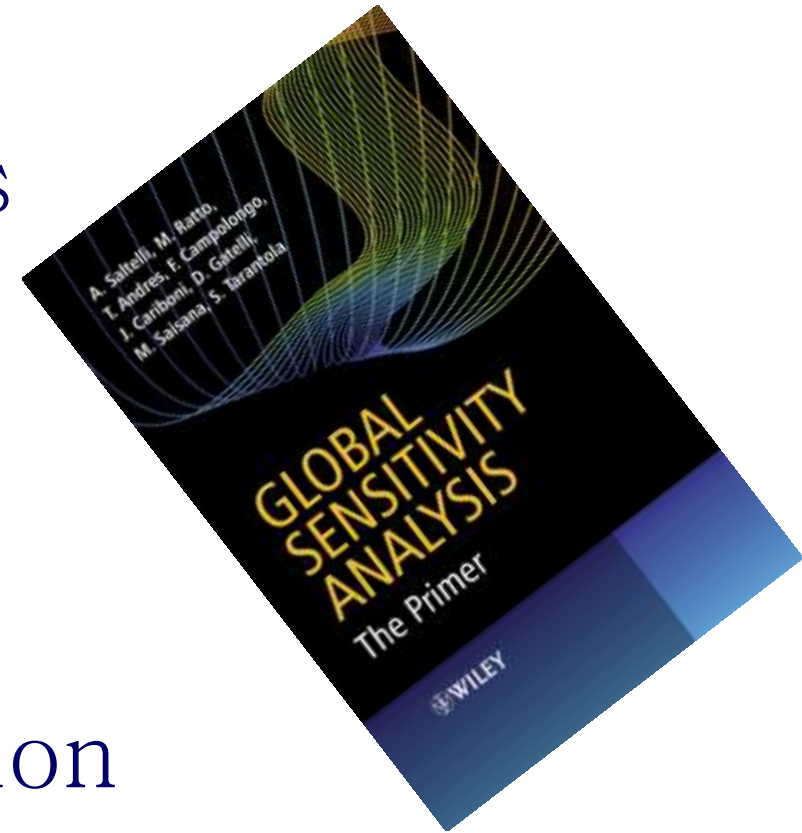
$$V_{X_i} \left(E_{\mathbf{X}_{\sim i}} (Y | X_i) \right)$$

First order effect, or top marginal variance = the expected reduction in variance that would be achieved if factor X_i could be fixed.

Why using variance-based
sensitivity analysis methods

Advantages with variance based methods:

- graphic interpretation scatterplots
- statistical interpretation
- expressed plain English
- working with sets
- relation to settings such as factor fixing and factor prioritization



Secrets of sensitivity analysis

Why should one
ever run a model
just once?

EC impact assessment guidelines: sensitivity analysis & auditing

The screenshot shows the 'Better Regulation Guidelines' page on the European Commission's website. The header features the European Commission logo and the title 'Better Regulation'. A navigation menu on the left lists various sections, with 'Guidelines' highlighted. The main content area provides an overview of the guidelines, explaining their purpose and structure. A sidebar on the right includes a search bar, social media links, and a list of latest documents. The footer contains a last update date and additional navigation links.

Press | Archives | Sitemap | About this site | Legal notice | Contact | Search | English (en)

Better Regulation

European Commission > Better Regulation > Guidelines

Better Regulation Guidelines

These guidelines explain what Better Regulation is and how it should be applied in the day to day practices when preparing new initiatives and proposals or managing existing policies and legislation.

They cover the whole policy cycle, from policy preparation and adoption to implementation and application, to evaluation and revision of EU law. For each of these phases there are a number of Better Regulation principles, objectives, tools and procedures to make sure that the EU has the best regulation possible. These relate to planning, impact assessment, stakeholder consultation, implementation and evaluation.

The **Better Regulation Guidelines** are structured into chapters which cover each of the instruments of the law-making process. The corresponding [toolbox](#) gives more detailed and technical information.

Better Regulation Guidelines are based on the outcomes of public consultation exercises carried out in 2013 and 2014.

- [Public consultation on the revision of the Commission's Impact Assessment Guidelines](#)
- [Stakeholder Consultation Guidelines](#)
- [Consultation on the draft Commission Evaluation Policy Guidelines](#)

Share

Search

Stay connected

Facebook Twitter YouTube EU Tube

Blog

Latest documents

- [19/05/2015 - Better Regulation Package](#)

Help us improve

Find what you wanted?

Yes ☐ No ☐

What were you looking for?

Any suggestions?

Send

Last update: 11/08/2015 | [Legal notice](#) | [Cookies](#) | [Contact](#) | [Search](#) | [Top](#)

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]



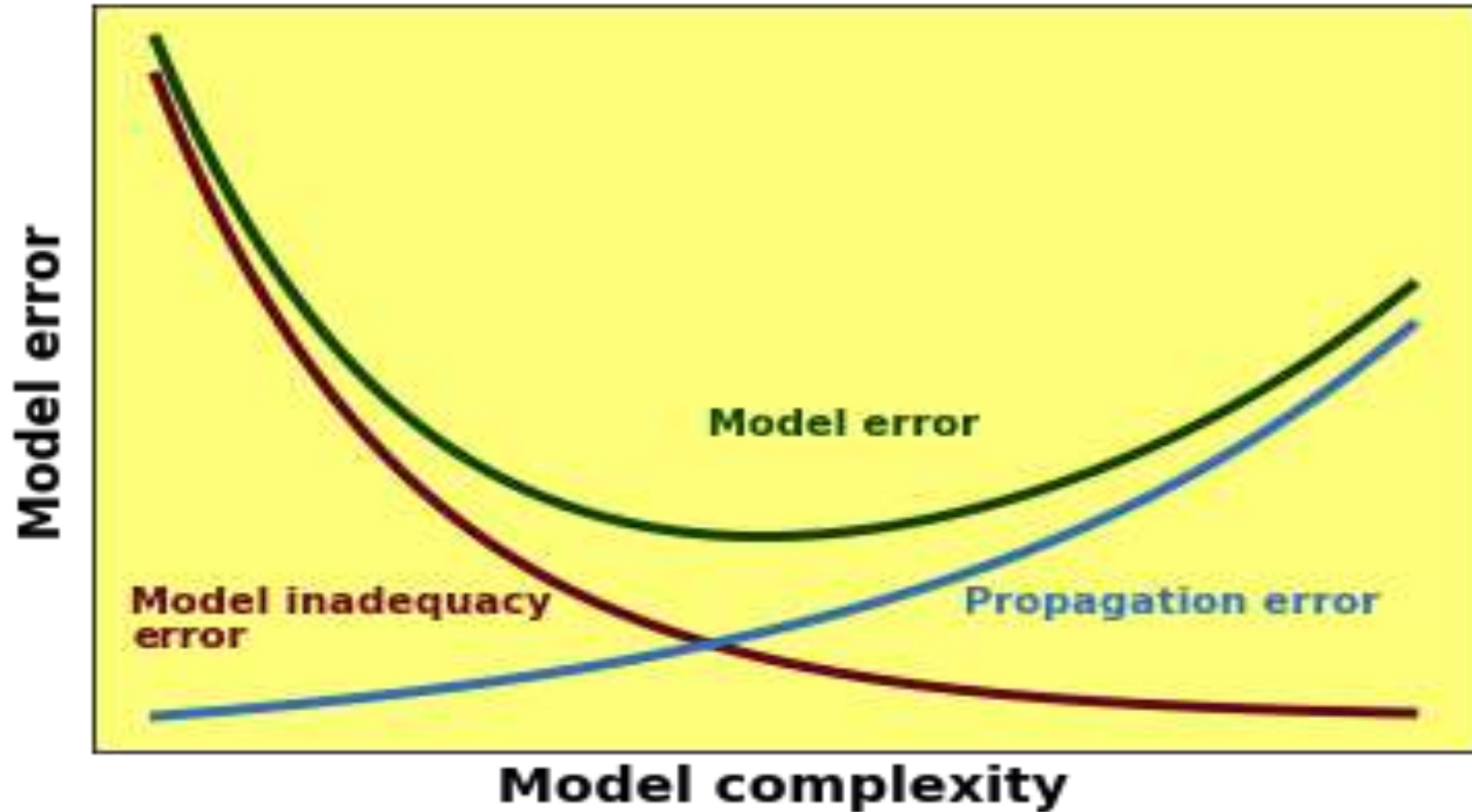
Third secret: If sensitivity analysis shows that a question cannot be answered by the model one should find another question or model

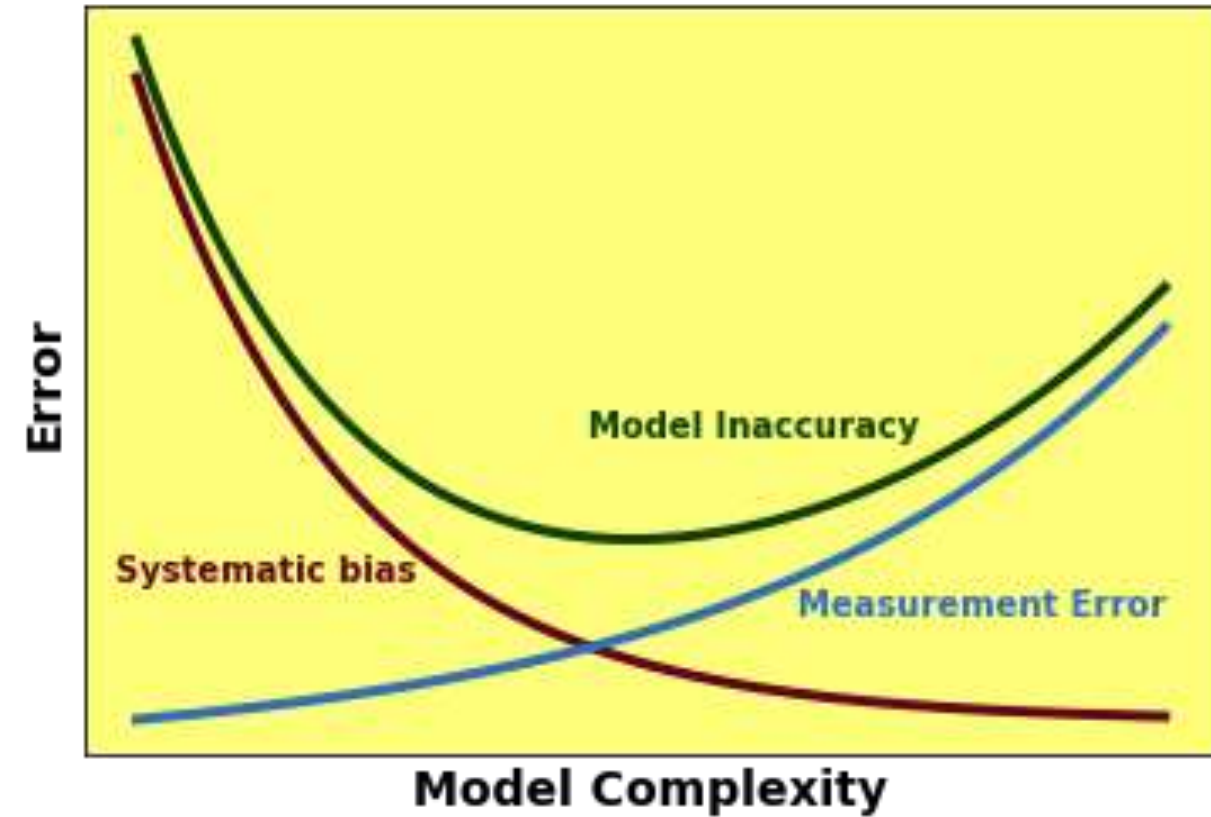
[Often the love for one's own model prevails]

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’

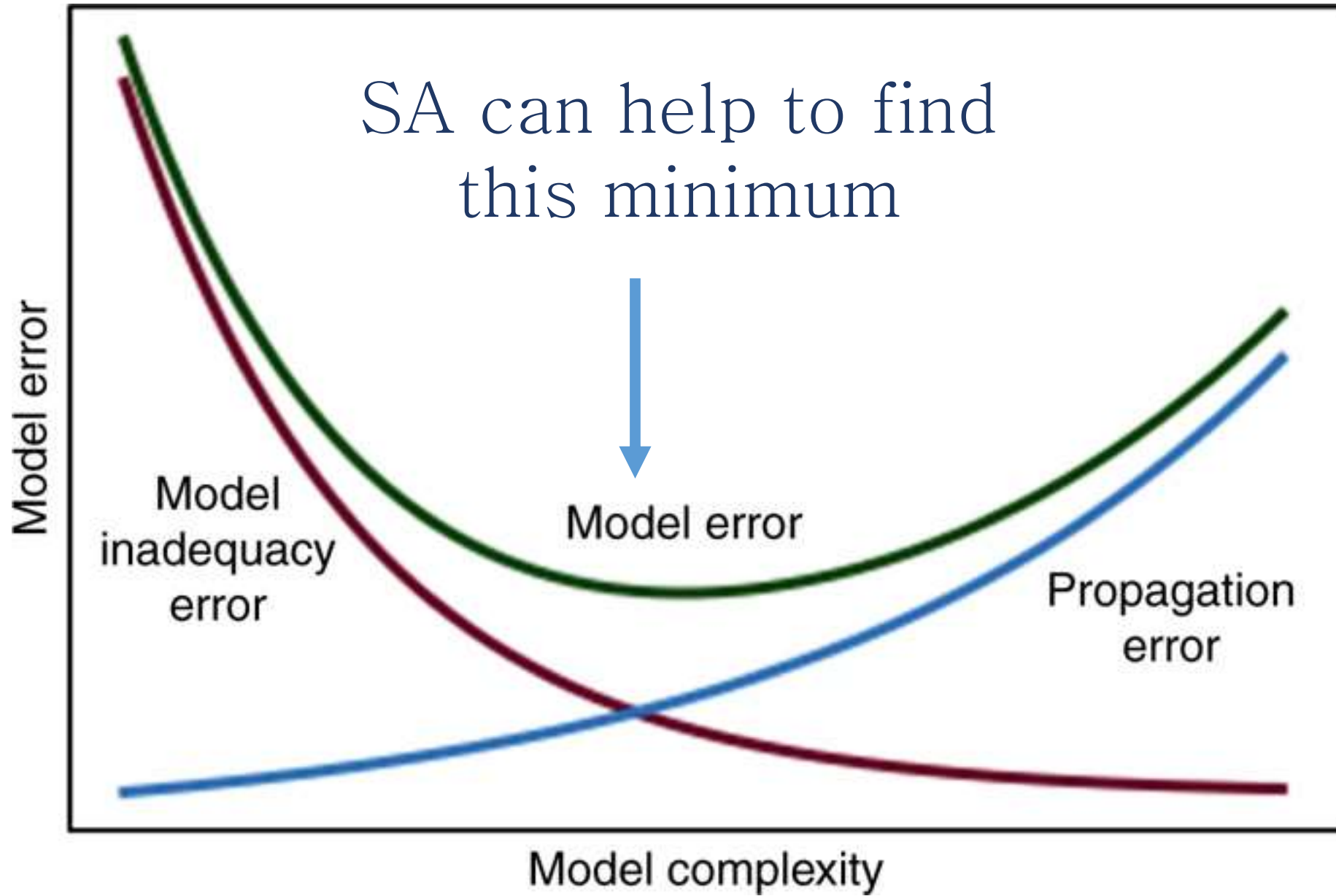
In M. G. Turner and R. H. Gardner,
“Introduction to Models” in Landscape
Ecology in Theory and Practice, New
York, NY: Springer New York, 2015, pp.
63–95.



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.



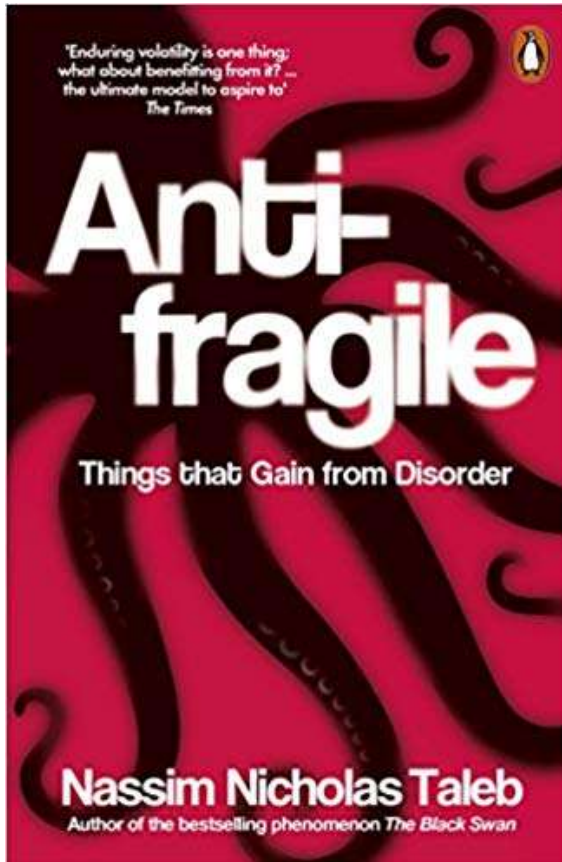
Sixth secret:

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?





Environmental Modelling & Software

Volume 114, April 2019, Pages 29-39



Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices

Andrea Saltelli ^{a, b}  , Ksenia Aleksankina ^c, William Becker ^d, Pamela Fennell ^e, Federico Ferretti ^d, Niels Holst ^f, Sushan Li ^g, Qiongli Wu ^h

Limit of SA: Often no SA (SA
conflated with UA e.g. in economics) or
one-factor-at-a-time SA

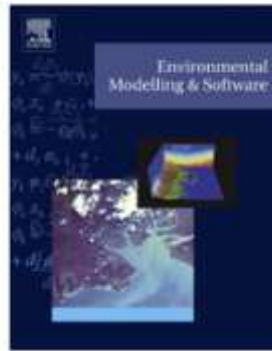
Why is OAT (one-factor-at-a-time) SA so bad?



Contents lists available at ScienceDirect

Environmental Modelling & Software

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

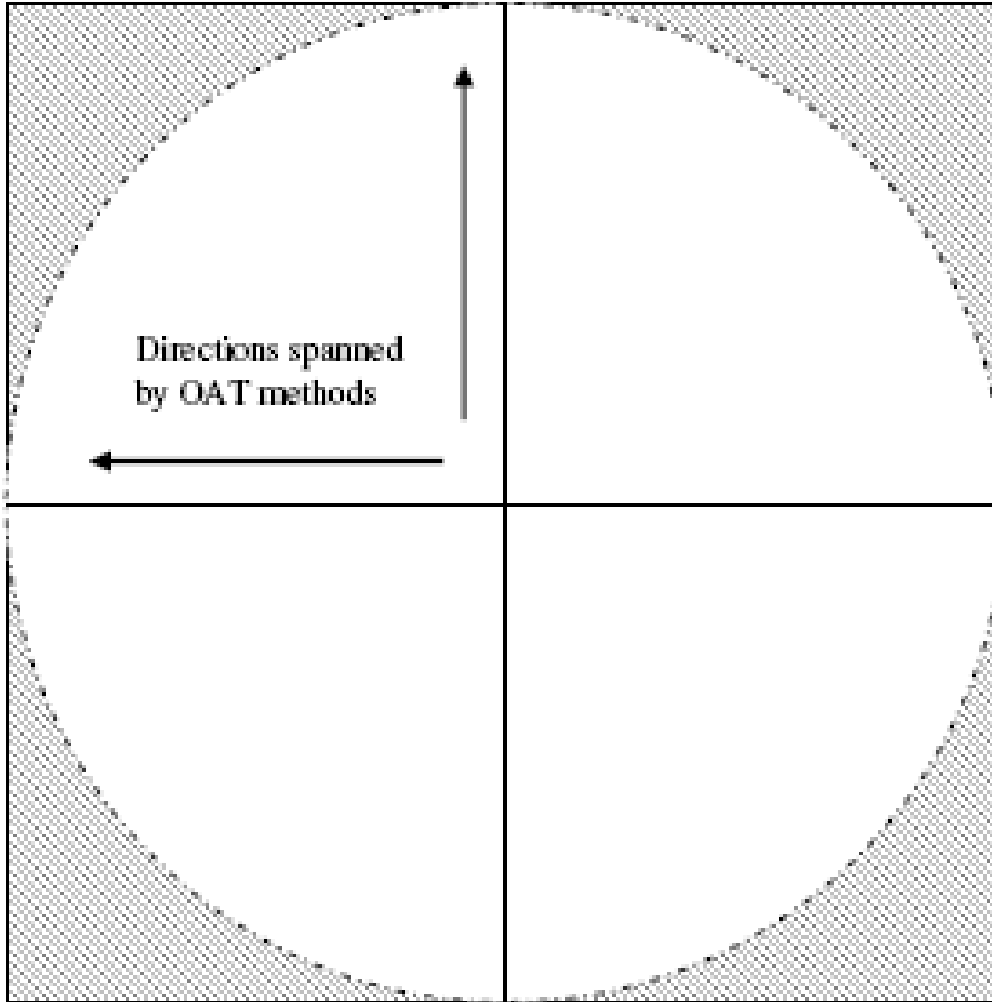


How to avoid a perfunctory sensitivity analysis

Andrea Saltelli*, Paola Annoni

Joint Research Center, Institute for the Protection and Security of the Citizen, via E.Fermi, 2749, Ispra VA 21027, Italy

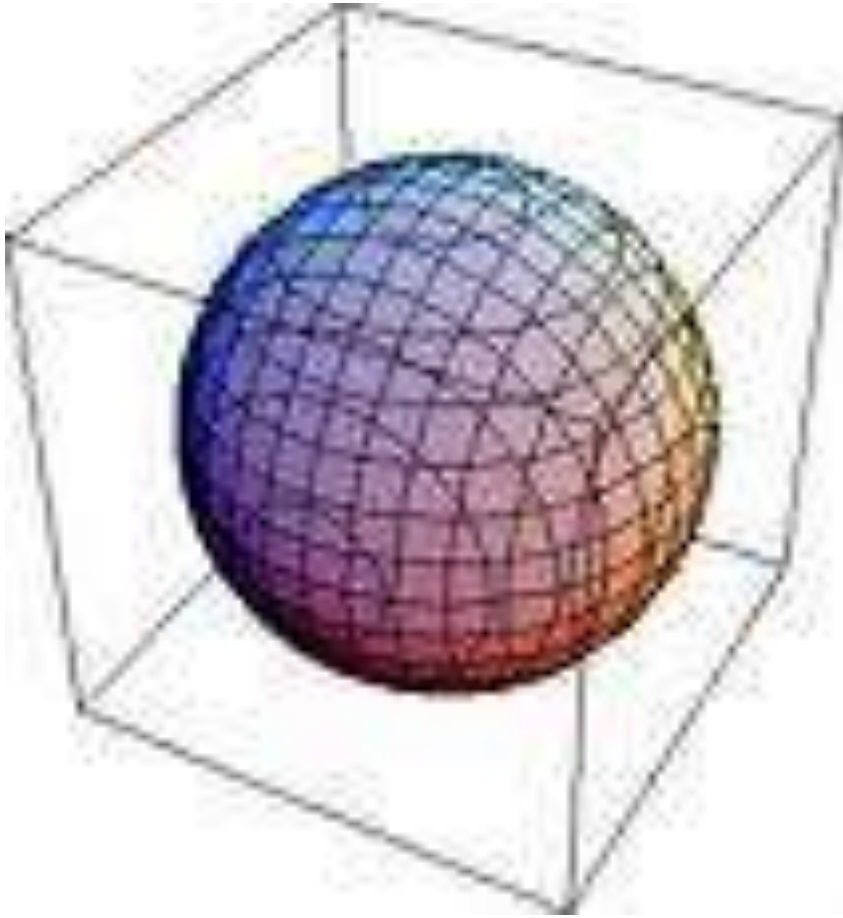
OAT in 2 dimensions



Area circle
/ area
square =?

$\sim 3/4$

OAT in 3 dimensions



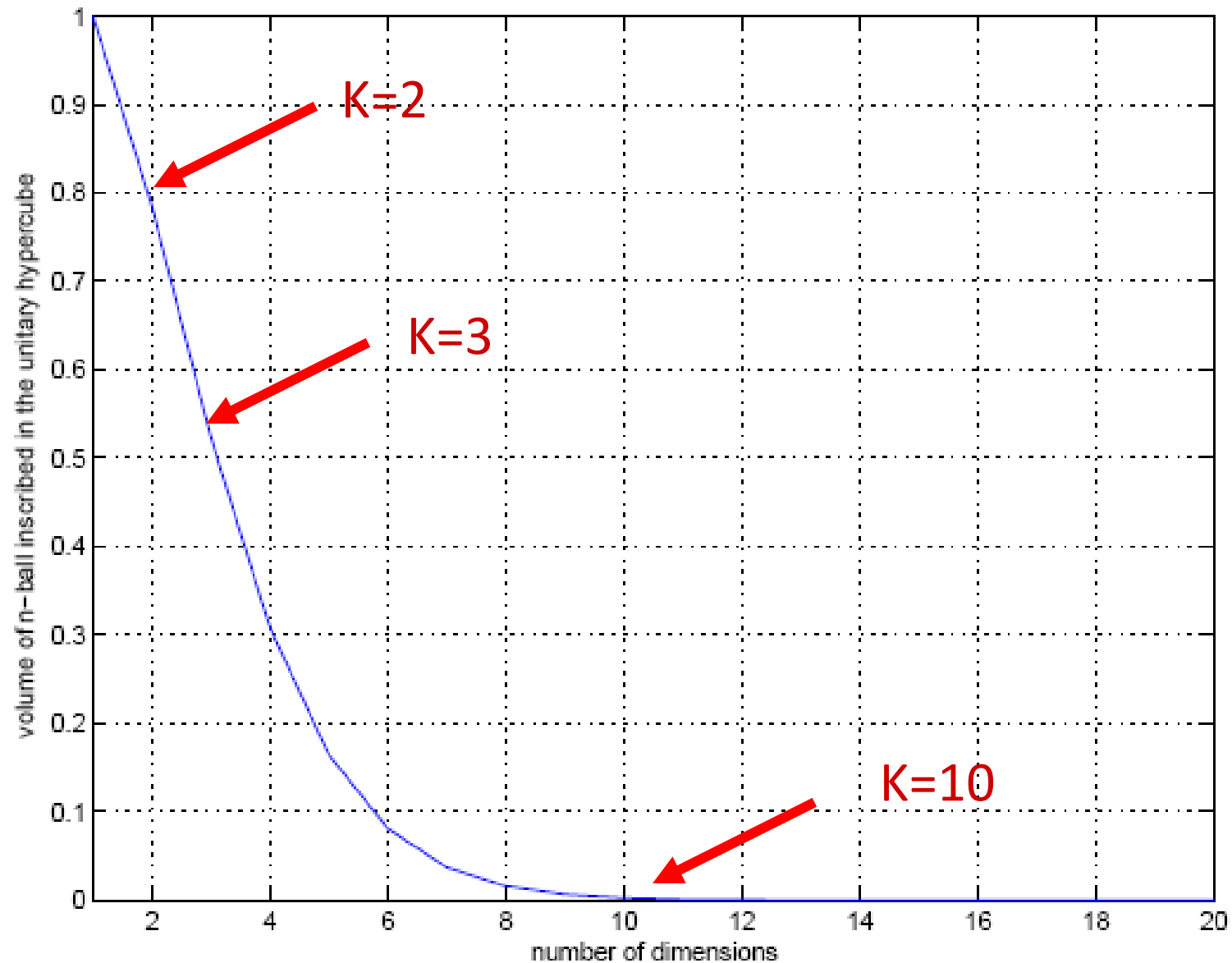
Volume sphere /
volume cube =?

$\sim 1/2$

OAT in 10 dimensions; Volume
hypersphere / volume ten dimensional
hypercube =? ~ 0.0025



OAT in k dimensions



Literature search in Scopus

Query: “sensitivity analysis” & “model/modelling”
& “uncertainty”; years 2012–2017; journal
articles; in English

➔ 6000 articles

- AgrBioSci (Agricultural and Biological Sciences)
- BiochemGenMBio (Biochemistry, Genetics and Molecular Biology)
- BusManAcc (Business, Management and Accounting)
- Chemi (Chemistry)
- ChemEng (Chemical Engineering)
- CompSci (Computer Science)
- DecSci (Decisional Science)
- EarthSci (Earth and Planetary Sciences)
- EconFin (Economy and Finance)
- Energy (Energy)
- Engineering (Engineering)
- EnvSci (Environmental Science)
- ImmunMicrobio (Immunology and Microbiology)
- MatSci (Material Science)
- Math (Math)
- Medicine (Medicine)
- PharTox (Pharmacology and Toxicology)
- PhysAstro (Physics and Astronomy)
- SocSci (Social Science)

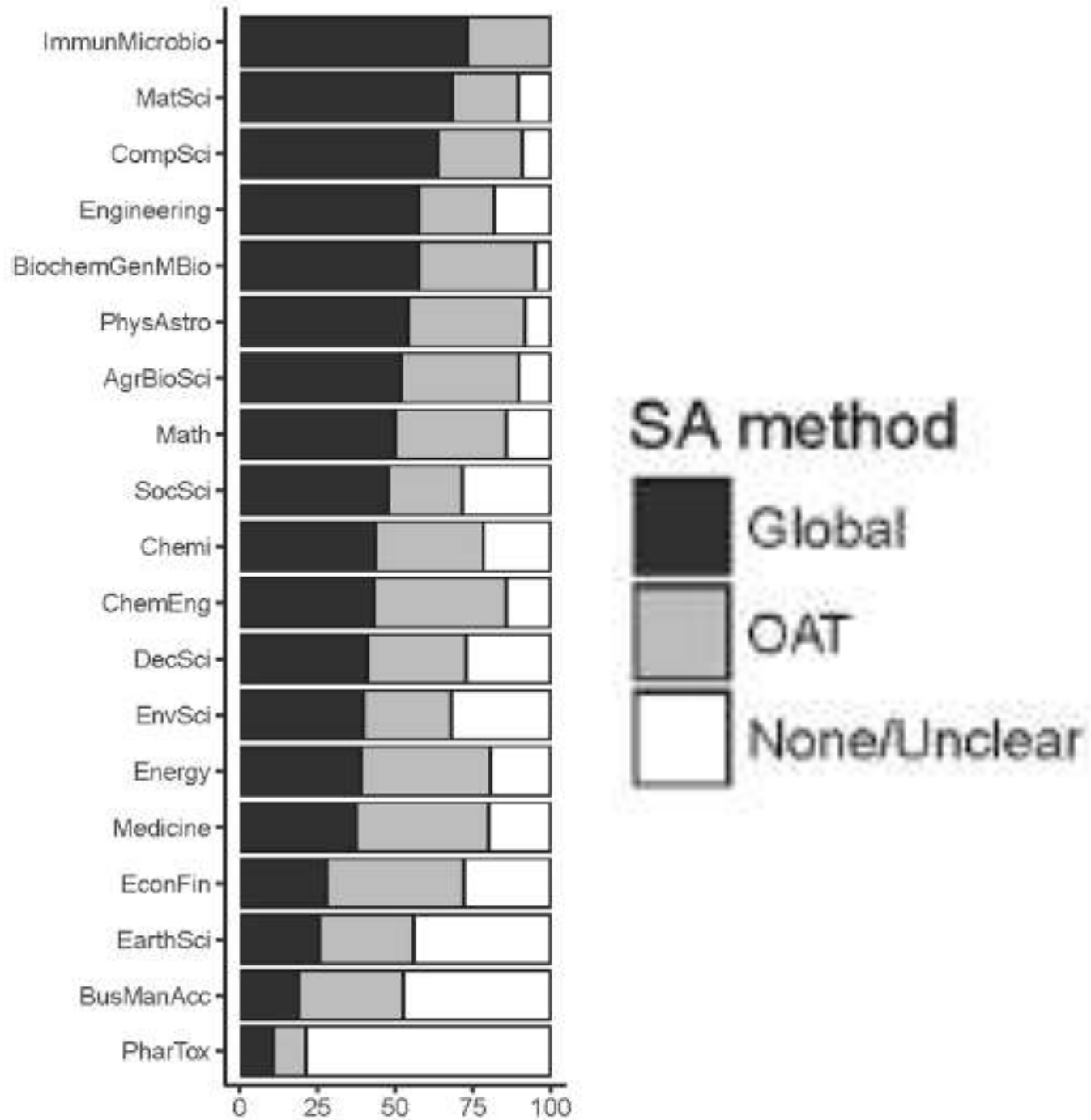
subject areas >100 articles

Taking the top twenty most-cited papers in each subject area:

→ 324 articles, divided among authors

Cleansing manually irrelevant articles:

→ 280 articles



Still many papers
apply an OAT SA:
65%

What if the model is truly linear?

Linear	7%
Nonlinear	61%
Unclear	32%

Linear	7%
Nonlinear	61%
Unclear	32%

65% highly cited articles are OAT

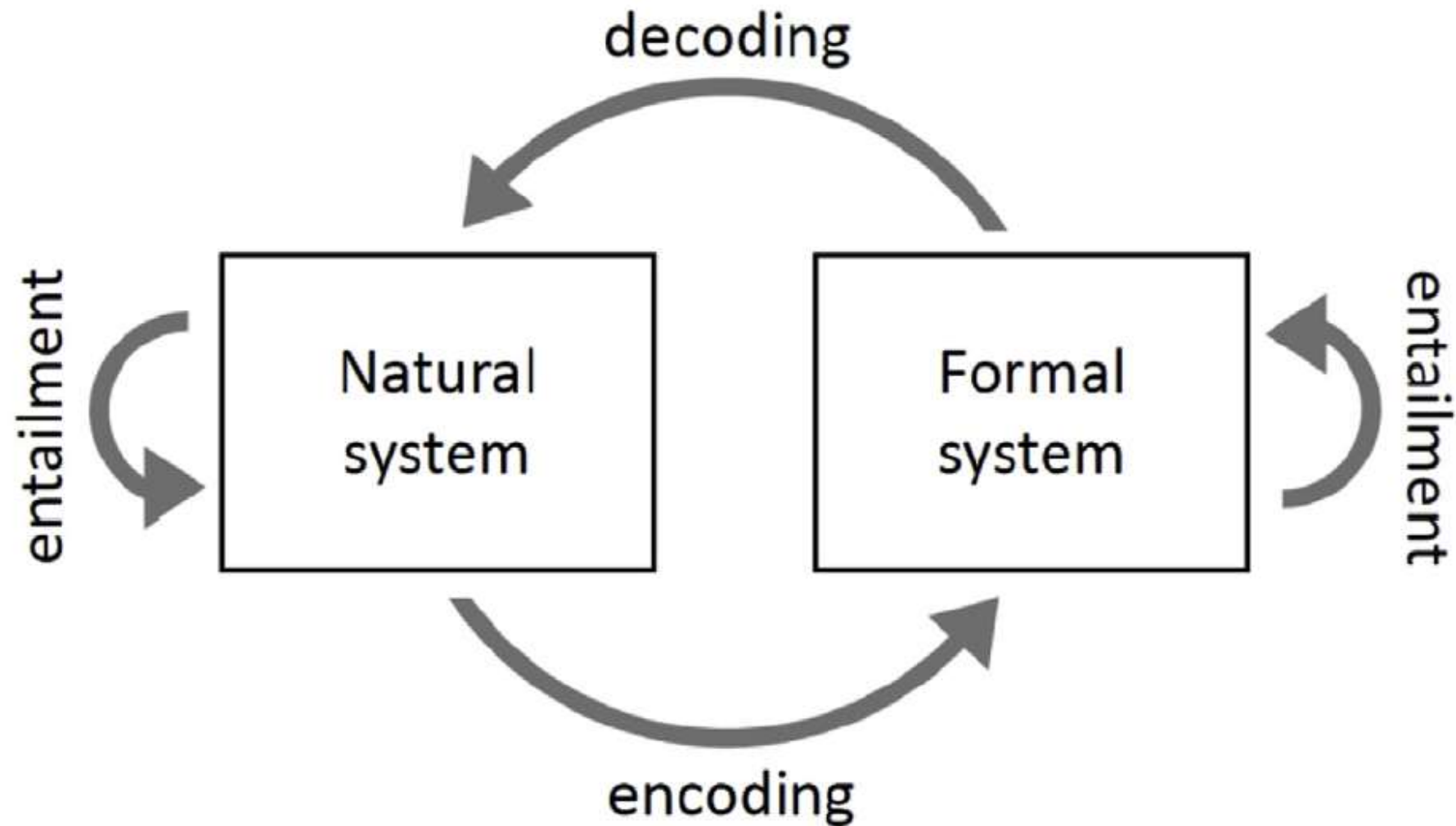
Taking all unclear = linear → still
over 20% of papers wrong
(OAT & non-linear model)

Why?

5. Discussion

5.1. Reasons for bad practice

Why? ➔ 1. Modelling as a craft



Why? ➔ 2. Each discipline going about modelling on its own separate way; pockets of SA practitioners (out of our 280 papers, 35 were methodological, of which 24 suggest global SA)

Why? ➔ 3. Mathematical modelling is not
a discipline

... mathematical modelling cannot do this:



**AMERICAN STATISTICAL ASSOCIATION RELEASES STATEMENT ON
STATISTICAL SIGNIFICANCE AND *P*-VALUES**

*Provides Principles to Improve the Conduct and Interpretation of Quantitative
Science*

March 7, 2016

Wasserstein, R.L. and Lazar, N.A., 2016. 'The ASA's statement on p-values: context, process, and purpose', *The American Statistician*, Volume 70, 2016 – Issue 2, Pages 129–133.

Need for a more structured, generalized and standardized approach to verification

Padilla, J. J., Diallo, S. Y., Lynch, C. J., & Gore, R. (2018). Observations on the practice and profession of modeling and simulation: A survey approach. *SIMULATION*, 94(6), 493–506.

Why? ➔ 4. Good practices require
training in statistics

Why? ➔ 5. More time is needed; though mature global sensitivity analysis methods around for more than 25 years researchers tend to emulate methods found in highly cited papers assuming that they are best practice

Why? ➔ 6. Strategic reasons: global SA is bad if one wants to play the uncertainty game, inflating or deflating uncertainties instrumentally

Solutions? 1. Statistics as a discipline
takes responsibility for statistical
methods for
model validation and verification

Example: who can authoritatively suggest
to modellers not to overinterpret results
from multi-model ensembles?

Climate Models as Economic Guides: Scientific Challenge or Quixotic Quest?

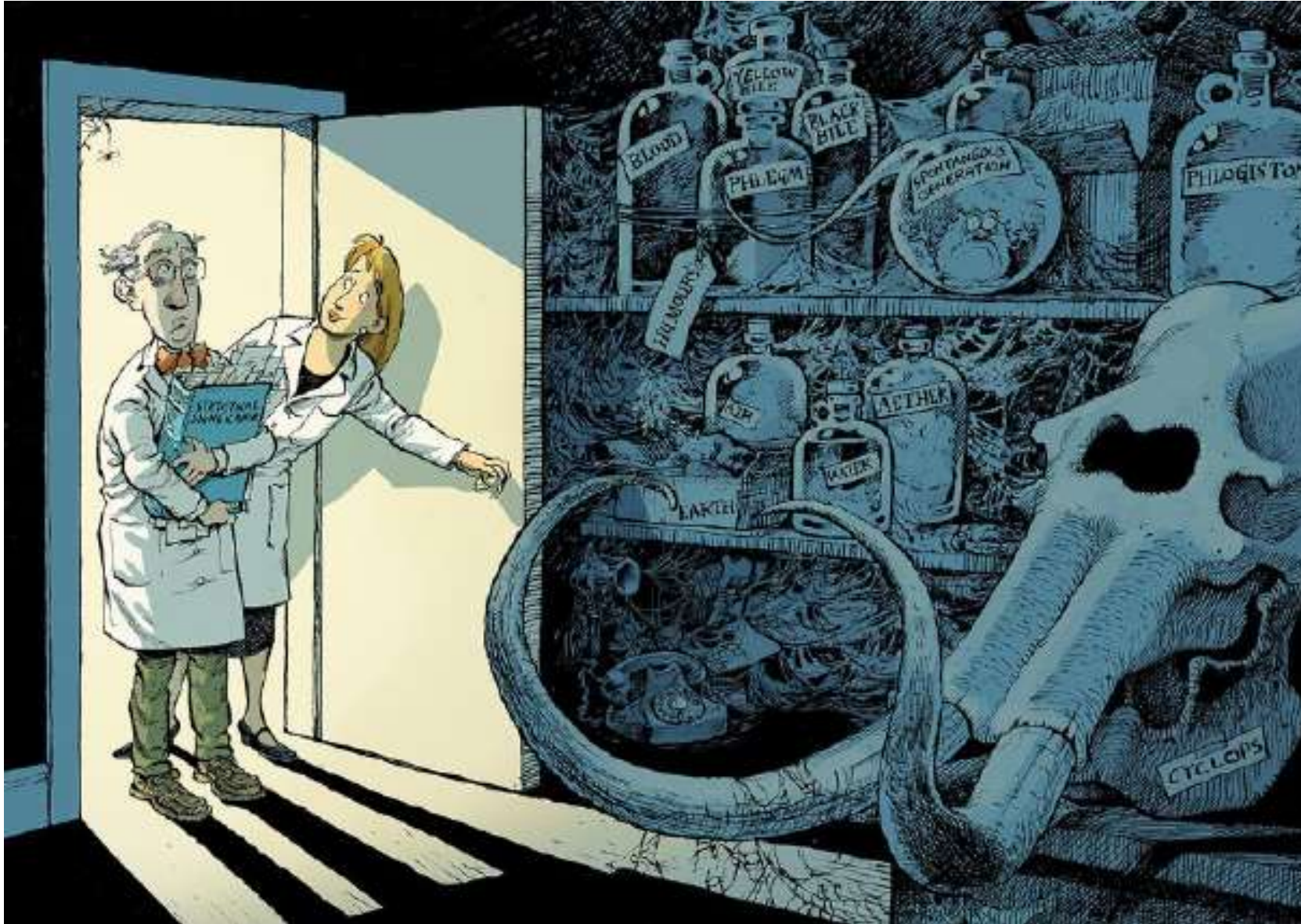
BY ANDREA SALTELLI, PHILIP B. STARK, WILLIAM BECKER, PAWEL STANO

Climate Models as Economic Guides: Scientific Challenge or Quixotic Quest?

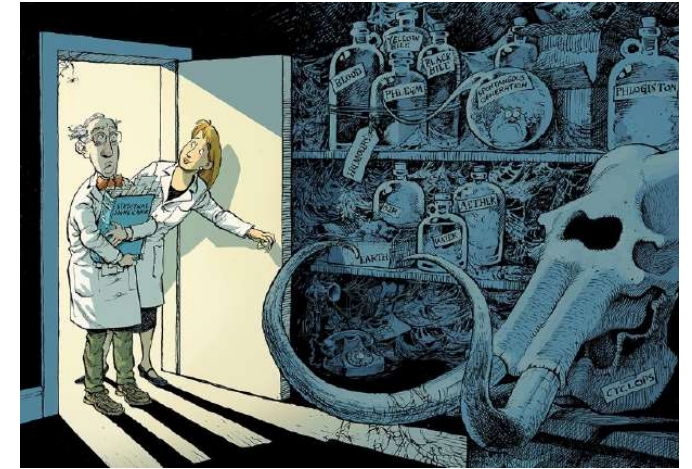
BY ANDREA SALTELLI, PHILIP B. STARK, WILLIAM BECKER, PAWEL STANO

A plea against audacious risk or cost-benefit analysis running over centennial time scales; example: crime rate as modified by climate change at US county level in 2100

Solutions? 2. Learn from what happens in statistics where the p-test crisis is being tackled head on



Throw away
the concept of
statistical
significance?



COMMENT • 20 MARCH 2019

Scientists rise up against statistical significance

Valentin Amrhein, Sander Greenland, Blake McShane and more than 800 signatories call for an end to hyped claims and the dismissal of possibly crucial effects.

Valentin Amrhein , Sander Greenland & Blake McShane

See the discussion on the blog of Andrew Gelman <https://statmodeling.stat.columbia.edu/>

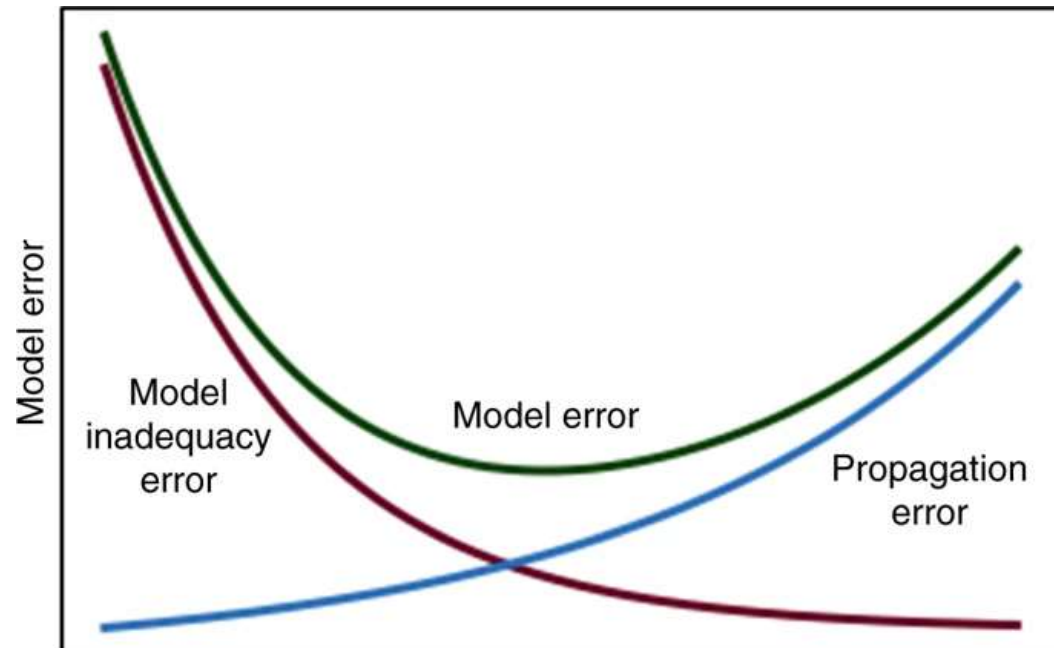


Cargo-cult statistics and scientific crisis

The mechanical, ritualistic application of statistics is contributing to a crisis in science. Education, software and peer review have encouraged poor practice – and it is time for statisticians to fight back. By **Philip B. Stark** and **Andrea Saltelli**

Lessons for sensitivity analysis

- Global SA
- UA and SA coupled
- Purpose- & context-specific
- The map is not the territory



- Memento

Beyond sensitivity
analysis: sensitivity
auditing

Sensitivity auditing

EC guidelines on impact assessment 2015, and
SAPEA report 2019





ISSUES

IN SCIENCE AND TECHNOLOGY

4 VOL. XXX, NO. 2, WINTER 2014

When All Models Are Wrong

BY ANDREA SALTELLI, SILVIO FUNTOWICZ

The rules of sensitivity auditing

1. Check against rhetorical use of mathematical modelling;
2. Adopt an “assumption hunting” attitude; focus on unearthing possibly implicit assumptions;
3. Check if uncertainty been instrumentally inflated or deflated.

4. Find sensitive assumptions before these find you; do your SA before publishing;
5. Aim for transparency; Show all the data;
6. Do the right sums, not just the sums right;
7. Perform a proper global sensitivity analysis.



Comment

Open Access

Published: 27 August 2019

A short comment on statistical versus mathematical modelling

Andrea Saltelli 

Is there a broader problem
affecting different instances of
quantification?

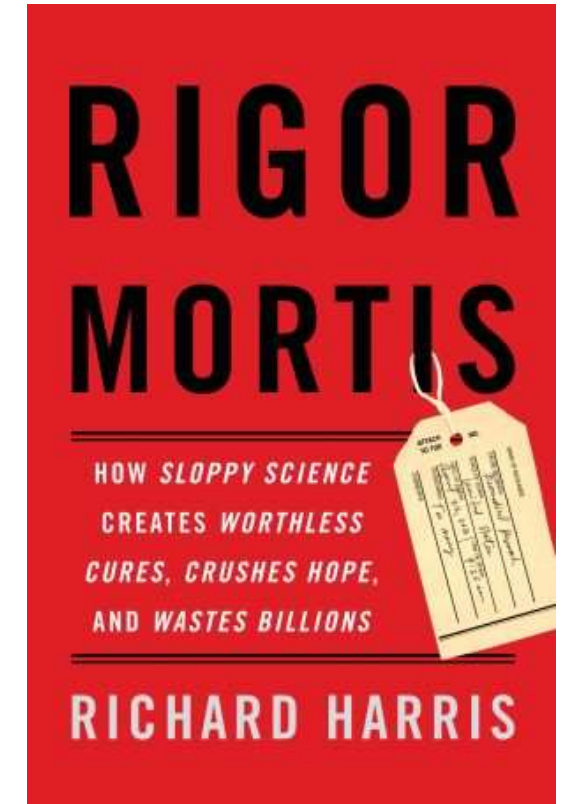
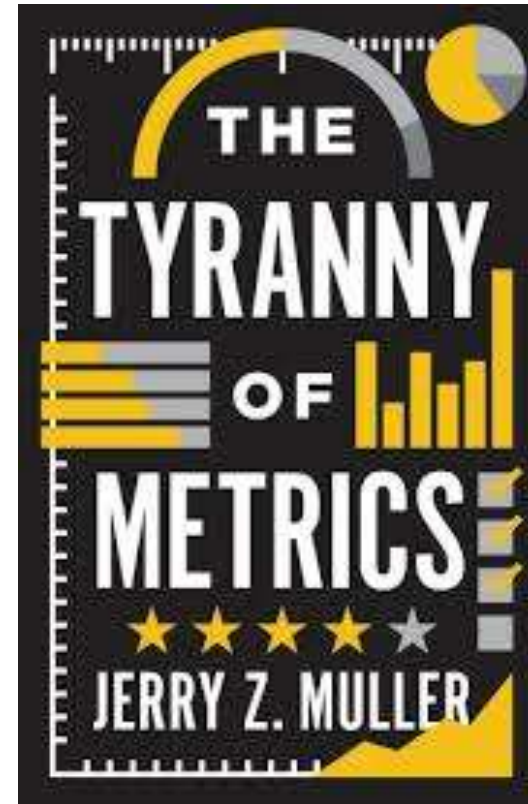
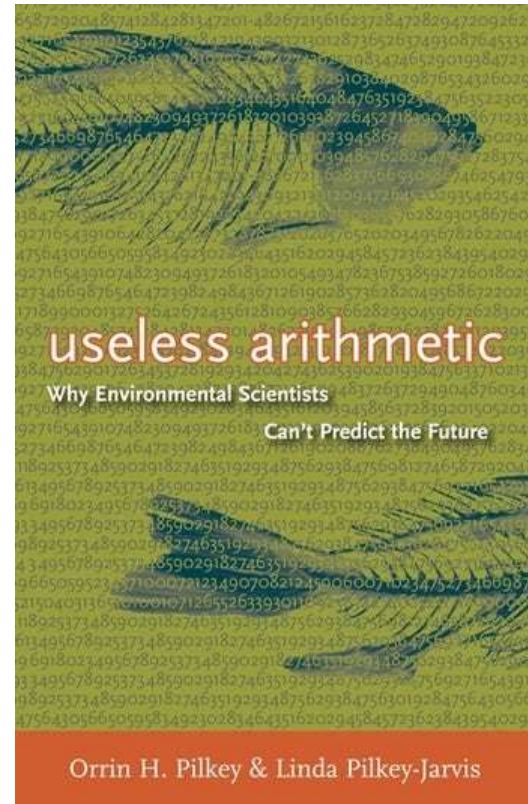
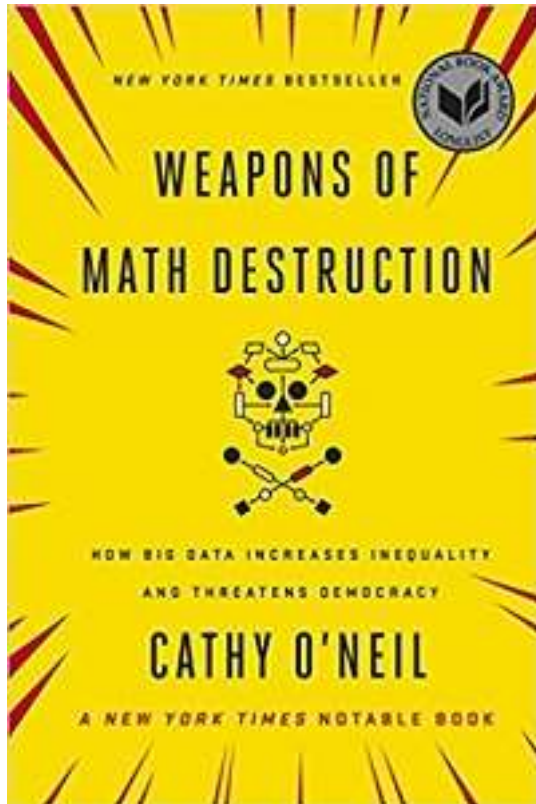
“what qualities are specific to rankings, or indicators, or models, or algorithms?”



E. Popp Berman

Popp Berman, E. & Hirschman, D. The Sociology of Quantification: Where Are We Now? *Contemp. Sociol.* 47, 257–266 (2018).

Algorithms, models, metrics, statistics



Common root causes?

Theodore M. Porter

TRUST IN NUMBERS

The Pursuit
of Objectivity
in Science and
Public Life

Alain Supiot

La Gouvernance
par les nombres

*Cours au Collège de France
(2012-2014)*



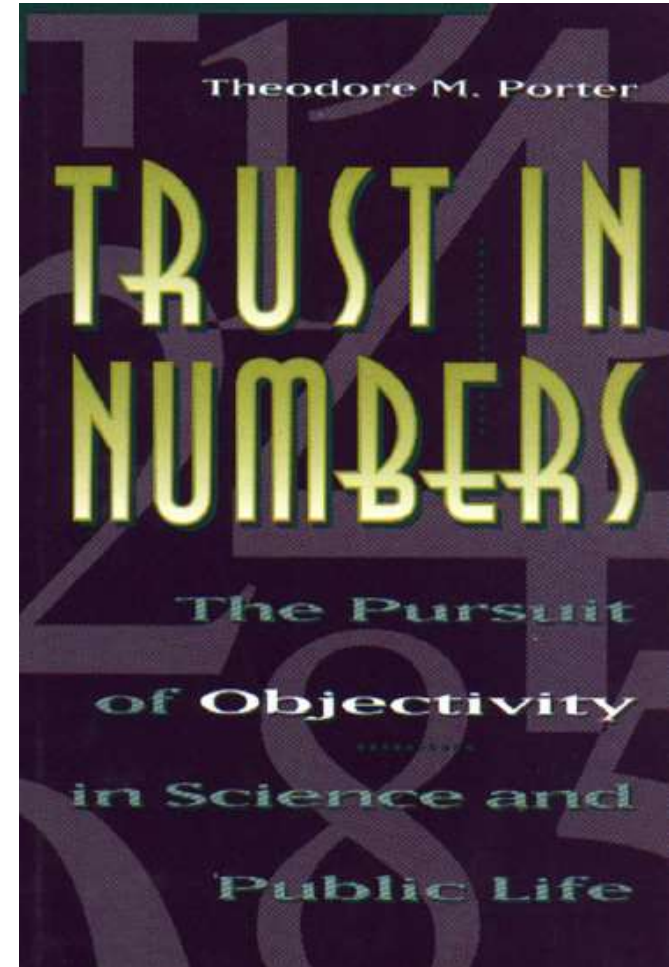
FAYARD
POIDS ET MESURES
DU MONDE

Can we learn
something from
sociology of
numbers?

From law?

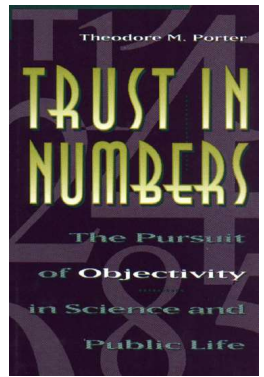


Theodor
M. Porter



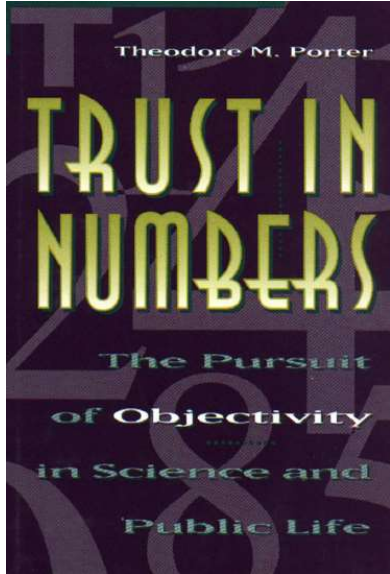
Theodore M. Porter, Trust in Numbers,
The Pursuit of Objectivity in Science and Public Life, Princeton 1995

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.”

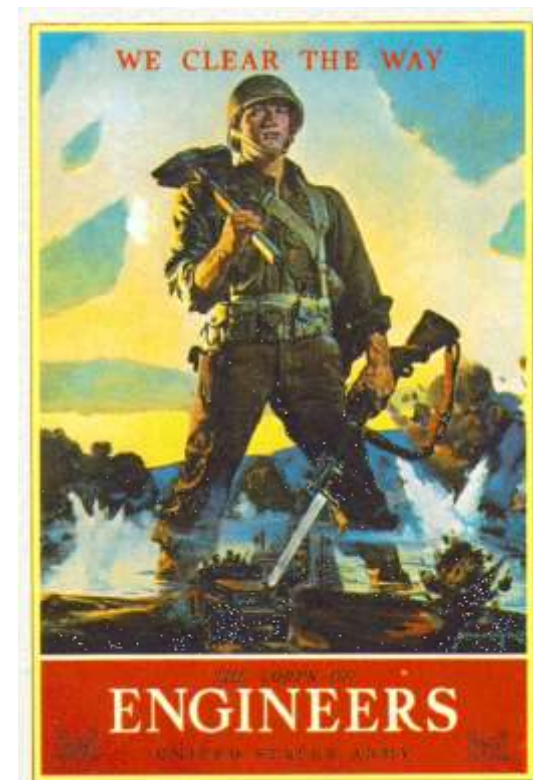
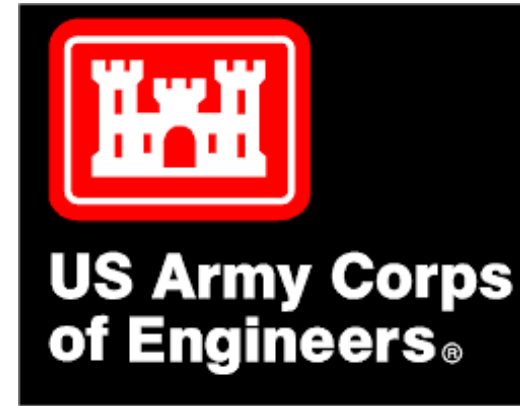


p. 8: “Scientific objectivity thus provides an answer to a moral demand for impartiality and fairness.

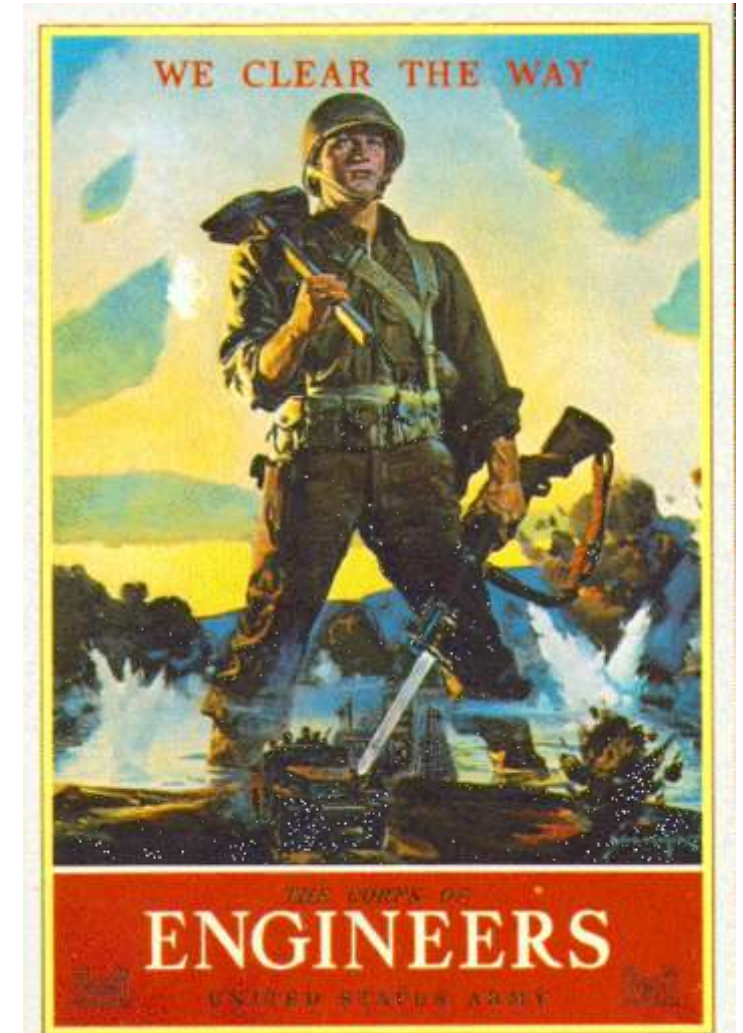
Quantification is a way of making decisions without seeming to decide.

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

Trust, authority and styles of quantification: two different stories



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





Charles Goodhart

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

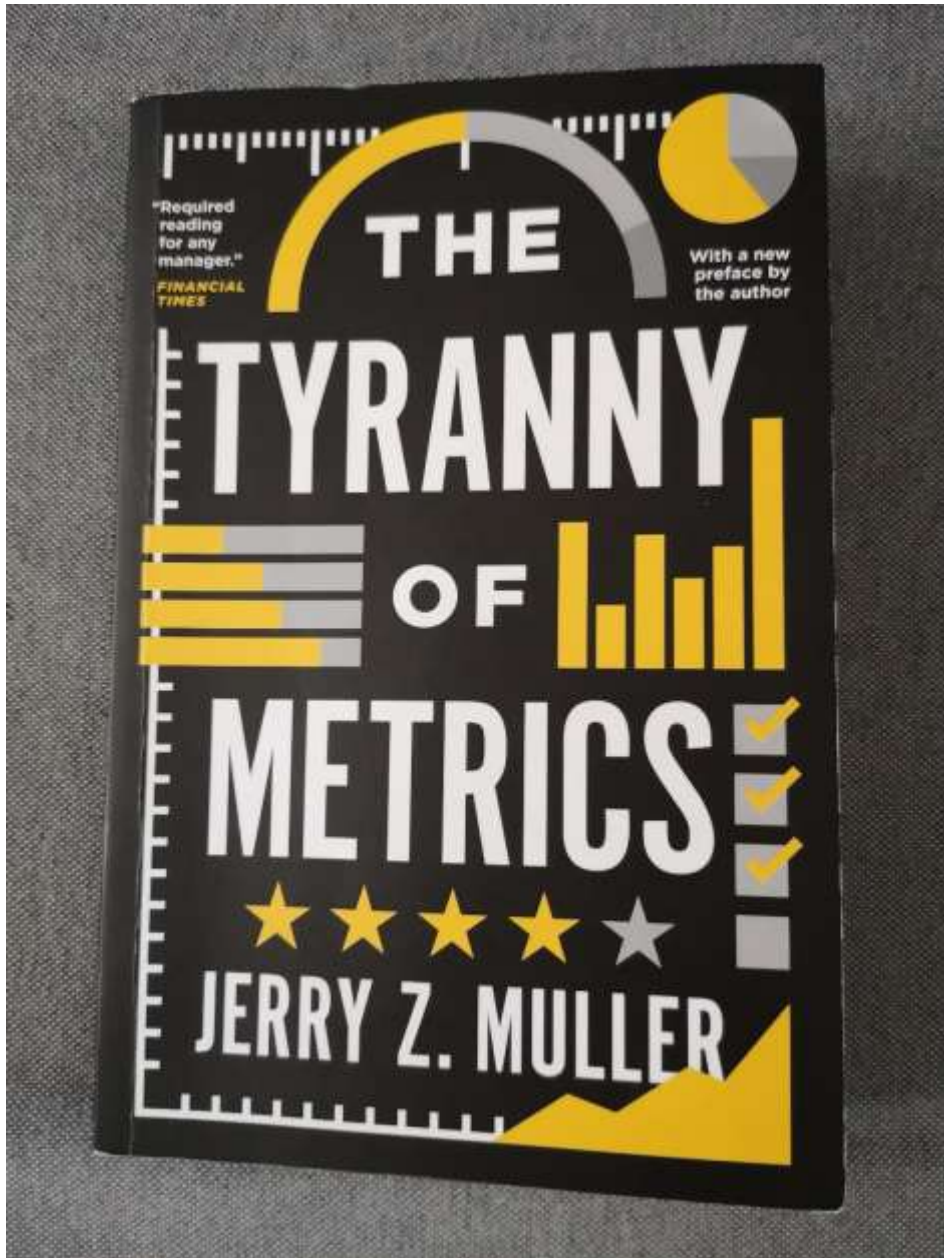
This is what we normally call Goodhart's law, from Charles Goodhart. "When a measure becomes a target, it ceases to be a good measure."

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

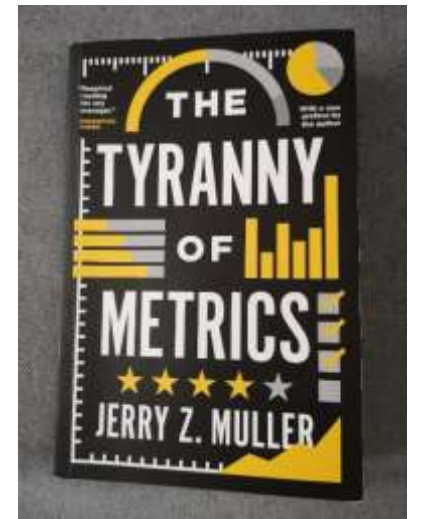
J. Z. Muller, *The tyranny of metrics*.
Princeton University Press , 2018.

Metric fixation, or the irresistible pressure to measure performance

Gaming of metrics (recall Goodhart law)

“The calculative is the enemy of the imaginative”

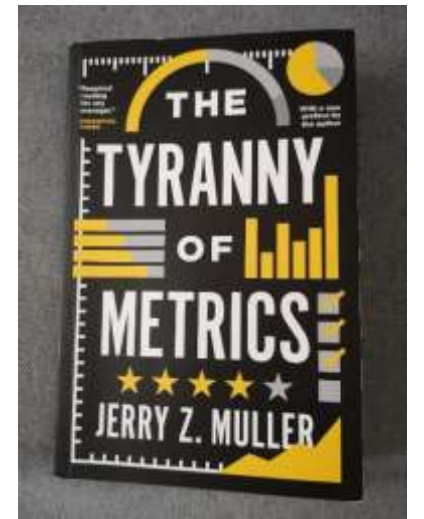
A wealth of case studies from education to war to medicine to foreign aid..



Critiques of metrics

From the left: metric fixation promotes deskilling

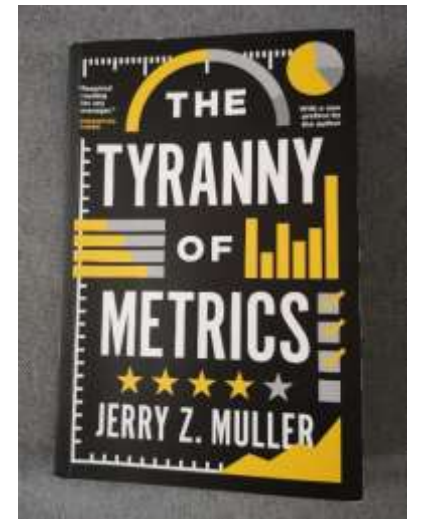
From the right (Friedrich Hayek):
metric fixation reproduces features of
the soviet system



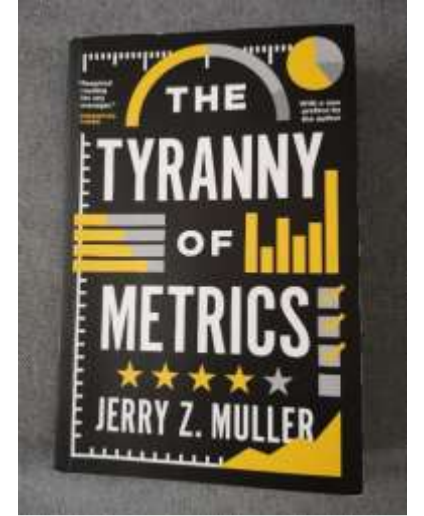
Critiques of metrics

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

(Greek concept of metis)



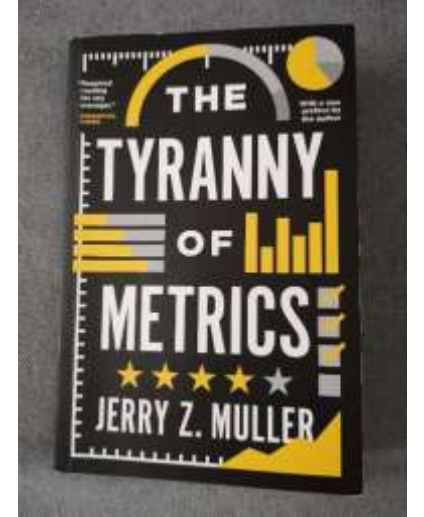
Unintended consequences: a litany



- 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

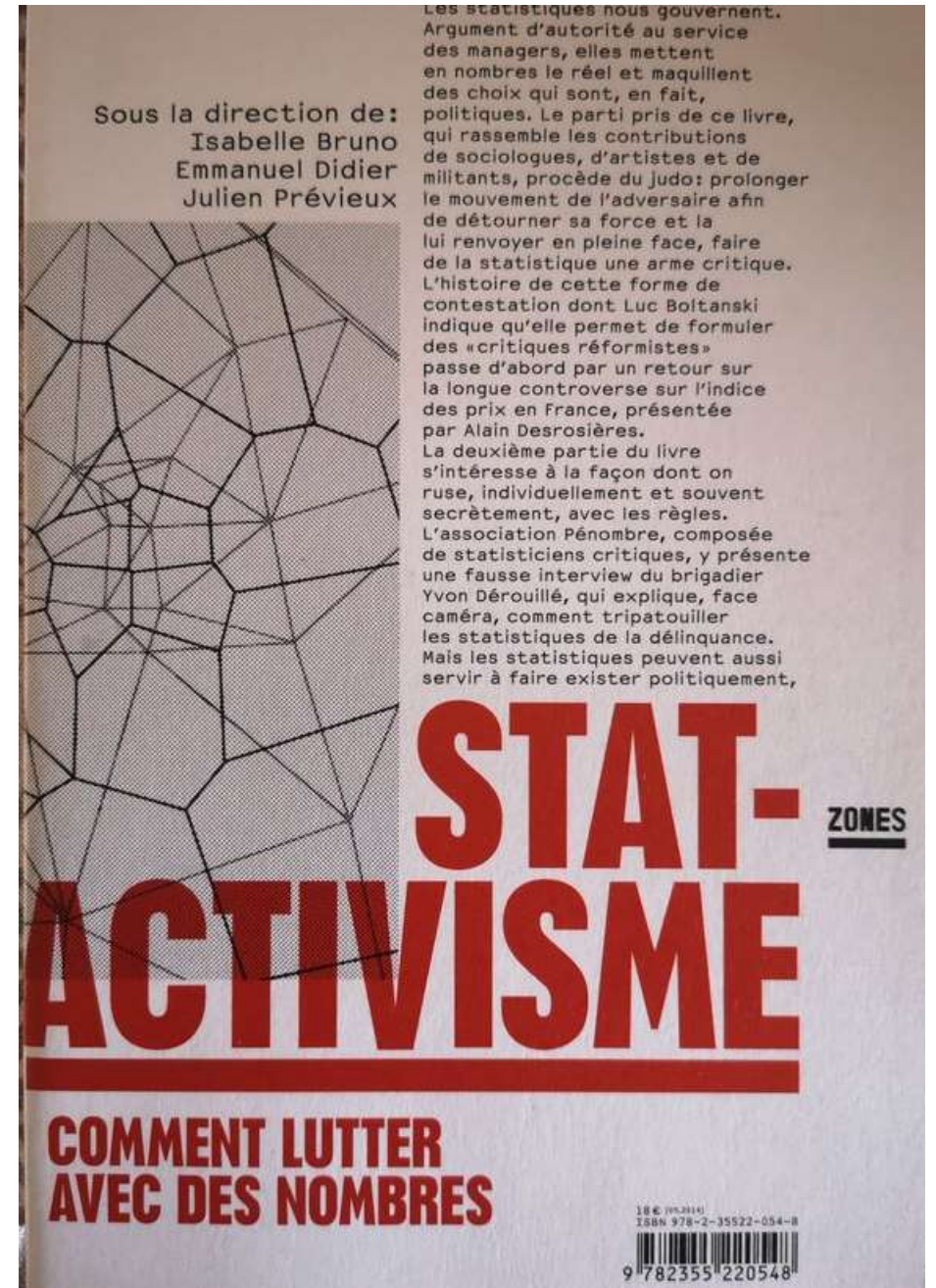
T. M. Porter, “Funny Numbers,” *Culture Unbound*, vol. 4, pp. 585–598, 2012

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

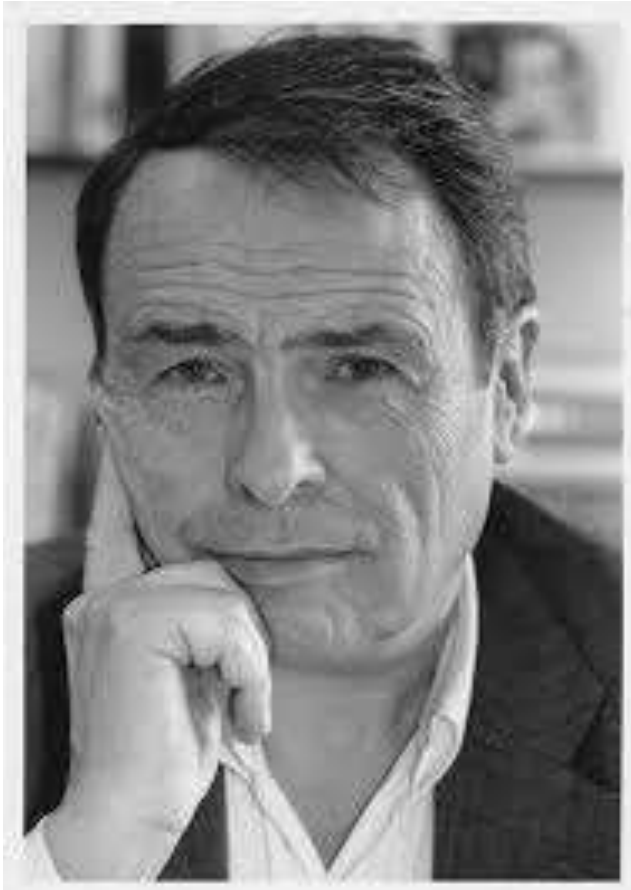
Do we need a movement of resistance?

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

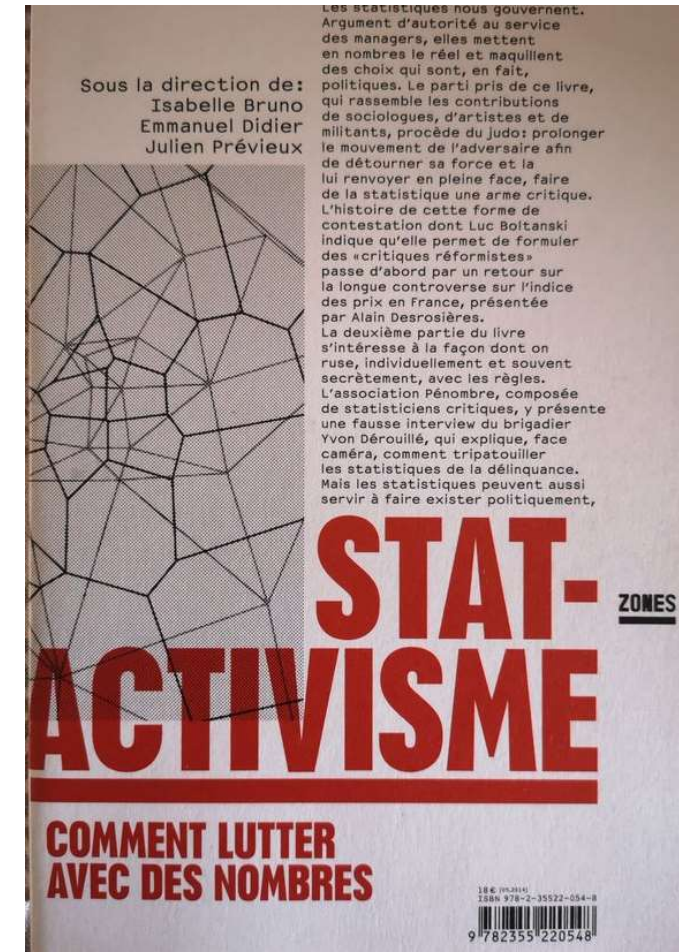
(Sociology must be fun)



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

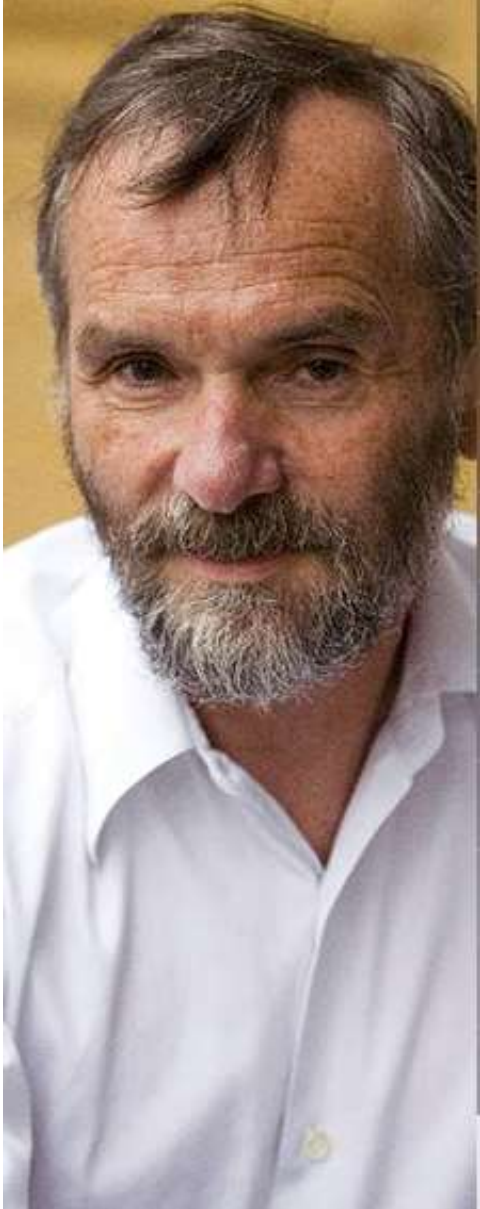


Important:

“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

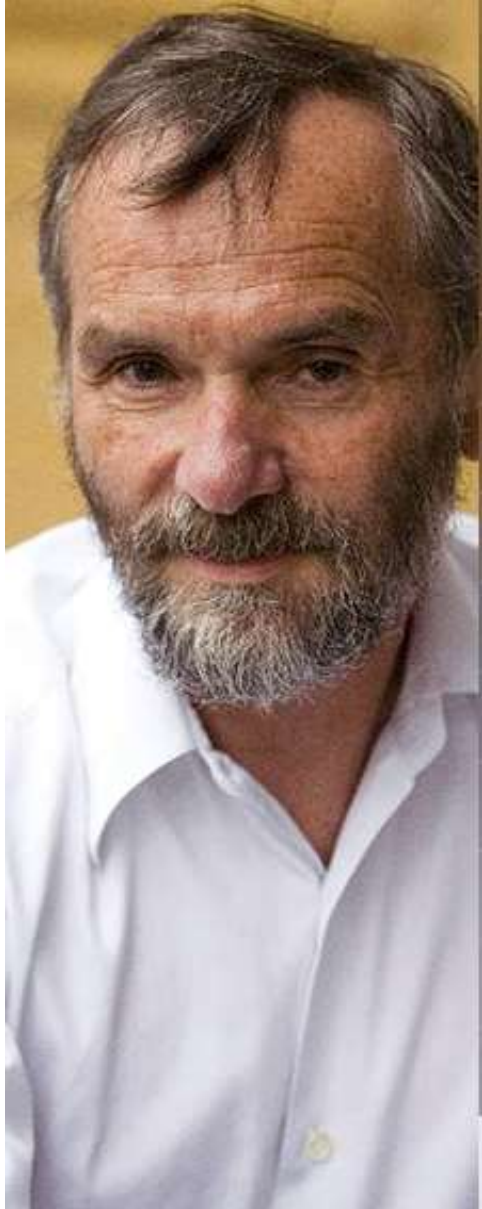


An indictment of the
Total Market and the
normative uses of
economic quantification

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



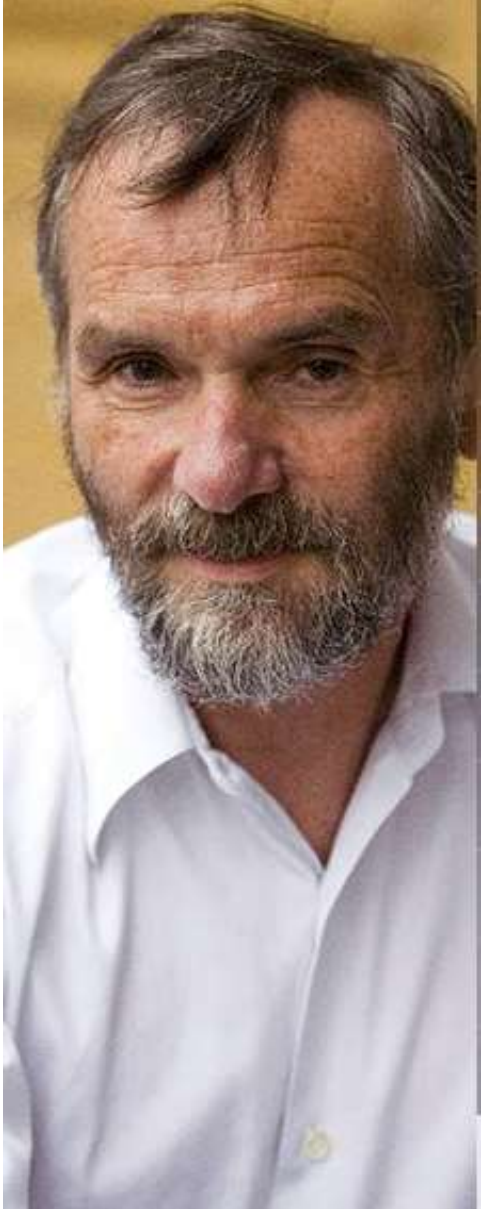
Alain Supiot



...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

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



The End

@andreasaltelli

