







### Forking paths, fishing expeditions and sensitivity analysis

Andrea Saltelli, Florence, December 15, 2023

Seminar at DISIA, University of Florence

#### Where to find this talk: www.andreasaltelli.eu



HOME

ABOUT ME

**PUBLICATIONS** 

**NEWS & VIDEOS** 

RESOURCES

#### 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.

\*\*\*

"The methods by which power insinuates itself



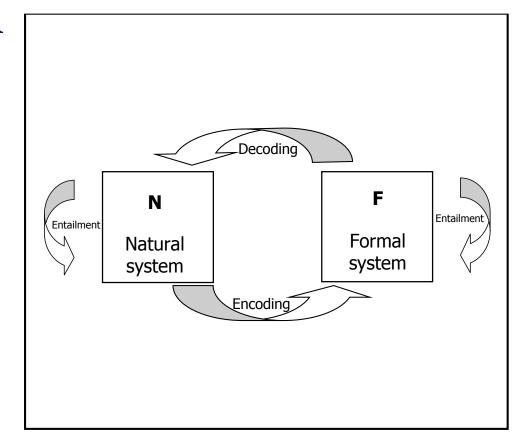
## Modelling is a craft more than a science

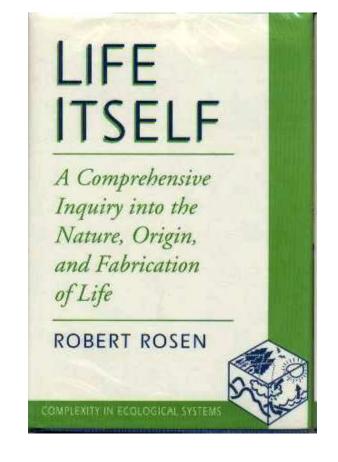
Forking paths?
Met by analysts in building their math or stat models

Fishing expeditions
Another expression for HARKing

Modelling as a craft rather than as a science for

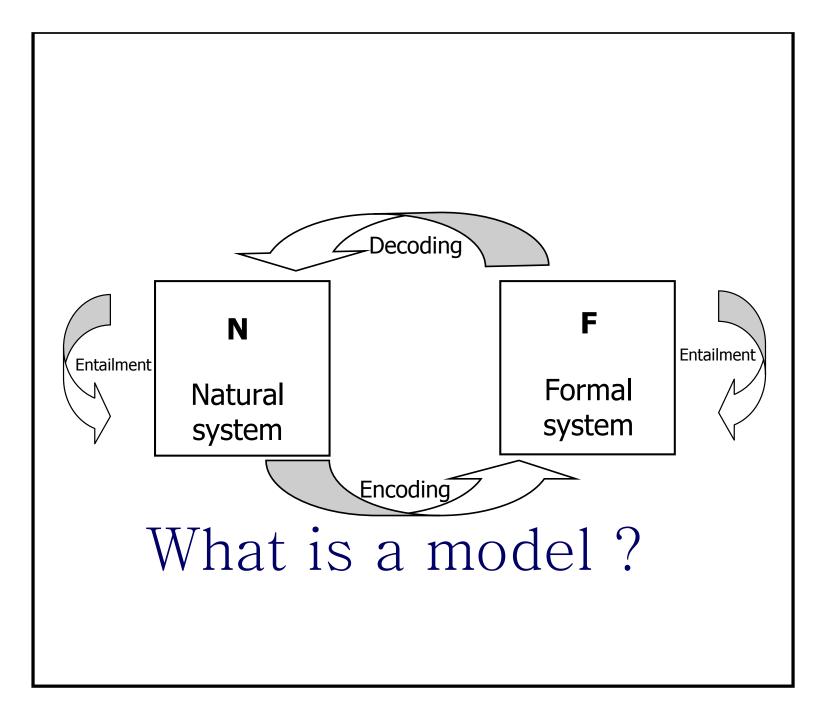
Robert Rosen





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

Louie, A.H. 2010. "Robert Rosen's Anticipatory Systems." Edited by Riel Miller. Foresight 12 (3): 18–29. https://doi.org/10.1108/14636681011049848.





Robert Rosen

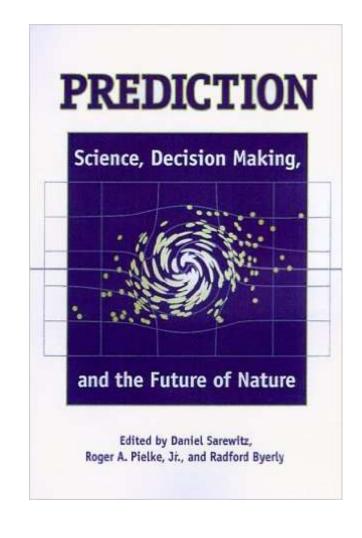
"models are most useful when they are used to challenge existing formulations, rather than to validate or verify them"



Naomi Oreskes

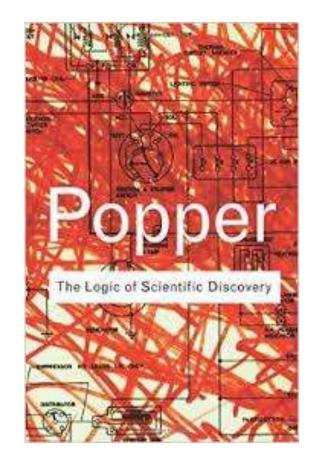
N. Oreskes, K. Shrader-Frechette, and K. Belitz, "Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences," Science, 263, no. 5147, 1994.

## 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)



"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"

# 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

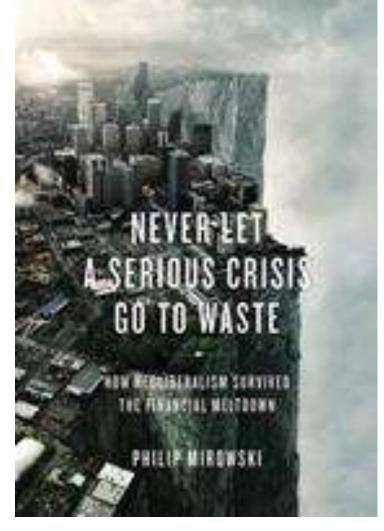
# Caeteris are never paribus

Ceteris paribus or caeteris paribus is a Latin phrase meaning "all other things being equal" or "other things held constant" or "all else unchanged" (Wikipedia)

## The case of DSGE, dynamic stochastic general equilibrium models

Rational expectations of agents Efficient market hypothesis





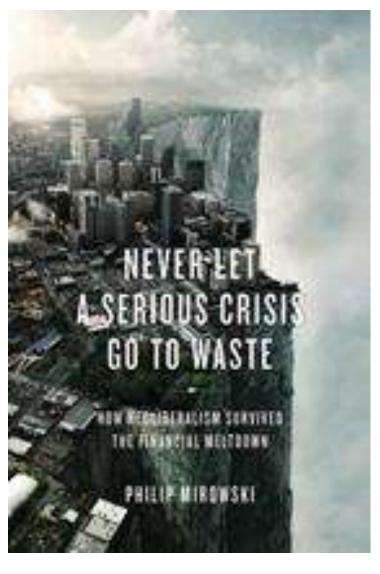
Philip Mirowski

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

The US senate and Queen Elisabeth perplexed…







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

#### Dangers of mathematization of economics









Wolfgang Drechsler

Erik S. Reinert

Paul Romer

Philip Mirowski

W. Drechsler, "On the possibility of quantitative-mathematical social science, chiefly economics," *J. Econ. Stud.*, vol. 27, no. 4/5, pp. 246–259, 2000.

E. S. Reinert, "Full circle: economics from scholasticism through innovation and back into mathematical scholasticism," *J. Econ. Stud.*, vol. 27, no. 4/5, pp. 364–376, Aug. 2000.

P. Romer, "Mathiness in the Theory of Economic Growth," Am. Econ. Rev., vol. 105, no. 5, pp. 89–93, May 2015.

Mirowski, Philip. 2013. Never Let a Serious Crisis Go to Waste: How Neoliberalism Survived the Financial Meltdown. Verso.





WORKING PAPER
WP 2021/07

#### Altered States: Cartesian and Ricardian dreams

Erik S. Reinert

Tallinn University of Technology

UCL Institute for Innovation and Public Purpose

Monica di Fiore

Institute for Cognitive Sciences and Technologies, Consiglio Nazionale delle Ricerche

Andrea Saltelli

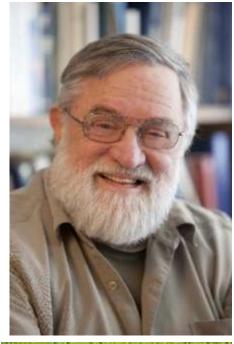
Open Evidence Research, Universitat Oberta de Catalunya (UOC)

Jerome R. Ravetz

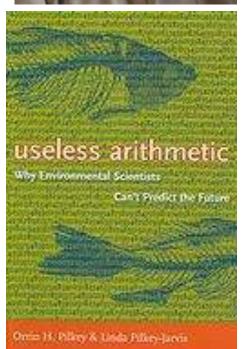
Institute for Science, Innovation and Society, University of Oxford

## Don't confuse the map with the territory

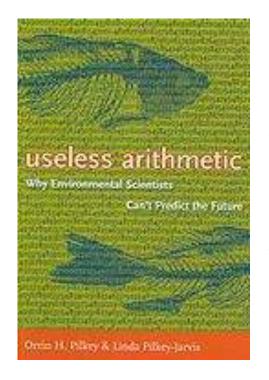
If you do, sensitivity analysis will not save you



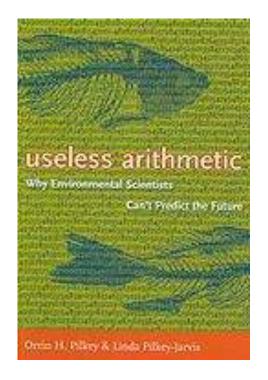
Orrin H. Pilkey



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



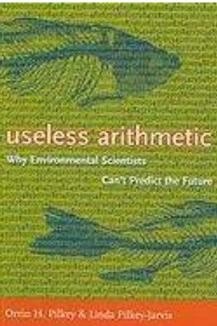
<<It is important, however, to recognize
that the sensitivity of the parameter in the
equation is what is being determined, not
the sensitivity of the parameter in
nature>>

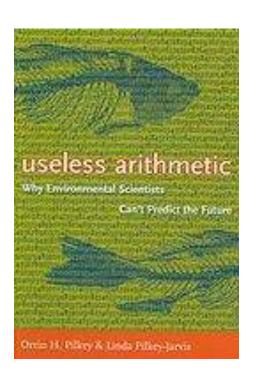


<< "If the model is wrong or if it is a poor representation of reality, determining the sensitivity of an individual parameter in the model is a meaningless pursuit>> One of the examples discussed concerns the Yucca Mountain repository for radioactive waste. TSPA model (for total system performance assessment) for safety analysis.

TSPA is Composed of 286 sub-models.

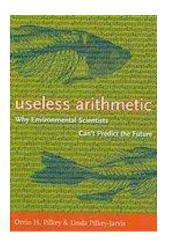


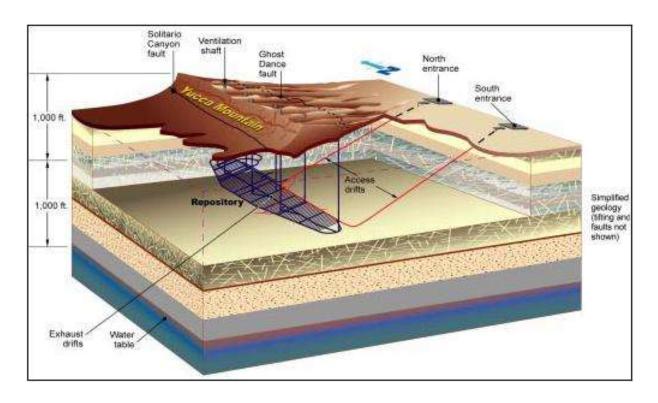




TSPA (like any other model) relies on assumptions → one is the low permeability of the geological formation → long time for the water to percolate from surface to disposal.







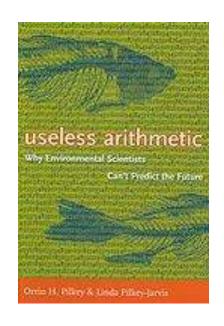
The confidence of the stakeholders in TSPA was not helped when evidence was produced which could lead to an upward revision of 4 orders of magnitude of this parameter

(the <sup>36</sup>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 (as per CUDOS)

Communalism, Universalism, Disinterestedness, Organized Skepticism, from sociology of science, Robert K. Merton.



Steve Rayner

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

### Rayner's (2012) strategies to deal with "uncomfortable knowledge".

Denial, Dismissal, Diversion, Displacement



Model based

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

Displacement: "The model we have developed tells us that real progress is being achieved" (The focus in now the model not the problem).

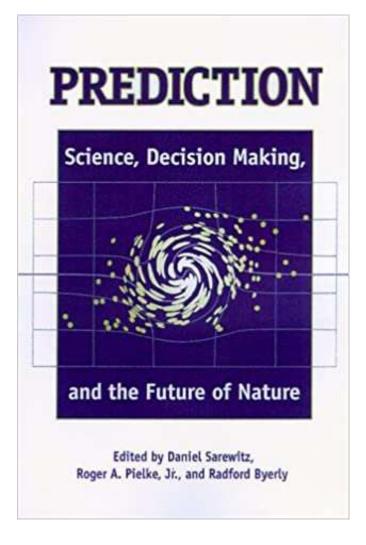
Rayner, S., 2012, Uncomfortable knowledge: the social construction of ignorance in science and environmental policy discourses, Economy and Society, 41:1, 107-125.

Example of displacement: Chesapeake Bay Program (CBP) modelling work

"Bay models are used to track nutrient loads to ensure the cap is not exceeded"

→ The model results – rather than the actual measurements, become the substance of use

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



Model GENESIS for beach erosion



Manipulated to support coastal-engineering projects

It neglected the role of extreme event

Sarewitz, D., Pielke, R. A. & Byerly, R. *Prediction: Science, Decision Making, and the Future of Nature* (Island Press, 2000).

## Beware the size of your model

Mind the conjecture of O'Neil





Comment Open Access Published: 27 August 2019

### A short comment on statistical versus mathematical modelling



Model complexity

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

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.

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.



Current Issue

First release papers

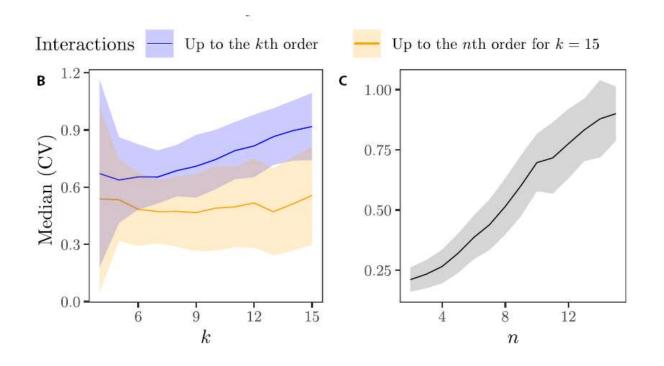
RESEARCH ARTICLE | MATHEMATICS

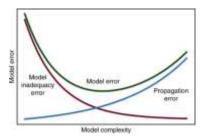


#### Models with higher effective dimensions tend to produce more uncertain estimates



SCIENCE ADVANCES • 19 Oct 2022 • Vol 8, Issue 42 • DOI: 10.1126/sciadv.abn9450

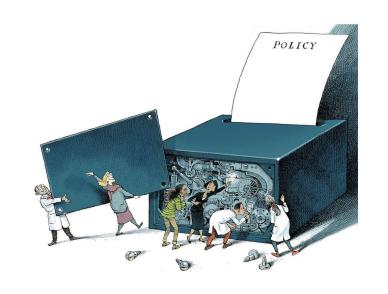






Empirical test using the SAbased concept of effective dimension

## COVID has put mathematical models in the limelight



→ Power & controversy

## Power

The New York Times

# POLICY

## Behind the Virus Report That Jarred the U.S. and the U.K. to Action

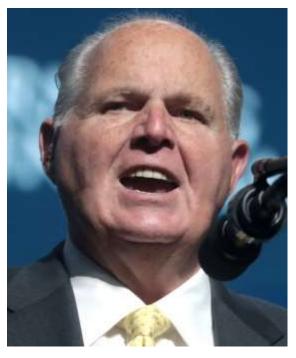
It wasn't so much the numbers themselves, frightening though they were, as who reported them: Imperial College London.

Landler, Mark, and Stephen Castle. 2020. Behind the Virus Report That Jarred the U.S. and the U.K. to Action – The New York Times.

## Conflicts, when questions of urgency, stakes, values and uncertainty collide

"Wild-Ass Covid numbers
... The minute I hear
anybody start talking about
models and modeling, I
blanch"

Rush Limbaugh



Rhodes, Tim, and Kari Lancaster. 2020. "Mathematical Models as Public Troubles in COVID-19 Infection Control: Following the Numbers", Health Sociology Review 1–18. doi: 10.1080/14461242.2020.1764376

#### WILEY Online Library









### What did COVID-19 really teach us about science, evidence and society?

Andrea Saltelli X, Joachim P. Sturmberg, Daniel Sarewitz, John P. A. Ioannidis

First published: 06 June 2023 | https://doi.org/10.1111/jep.13876

During the pandemic crisis of 2020–2023, it is likely that the successes of biomedical science were more than offset by its failures. These failures continue to undermine once-powerful ideals and hopes for science's role in societal betterment. Here, we dissect the underpinnings of these failures and argue that restoring such ideals first requires systemic reform of science itself.



#### Successes

Speed - real-time information sharing, record-time development of vaccines, and unprecedented rapidity in getting results from some large randomized trials of interventions

#### Failures

Demonizing dissenting
Instrumental use of 'following the science'
Proliferation of 'influencers'



"COVID-19 policies allocated sacrifice, privation and suffering across all walks of society [but] radically different responses from nation to nation—from draconian lockdowns, to relatively permissive and flexible pandemic regimes—made obvious to all that the value of scientific evidence was to support what was politically desirable and possible in different contexts

## Mind the assumptions

Assess uncertainty and sensitivity

#### Mind the hubris

Complexity can be the enemy of relevance

## Mind the framing

Match purpose and context

## Mind the consequences

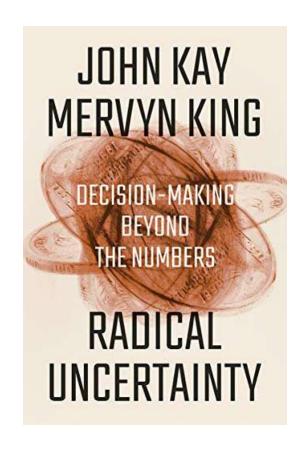
Quantification can backfire.

### Mind the unknowns

Acknowledge ignorance



## Models ask as input information which we don't have - The case of WEBTAG





John Kay

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

Journey Purpose	Weekday						
	7am- 10am	10am- 4pm	4pm-7pm	7pm-7am	Weekday Average	Weekend	All Week
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

Source: J. A. Kay, "Knowing when we don't know," 2012, https://www.ifs.org.uk/docs/john\_kay\_feb2012.pdf

## Mind the assumptions

Assess uncertainty and sensitivity

#### Mind the hubris

Complexity can be the enemy of relevance

## Mind the framing

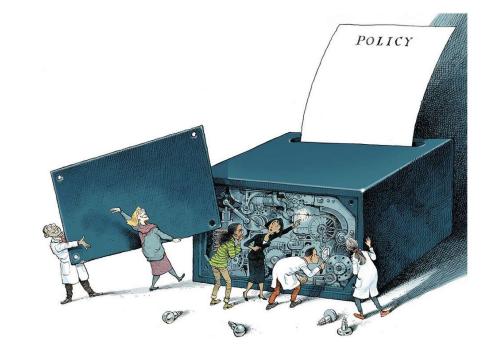
Match purpose and context

## Mind the consequences

Quantification can backfire.

#### Mind the unknowns

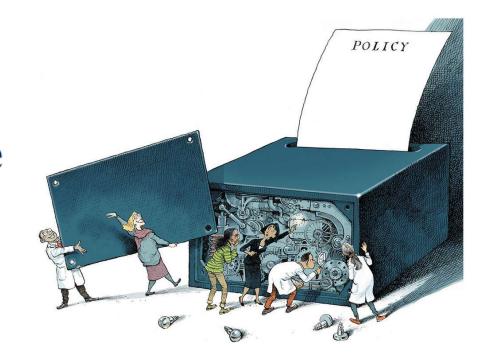
Acknowledge ignorance



## Mind the hubris

## Complexity can be the enemy of relevance

··· many are seduced by the idea of adding complexity in an attempt to capture reality more accurately, but ···



Propagation

error

#### SUPPLEMENTARY INFORMATION

>260 references

Model error

Model

inadequacy

error

Model error

Model complexity

## Mind the assumptions

Assess uncertainty and sensitivity

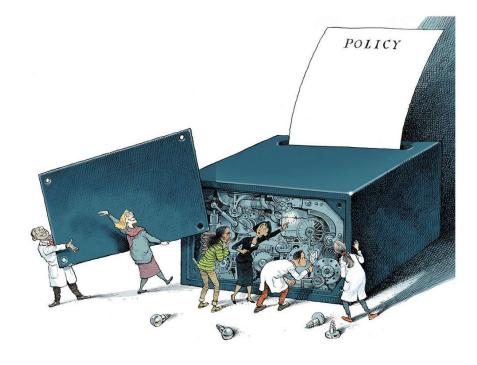
#### Mind the hubris

Complexity can be the enemy of relevance



## Mind the framing

Match purpose and context



## Mind the consequences

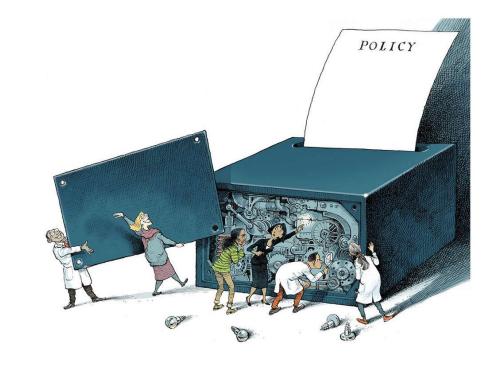
Quantification can backfire.

#### Mind the unknowns

Acknowledge ignorance

## Mind the framing

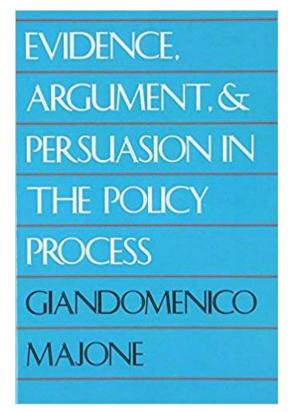
Match purpose and context



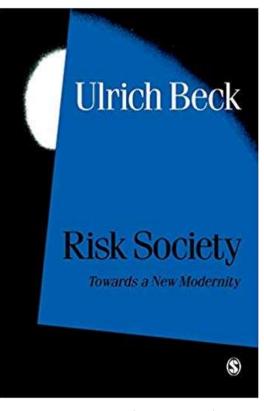
··· models will reflect the interests, disciplinary orientations and biases of the developers···

SUPPLEMENTARY INFORMATION

## From Ulrich Beck to Giandomenico Majone: the technique is never neutral







1992 (1986)



Ulrich Beck (1944 –2015)



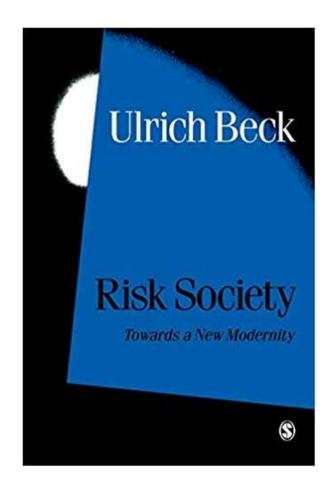
Environmental Science & Policy Volume 106, April 2020, Pages 87-98



The technique is never neutral. How methodological choices condition the generation of narratives for sustainability

Andrea Saltelli <sup>a, b</sup>  $\stackrel{\bowtie}{\sim}$   $\stackrel{\bowtie}{\bowtie}$ , Lorenzo Benini <sup>c</sup>, Silvio Funtowicz <sup>a</sup>, Mario Giampietro <sup>d, e</sup>, Matthias Kaiser <sup>a</sup>, Erik Reinert <sup>a, f</sup>, Jeroen P. van der Sluijs <sup>a, g, h</sup>

"It is not uncommon for political programs to be decided in advance simply by the choice of what expert representatives are included in the circle of advisers."



1992 (1986)



Ulrich Beck (1944 –2015)

# The technique is never neutral. How methodological choices condition the generation of narratives for sustainability



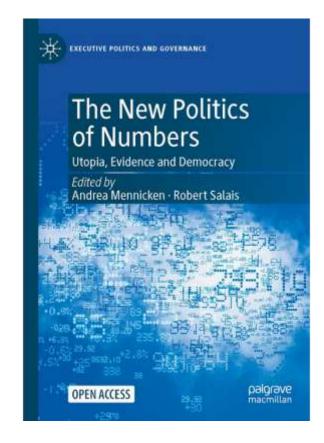
Combine more lenses, including Post-normal science (PNS), Bioeconomics, and Non-Ricardian economics

Since the technique is never neutral a technical proof of quality is illusory without a parallel investigation of normative quality

**Technical Quality** 

Normative quality

How the numbers of neoliberalism (New Public Management) constitute a regime of ademocracy; the example of indicators of employment



Salais, R. (2022). "La donnée n'est pas un donné": Statistics, Quantification and Democratic Choice. In *The New Politics of Numbers: Utopia, Evidence and Democracy*, Andrea Mennicken and Robert Salais, Palgrave Macmillan, pp. 379-415.

## Mind the assumptions

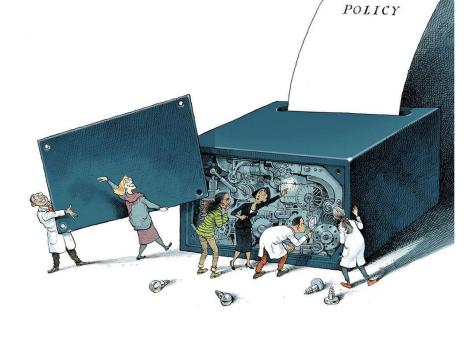
Assess uncertainty and sensitivity

#### Mind the hubris

Complexity can be the enemy of relevance



Match purpose and context





### Mind the consequences

Quantification can backfire.

#### Mind the unknowns

Acknowledge ignorance

## Mind the consequences

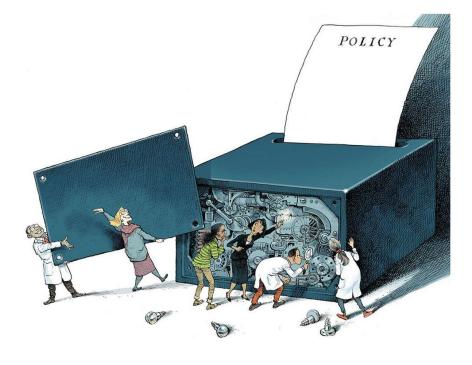
## Quantification can backfire.

« Back to Article WIRED MAGAZINE: 17.03

Recipe for Disaster: The Formula That Killed Wall Street

By Felix Salmon 02-23.00







Here's what killed your 401(k) David X Li's Gaussian copula function as first published in 2000. Investors exploited it as a quick—and fatally flawed—way to assess risk. A shorter version appears on this month's cover of Wired.

Here is what killed your 401(k)...

Li's Gaussian copula function ···

Nassim Nicholas Taleb, hedge fund manager and author of *The Black Swan*, is particularly harsh when it comes to the copula. "People got very excited about the Gaussian copula because of its mathematical elegance, but the thing never worked," he says. "Co-association between securities is not measurable using correlation," because past history can never prepare you for that one day when everything goes south. "Anything that relies on correlation is charlatanism."

Felix Salmon, Wired, February 2009



Source: https://www.wired.com/2009/02/wp-quant/

## Mind the assumptions

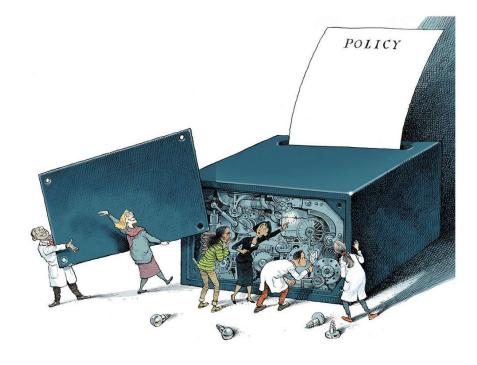
Assess uncertainty and sensitivity

#### Mind the hubris

Complexity can be the enemy of relevance

## Mind the framing

Match purpose and context



## Mind the consequences

Quantification can backfire.

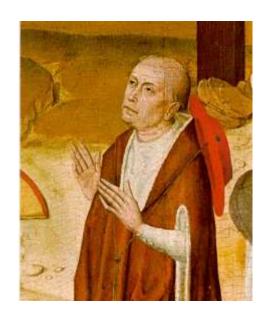


Mind the unknowns

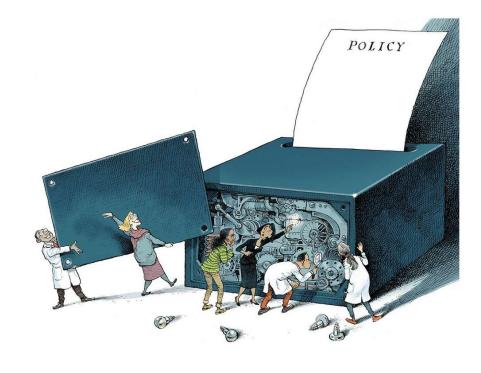
Acknowledge ignorance

## Mind the unknowns

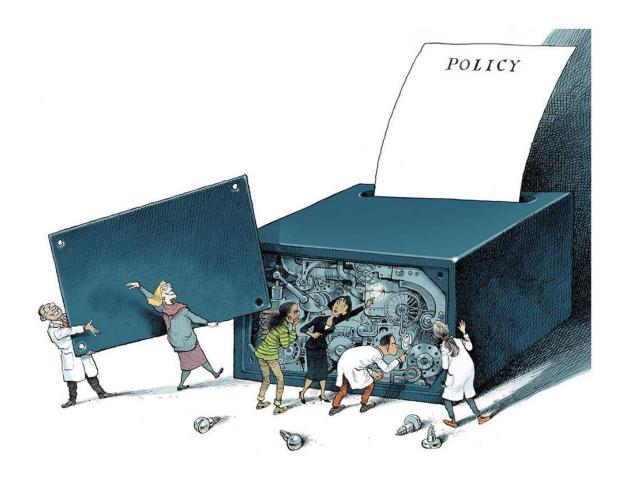
## Acknowledge ignorance







From Socrates's "knowing of not knowing" to Nicolaus Cusanus' Docta Ignorantia was a virtue until Descartes





## the politics of modelling

numbers between science and policy

edited by
Andrea Saltelli
& Monica Di Fiore

OXFORD

## "There is no number-answer to your question"





Anthony Fauci

March 12, 2020, Anthony Fauci before the House Oversight and Reform Committee https://archive.org/details/CSPAN\_20200314\_141500\_Dr.\_Redfield\_Dr.\_Fauci\_\_Other s\_Testify\_on\_Coronavirus\_Response\_Part\_1

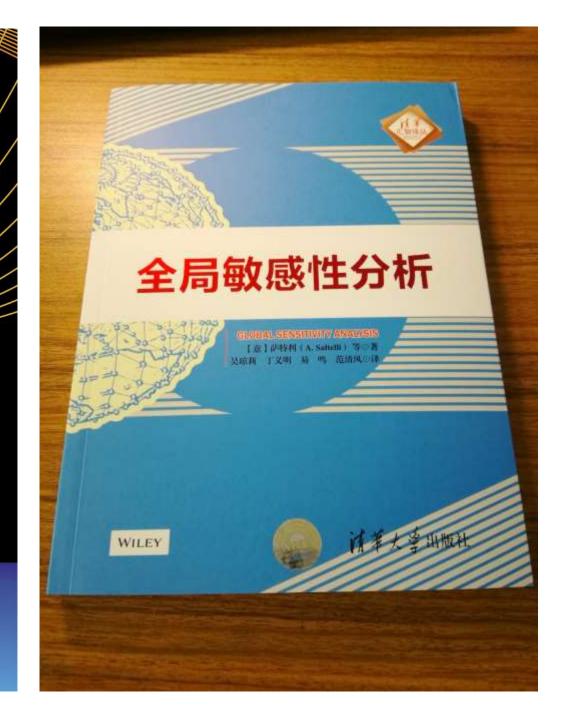
## Sensitivity analysis



## GLOBAL SENSITIVITY ANALYSIS

The Primer

**WILEY** 



## Sensitivity auditing

## EC impact assessment guidelines: sensitivity analysis & auditing



European Commission. November 2021. "Better Regulation: Guidelines and Toolbox." https://ec.europa.eu/info/law/law-making-process/planning-and-proposing-law/better-regulation-why-and-how/better-regulation-guidelines-and-toolbox\_en

## Sensitivity auditing in the EC Guidelines (p. 563)

"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."

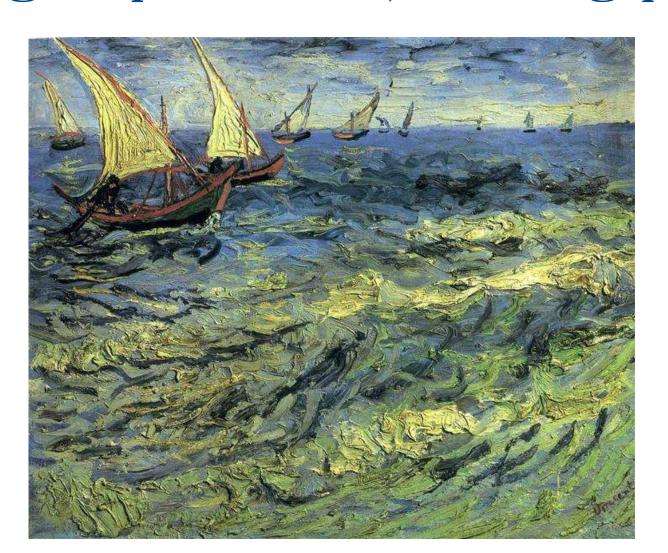


## The rules of sensitivity auditing

- 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.

## Why is all this needed? Fishing expeditions, forking paths ...





(1899 - 1986)

The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time\*

Andrew Gelman<sup>†</sup> and Eric Loken<sup>‡</sup> 14 Nov 2013 The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time\*

Andrew Gelman<sup>†</sup> and Eric Loken<sup>‡</sup> 14 Nov 2013

Why this matters?





**RESEARCH ARTICLE** 

SOCIAL SCIENCES



## Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty

Edited by Douglas Massey, Princeton University, Princeton, NJ; received March 6, 2022; accepted August 22, 2022



"Will different researchers [73 teams] converge on similar findings when analyzing the same data?

. . .

...teams' results varied greatly, ranging from large negative to large positive effects"

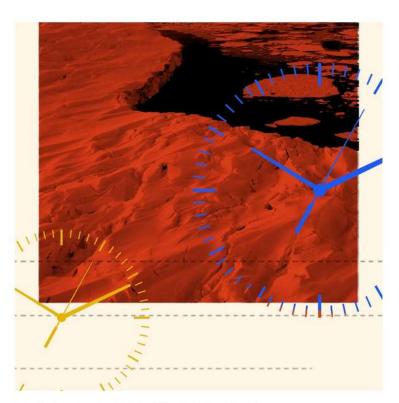
(Massey et al. 2022)

# Do we live immersed in fantastic numbers?

## 'The Most Important Number You've Never Heard Of'

## "social cost of carbon:

Sept. 17, 2021



=\$56 a ton on average at a 3 percent discount rate

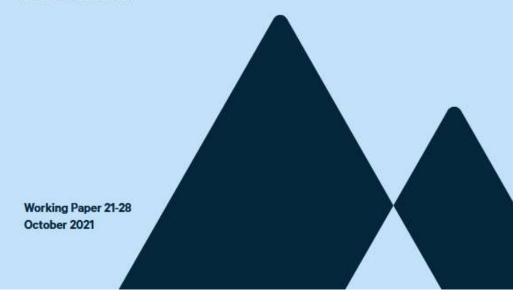
=\$171 a ton on average at a 2 percent discount rate"

The New Hork Times



# The Social Cost of Carbon: Advances in Long-Term Probabilistic Projections of Population, GDP, Emissions, and Discount Rates

Kevin Rennert, Brian C. Prest, William A. Pizer, Richard G. Newell, David Anthoff, Cora Kingdon, Lisa Rennels, Roger Cooke, Adrian E. Raftery, Hana Ševčíková, and Frank Errickson



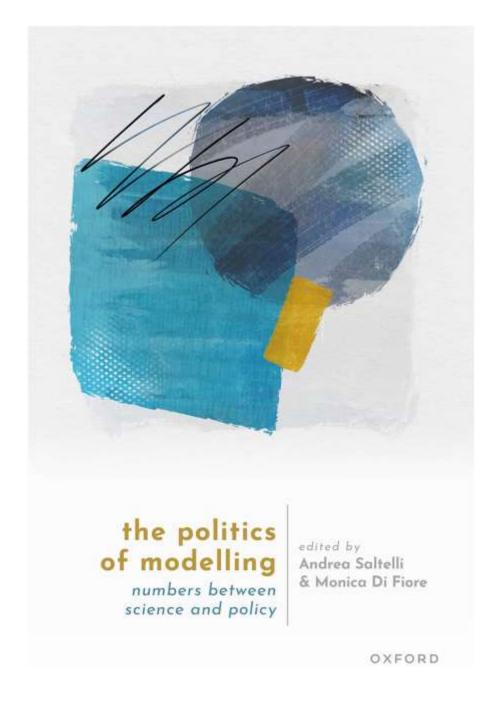
Averaged till year 2300

Feeds into policy design

We have perhaps reached a complex epistemic state, where on the one hand 'everybody knows' that some numbers are pseudo-precise and that numbers can be gamed, while the game works only because most people don't know about it



Jerome R. Ravetz



## The politics of modelling. Numbers between science and policy

Andrea Saltelli and Monica Di Fiore Eds.



From
Epilogue: these special
models, by the editors



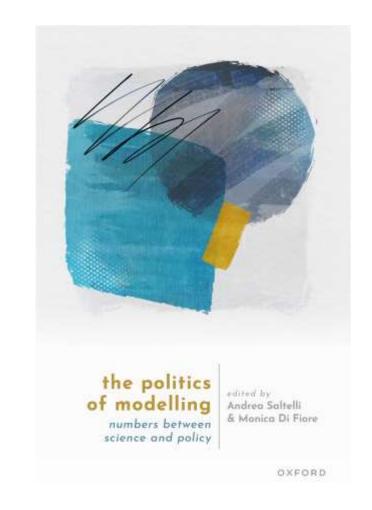
## Are models 'special'?

Unlimited repertoire of methods

Not a discipline

Escape sociology of quantification

Epistemic authority from mathematics



## Consequences?

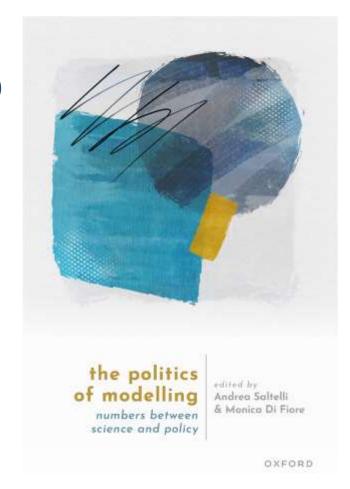


A pretence of neutrality (Luhmann's deparadoxification)

No antibodies to fight degeneration (Ravetz)

A ground for trans-science (Weinberg)

Ritual use (Gigerenzer)



## Solutions to resolve the state of exception

- Thinking about reproducibility of models (Ioannidis)
- Complexity of interpretation rather than complexity of construction
- Follow the example of statistics' statactivism
- Reciprocal domestication between models and society
- Defog the mathematics of uncertainty (Funtowicz & Ravetz, 1990)
- Practice assumption hunting / modelling of the modelling process / sensitivity analysis and auditing





### Reproducibility

is a necessary condition for

### Transparency

is a necessary condition for

### Legitimacy

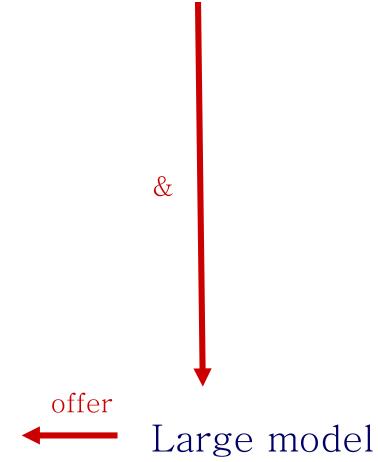
is a necessary condition for



Gianus Bifrons, Vatican Museum. Source: Wikipedia Common

Epistemic authority

## Important institution



Different political economies of modelling



## The End



https://mstdn.social/@AndreaSaltelli/

Abstract: Based on recent works on the potential multiplicity of results from statistical or mathematical models (Breznau et al., 2022) we show how sensitivity analysis (Saltelli et al., 2021) and the related approach of sensitivity auditing can help in a quest that we call "Modelling of the modelling process". We shall revisit some general principles and practices of sensitivity analysis, then introduce sensitivity auditing also following a recent volume published by Oxford in August 2023 (Saltelli & Di Fiore, 2023).

Breznau, N., Rinke, E. M., Wuttke, A., Nguyen, H. H. V., Adem, M., Adriaans, J., Alvarez-Benjumea, A., Andersen, H. K., Auer, D., Azevedo, F., Bahnsen, O., Balzer, D., Bauer, G., Bauer, P. C., Baumann, M., Baute, S., Benoit, V., Bernauer, J., Berning, C., ··· Żółtak, T. (2022). Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty. Proceedings of the National Academy of Sciences, 119(44), e2203150119. https://doi.org/10.1073/pnas.2203150119

Saltelli, A., & Di Fiore, M. (Eds.). (2023). The politics of modelling. Numbers between science and policy. Oxford University Press.

Saltelli, A., Jakeman, A. J., Razavi, S., & Wu, Q. (2021). Sensitivity analysis: A discipline coming of age. Environmental Modelling and Software, 146,(105226). https://doi.org/10.1016/j.envsoft.2021.105226