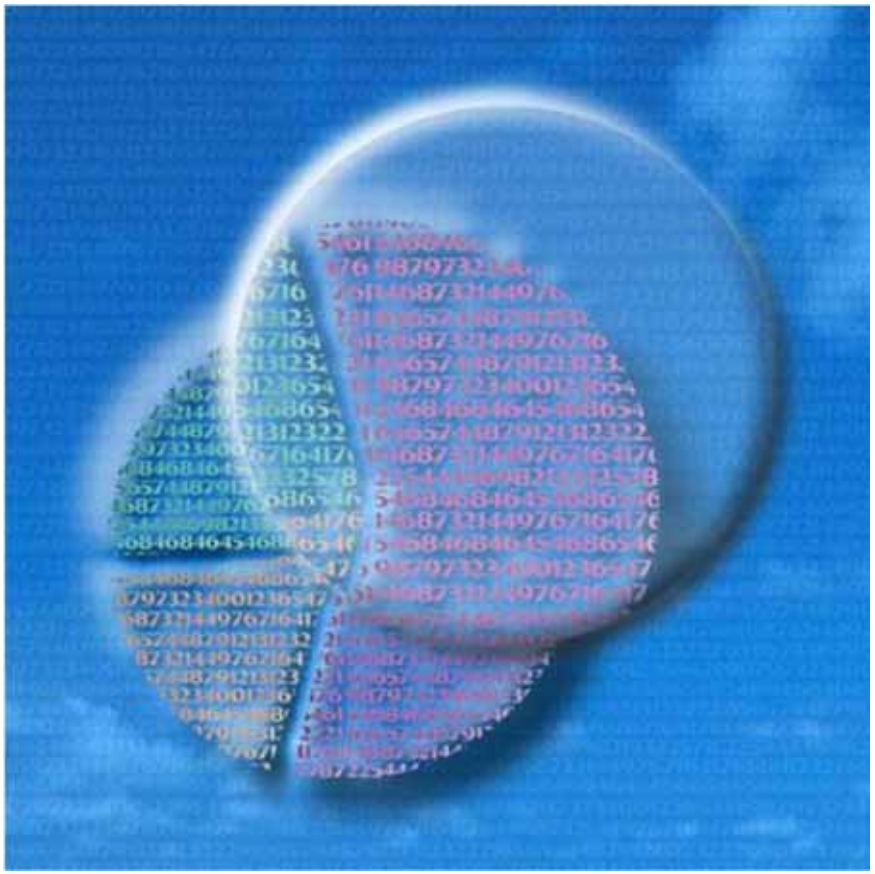
Sensitivity auditing



Andrea Saltelli, European Commission,
Joint Research Centre

Syllabus on impact assessment

April 4th 2014
Brussels, Berlaymont Room 6/A

Sensitivity auditing
or
Impact Assessment Modelling Scrutiny Service
since February 2013

Minutes of the IAWG of 7 February 2013:

"M. Gremminger informed the group that in the course of 2012 the SG has tested JRC competence in providing critical appraisal of modelling-rich IA cases. As a result he suggests that this type of analysis be anticipated at the very beginning of the impact assessment cycle, with JRC providing to each service willing to use it an introduction to the issue of robustness ands plausibility. This should take place possible at the inception (framing) phase."

George Box: ‘all models are wrong, some are useful’

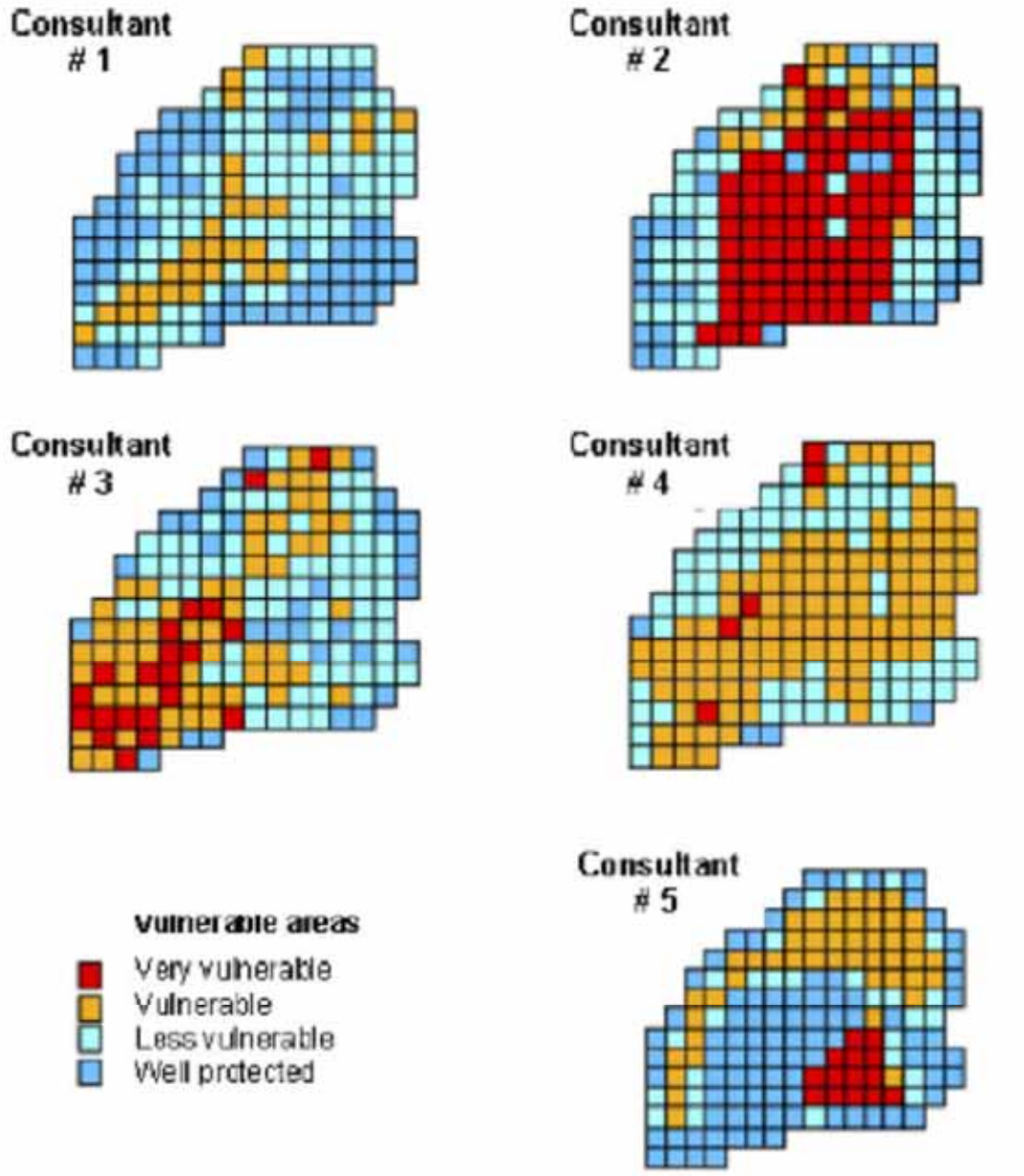
Box, G.E.P., Robustness in the strategy of scientific model building, in Robustness in Statistics, R.L. Launer and G.N. Wilkinson, Editors. 1979, Academic Press: New York.

Copernicus Institute

Model structure uncertainty…

5 consultants, each using a different model were given the same question:
"which parts of this particular area are most vulnerable to pollution and need to be protected?"

(Refsgaard et al, 2006)



Courtesy of Dr. Jeroen P. van der Sluijs (1965)
Copernicus Institute of Sustainable Development, Utrecht University

The IFPRI had raised about $460,000 for the modeling, which would have provided insights to help policymakers […]

NEWSFOCUS

Dueling Visions For a Hungry World

Sparks began to fly when scientists and activists against genetically modified crops came together to assess agricultural knowledge and the role of biotech in development

When economist Carl Pray heard about plans for the first international assessment of agricultural research, a gold standard sprang to mind: the Intergovernmental Panel on Climate Change (IPCC). But things didn't turn out the way he expected.

IPCC has been pivotal in proving that climate change is real and linking it to human activities. As an agricultural economist at Rutgers University who has worked in many poor countries, Pray is convinced that agricultural research—and genetic modification in

mentally, socially and economically sustainable development through the generation, access to, and use of agricultural knowledge, science and technology?" Critics say this broad mandate made conflict inevitable and stunted the assessment's analytical rigor.

On several key issues, consensus proved elusive. Industry scientists and some academics—mainly agricultural economists and plant biologists—believe the assessment was "hijacked" by participants who oppose genetically modified (GM) crops and other common

the outcome. They note that the voice and experience of small-scale farmers, particularly women, have finally been brought to the fore by the assessment. "It really deals with issues of power, influence, and benefits," says Marcia Ishii-Eiteman of the Pesticide Action Network North America in San Francisco, California. Toby Kiers, who studies sustainable agriculture at Vrije University in Amsterdam, the Netherlands, agrees. "For technology to be most effective, farmers must be at the center, influencing how it is developed, delivered, and

[…] But Greenpeace […] objected that the models were not "transparent".

Source: Dueling visions for an hungry world, Erik Stokstad, 14 MARCH 2008, 319 SCIENCE

we reduce hunger and poverty, improve rural livelihoods, and facilitate equitable, environ-

community-based knowledge.

- Create space for diverse voices and include social scientists in policy.

suggested that the bank review the entire range of agricultural technologies and policies. Convinced that agricultural research should be considered in the context of the myriad factors

We just can’t predict, says N. N. Taleb, and we are victims of the ludic fallacy, of delusion of uncertainty, and so on. Modelling is just another attempt to ‘Platonify’ reality…

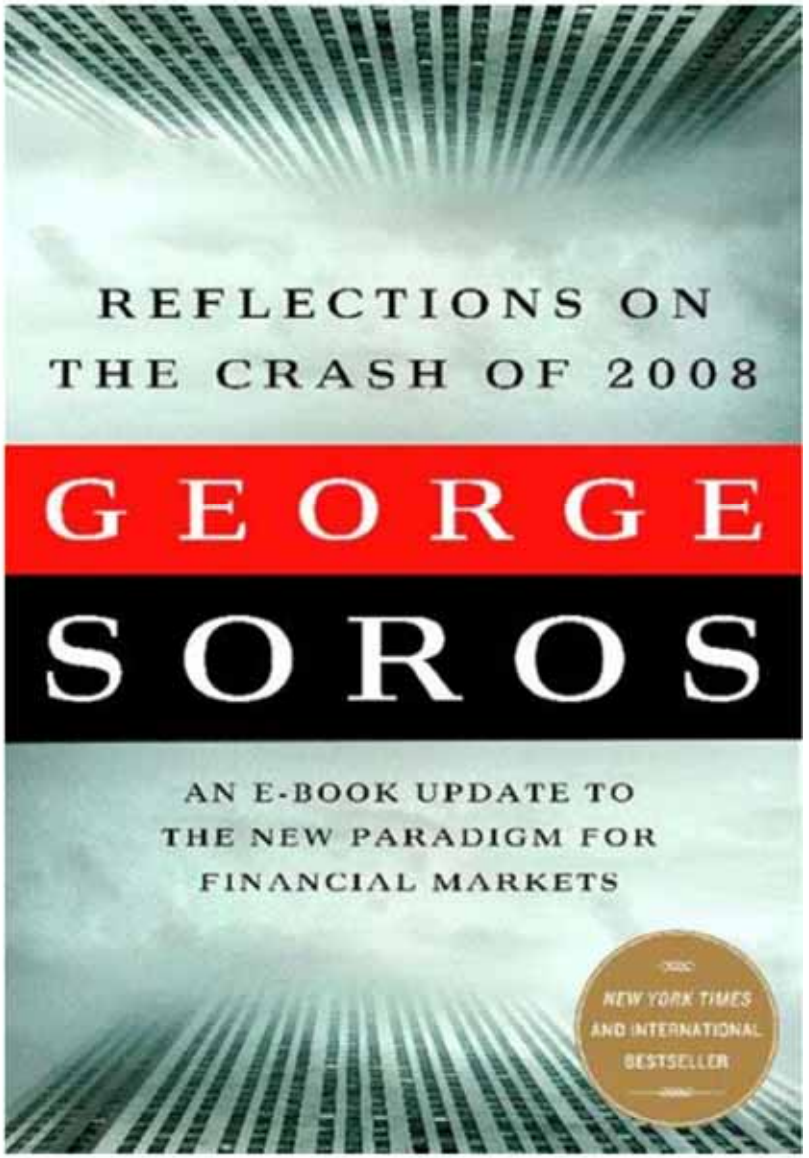
Written before the financial crisis

Nassim Nichola Taleb, The Black Swan, Penguin, London 2007

Postulate of 'radical fallibility':

"Whenever we acquire some useful knowledge, we tend to extend it to areas where it is no longer applicable"

(Taleb's -Platonification')



Models by their nature are like blinders. In leaving out certain things, they focus our attention on other things. They provide a frame through which we see the world.

Joseph E. Stiglitz, 2011, RETHINKING MACROECONOMICS: WHAT FAILED, AND HOW TO REPAIR IT, Journal of the European Economic Association August 2011 9(4):591–645

Caeteris are never paribus!

FT.com
FINANCIAL TIMES

The follies and fallacies of our forecasters

By Samuel Brittan
Published: May 26 2011 23:22 | Last updated: May 26 2011 23:22

Nothing has done more to discredit serious economic analysis than its identification with the guesses about output, employment, prices and so on which politicians feel obliged to make. […] True scientific predictions are conditional. They assert that certain changes […] will, granted other conditions are met, […], lead to a certain state of affairs […]. But they cannot tell us that the required conditions will be fulfilled.

Keynes' take

The rethorical question Keynes asks is (Keynes, 1940):

"It will be remembered that the seventy translators of the Septuagint were shut up in seventy separate rooms with the Hebrew text and brought out with them, when they emerged, seventy identical translations. Would the same miracle be vouchsafed if seventy multiple correlators were shut up with the same statistical material?"

Keynes, J. M. , 1940, On a Method of Statistical Business-Cycle Research. A Comment, The Economic Journal, Vol. 50, No. 197 (Mar., 1940), 154-156.

Mirowski's take

Philip Mirowski
See blog on
*http://www.nakedcapitalism.com/2011/12/philip-mirowski-**the-seekers-or-how-mainstream-economists-have-defended-their-discipline-since-2008-**%E2%80%93%C2%A0part-iv.html*

Mirowski's take

July 20, 2010 hearing on DSGE

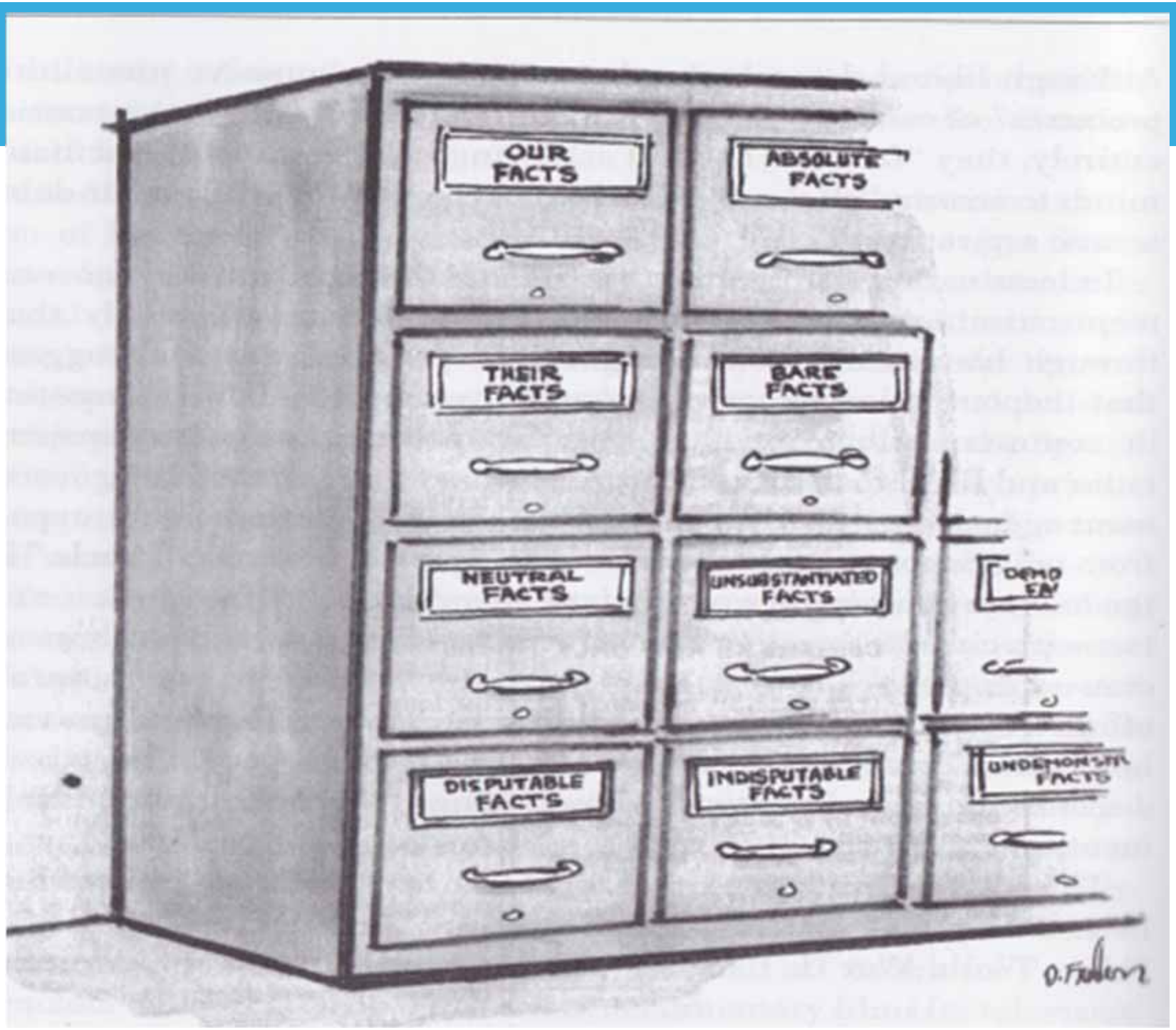
[...] The basic stance of the hearings was defined in the opening comments by Chairman Brad Miller:

"According to the model's most devoted acolytes, the model's insights rival the perfect knowledge Paul described in the First Letter to the Corinthians; but unlike the knowledge Paul described, DSGE's insights are available in the here and now...

Mirowski's take

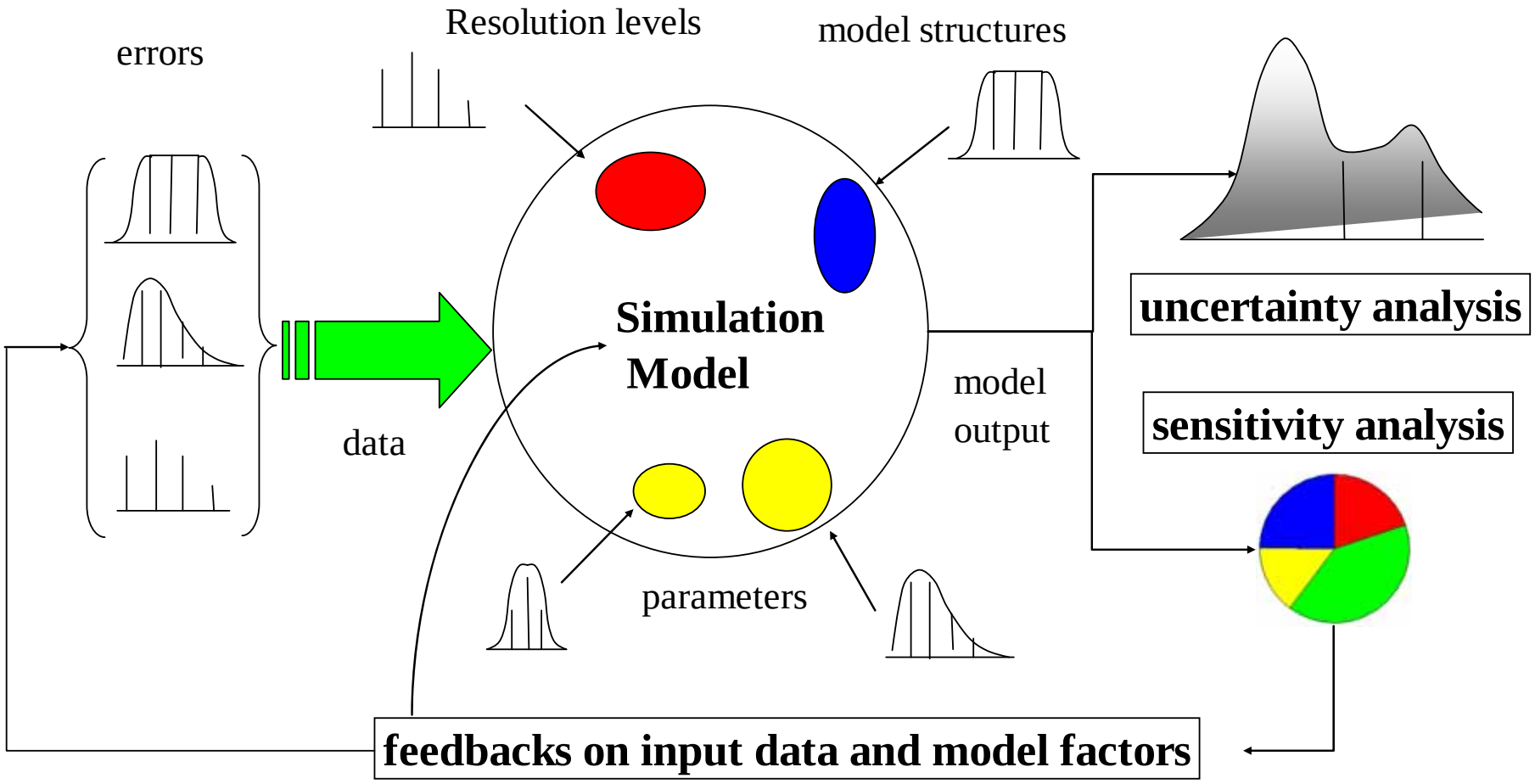
…To be fair, DGSE and similar macroeconomic models were first conceived as theorists' tools. But why, then, are they being relied on as the platform upon which so much practical policy advice is formulated? And what has caused them to become, and to stay, so firmly entrenched? And, finally, the most important question of all: What do we get when we apply the various tools at our disposal to the urgent economic problems we're facing today?"

The Hearing Charter of the House Committee on Science and Technology and sworn testimony of economists Sidney Winter, Scott Page, Robert Solow, David Colander and V.V. Chari


OUR FACTS
ABSOLUTE FACTS
THEIR FACTS
BARE FACTS
NEUTRAL FACTS
UNSUBSTANTIATED FACTS
DISPUTABLE FACTS
INDISPUTABLE FACTS

From sensitivity analysis to sensitivity auditing

Sensitivity Analysis



Sensitivity Analysis

"The study of how the uncertainty in the output of a mathematical model or system (numerical or otherwise) can be apportioned to different sources of uncertainty in its inputs"

Saltelli, A., 2002, Sensitivity analysis for importance assessment. *Risk Analysis*, 22(3):579-590.

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., Funtowicz, S., When all models are wrong: More stringent quality criteria are needed for models used at the science-policy interface, *Issues in Science and Technology*, Winter 2014, 79-85.

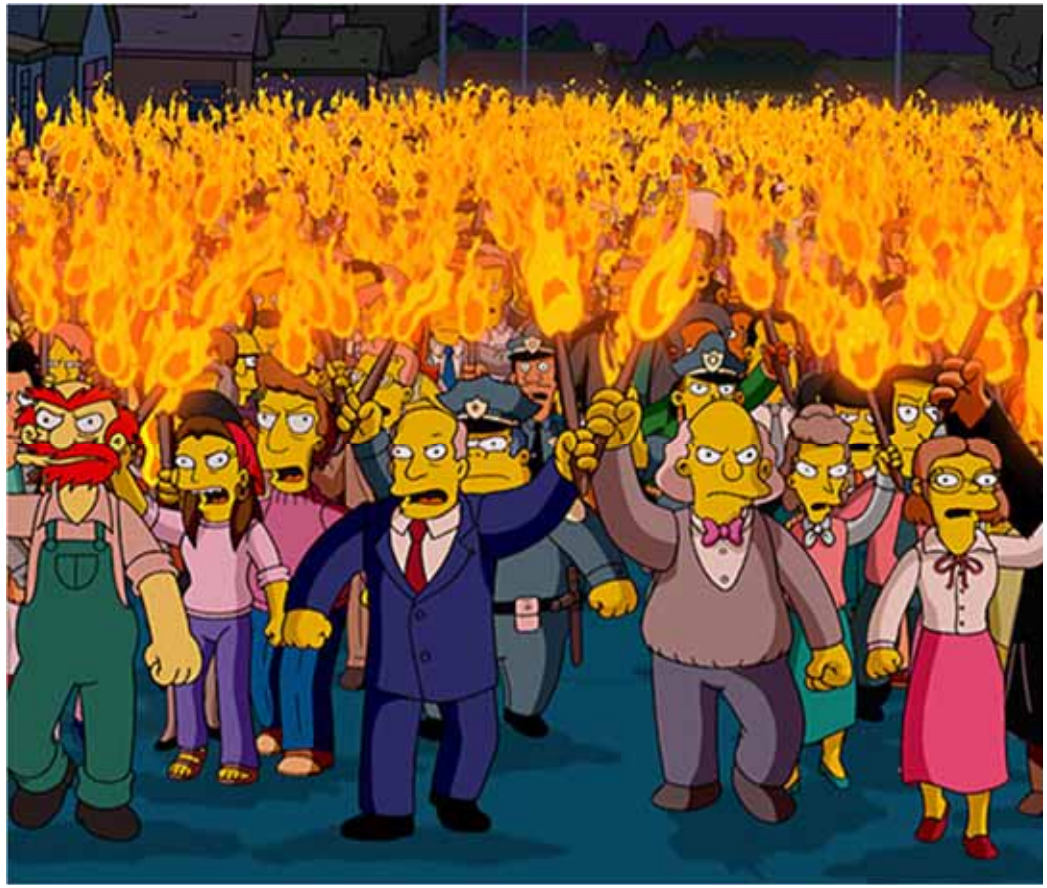
Sensitivity analysis, mandated by existing guidelines as a good practice to use in conjunction to mathematical modelling, is as such insufficient to ensure quality in the treatment of uncertainty of science for policy.

If policy-related science calls for an extension of the traditional internal, peer review-based methods of quality assurance to higher levels of supervision, where extended participation and explicit value judgments are necessary […]

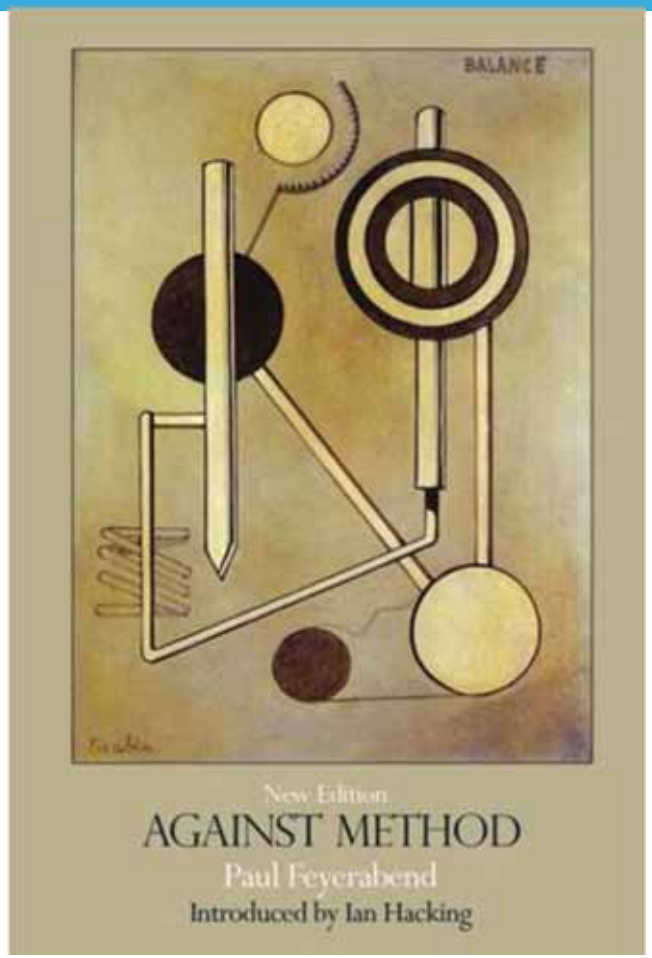
then sensitivity analysis must extend beyond the technical exploration of the space of uncertain assumptions

In an adversarial context not only the nature of the evidence, but also the degree of certainty and uncertainty associated to the evidence will be the subject of partisan interests

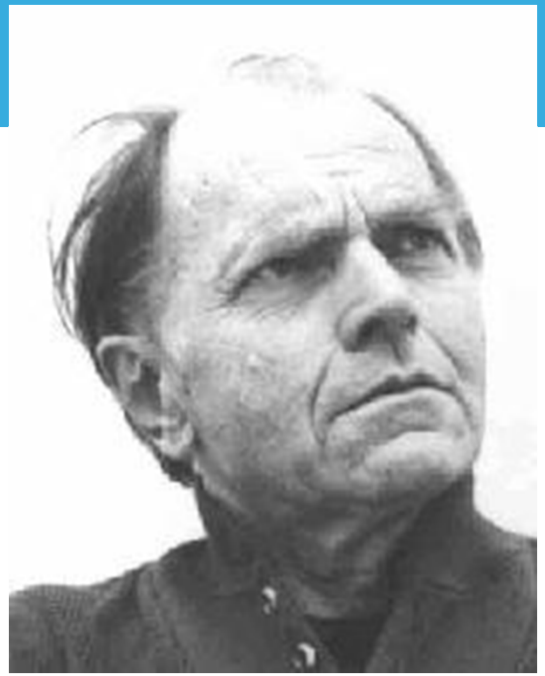
➔ Extended peer review



[…] in a democracy local populations not only will, but also should, use the sciences in ways most suitable to them. The objections that citizens do not have the expertise to judge scientific matters overlooks that important problems often lie across the boundaries of various sciences so that scientists within these sciences don't have the needed expertise either.



Moreover doubtful cases always produce experts from one side, experts for the other side, and experts in between. But the competence of the general public could be vastly improved by an education that exposes expert fallibility instead of acting as if it did not exist. (Paul Feyerabend, Against Method)



Paul Feyerabend

Transactions of the Institute of British Geographers
Royal Geographical Society with IBG

Doing flood risk science differently: an experiment in radical scientific method

S N Lane*, N Odoni*, C Landström**, S J Whatmore**, N Ward† and S Bradley‡

Trans Inst Br Geogr NS 36 15–36 2011
ISSN 0020-2754 © 2010 The Authors.

[…] knowledge regarding flooding was co-produced. This illustrates a way of working with experts, both certified (academic natural and social scientists) and noncertified (local people affected by flooding), […] We reveal a deep and distributed understanding of flood hydrology across all experts, certified and uncertified, …

Joint Research Centre

[…] the purpose of our experiment became as much about creating a new public capable of making a political intervention in a situation of impasse, as it was about producing the solution itself. The practice of knowledge generation, the science undertaken, worked with the hybridisation of science and politics rather than trying to extract science from it.

From sensitivity analysis to sensitivity auditing; Seven rules

1. Check against rhetoric use of mathematical modeling [is the m
obfuscate?]:

2. Adopt an 'assumption hunting' attitude [what
possibly normative assumptions underlying the

3. Detect Garbage In Garbage O
d in order to achieve
a desired inference at a desired

4. Find sensiti

5. A
ake sense of, and possibly replicate, the results
of t

6.
han 'Do the sums right'; is the viewpoint of a relevant
stakeh

7. Fo
he key question answered by the model, exploring holistically the entire space of
the assump
do perfunctory analyses changing one factor at a time].

RULE ONE: Check against rhetorical use of mathematical modelling

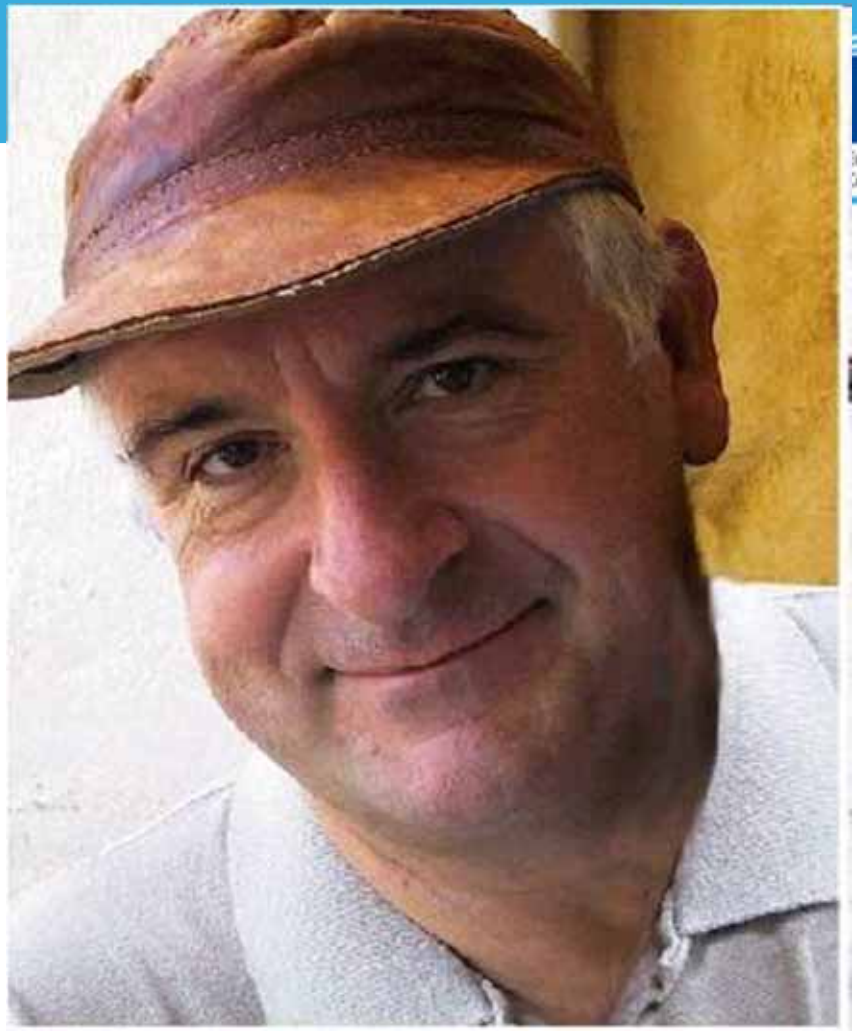
The instrumental use of mathematical modelling to advance one's agenda can be termed rhetorical, or strategic, like the use of Latin by the elites and the clergy in the classic age.

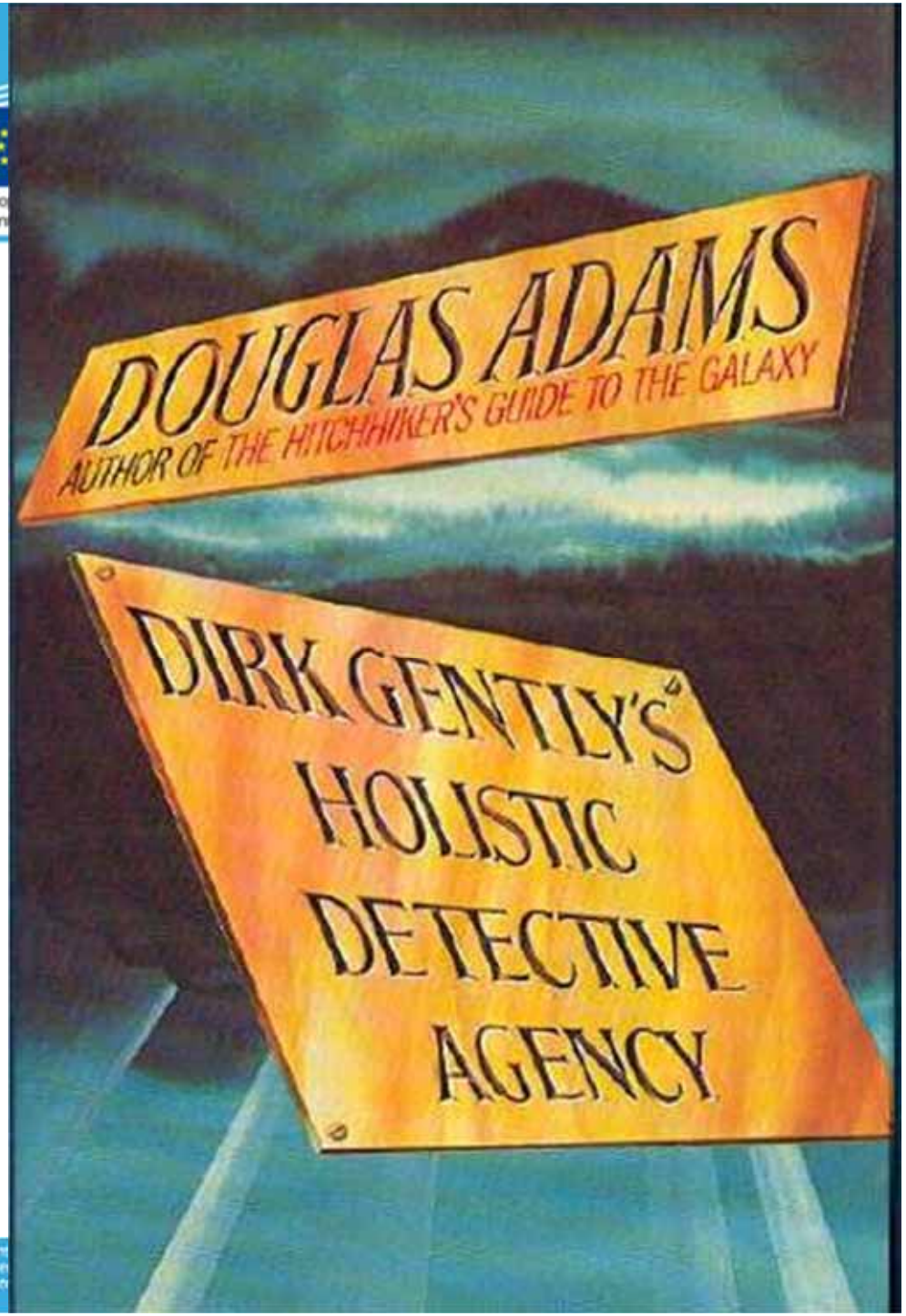
RULE ONE: Check against rhetorical use of mathematical modelling

George M. Hornberger, Professor at University of Viginia

<<[…] most simulation models will be complex, with many parameters, state-variables and non linear relations. Under the best circumstances, such models have many degrees of freedom and, with judicious fiddling, can be made to produce virtually any desired behaviour, often with both plausible structure and parameter values.>>

HORNBERGER and Spear (1981).





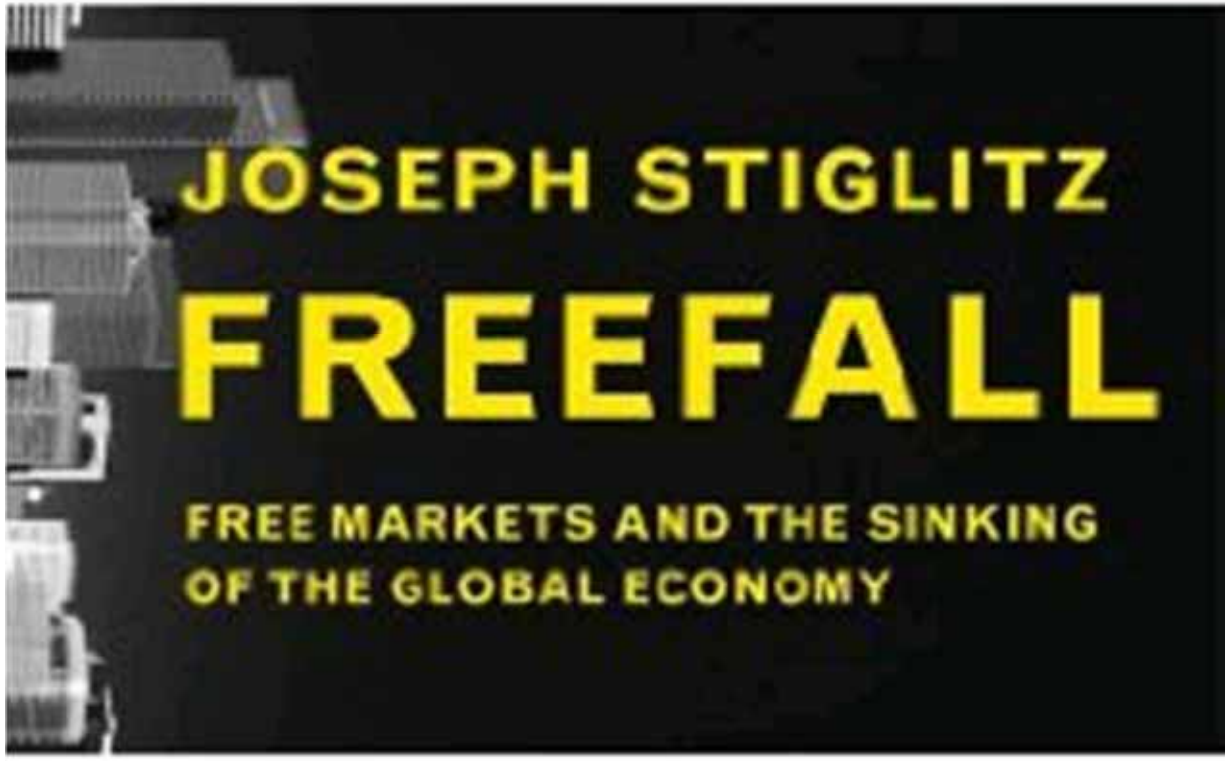
Pocket Books 1987, p.69

RULE ONE: Check against rhetorical use of mathematical modelling

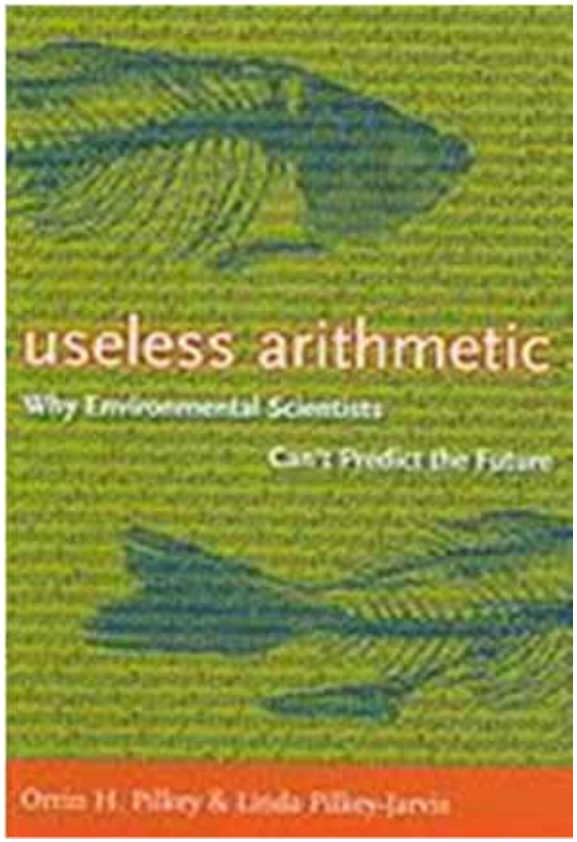
"Well, Gordon's great insight was to design a program which allowed you to specify in advance what decision you wished it to reach, and only then to give it all the facts. The program's task, […], was to construct a plausible series of logical-sounding steps to connect the premises with the conclusion."

‘Perverse incentives and flawed models – accelerated by a race to the bottom’, p. 92

[…] Part of the agenda of computer models was to maximize the fraction of, say, a lousy sub-prime mortgage that could get an AAA rating, then an AA rating, and so forth,[…] This was called rating at the margin, and the solution was still more complexity”, p. 161



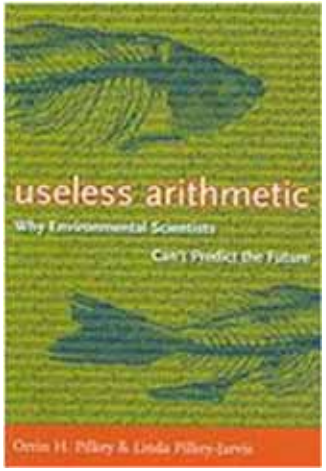
RULE ONE: Check against rhetorical use of mathematical modelling



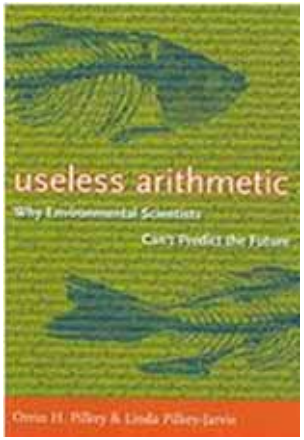
Useless Arithmetic: Why Environmental Scientists Can't Predict the Future
by Orrin H. Pilkey and Linda Pilkey-Jarvis

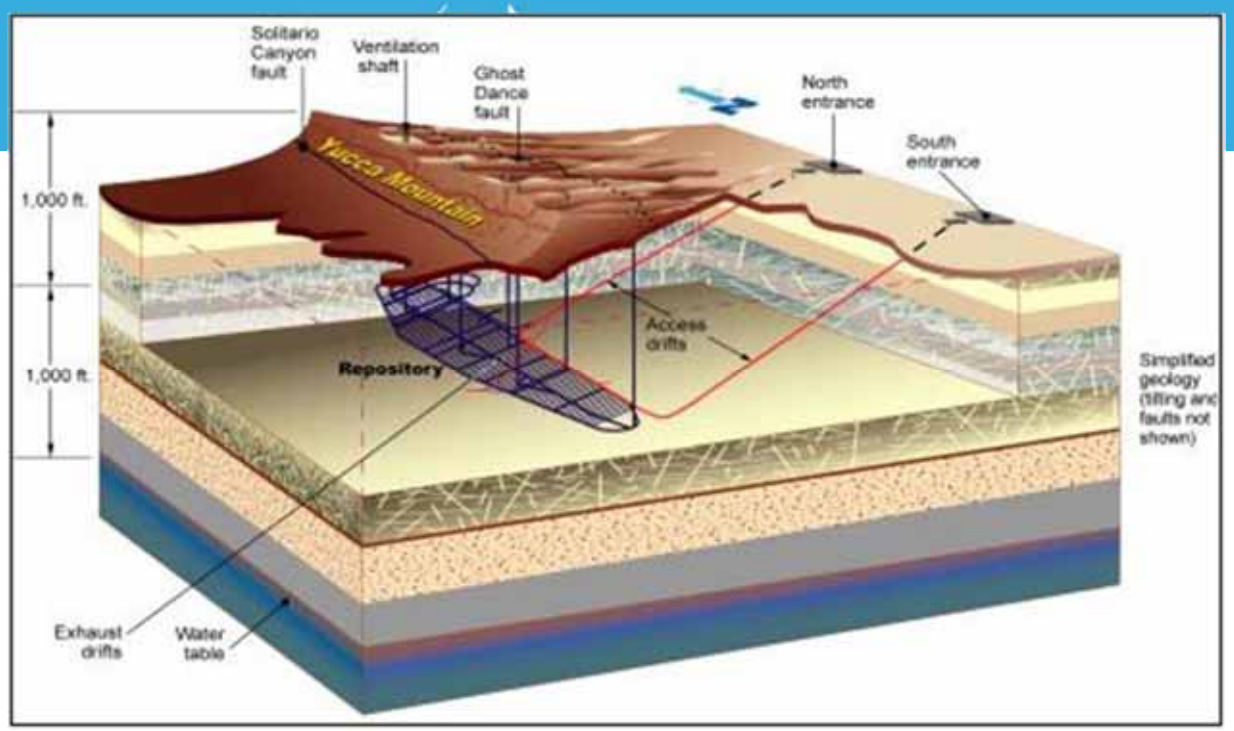
'Quantitative mathematical models used by policy makers and government administrators to form environmental policies are seriously flawed'

RULE ONE: Check against rhetorical use of mathematical modelling



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 36Cl story)

Joint Research Centre

RULE ONE: Check against rhetorical use of mathematical modelling

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.

RULE TWO: Adopt an ‘assumption hunting’ attitude;

What was ‘assumed out’? What are the tacit, pre-analytic, possibly normative assumptions underlying the analysis?

E.g. in ‘Bogus Quantification: Uses and Abuses of Models’ John Kay uncovers that the UK transport WebTAG model (the standard for transport policy simulation) needs as input ‘Annual Percentage Change in Car Occupancy up to 2036.’

John Kay, London School Economics, Columnist Financial Times

A wise man knows one thing – the limits of his knowledge, John Kay, FT November 30, 2011.

"[…] The models share a common approach. They pose the question: "How would we make our decision if we had complete knowledge of the world?" With such information you might make a detailed assessment […]. But little of this knowledge exists. So you make the missing data up. You assume the future will be like the past […]. The impression of rationality these procedures convey is spurious."

"[…] If you do not know the answer to a question, the right response is not to make a number up, but to rethink and frame an alternative question that is capable of being answered."

John Kay’s approach is called ‘Assumptions hunting’ in Dutch circles …

Contents lists available at ScienceDirect

Energy Policy

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

On the contribution of external cost calculations to energy system governance: The case of a potential large-scale nuclear accident

Erik Laes [a,*], Gaston Meskens [b], Jeroen P. van der Sluijs [c]

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

A method for the analysis of assumptions in model-based environmental assessments

Penny Kloprogge [a], Jeroen P. van der Sluijs [a,b,*], Arthur C. Petersen [c]

Joint Research Centre

Contents lists available at ScienceDirect

Energy Policy

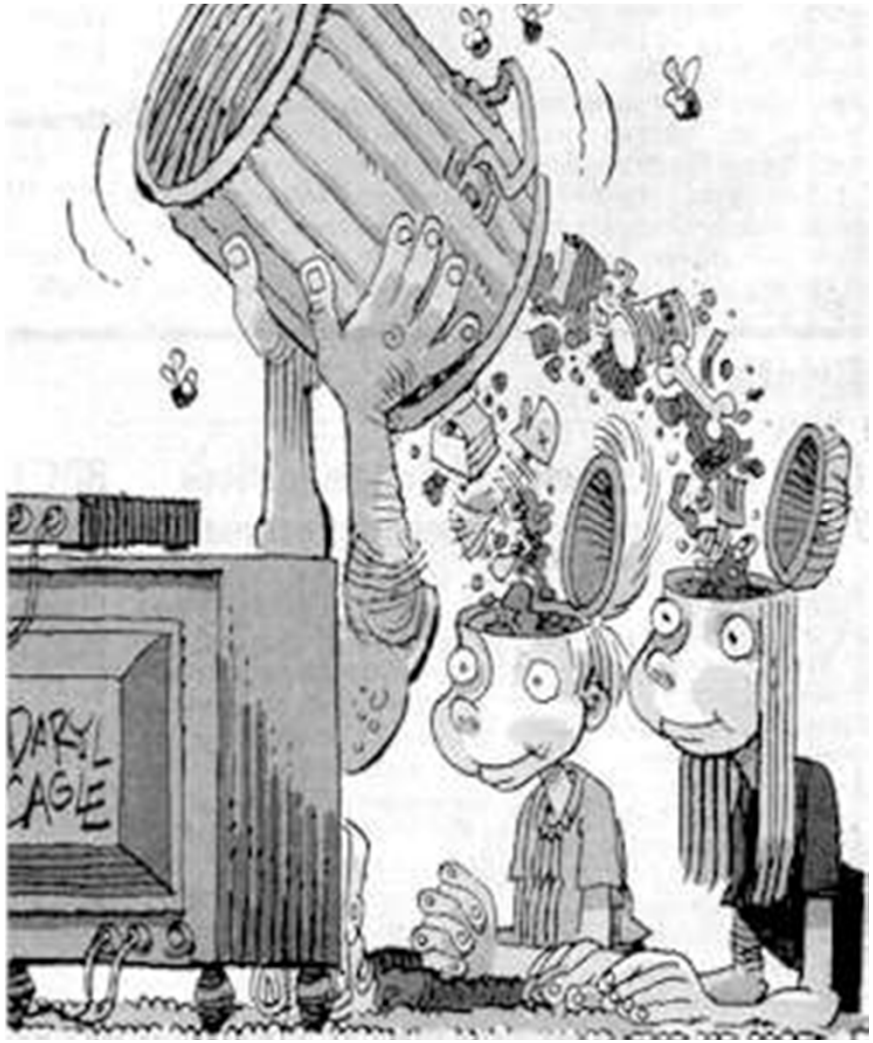
journal homepage: www.elsevier.com/locate/enpol

On the contribution of external cost calculations to energy system governance: The case of a potential large-scale nuclear accident

Erik Laes [a,*], Gaston Meskens [b], Jeroen P. van der Sluijs [c]

‘[…] calculation of the external costs of a potential large-scale nuclear accident […] ‘An [analysis] resulted in a list of 30 calculation steps and assumptions’ …

RULE THREE: detect GIGO (Garbage In, Garbage Out) Science or pseudo-science



RULE THREE: detect GIGO (Garbage In, Garbage Out) Science or pseudo-science

"where uncertainties in inputs must be suppressed lest outputs become indeterminate"

From: Uncertainty and Quality in Science for Policy
by Silvio Funtowicz and Jerry Ravetz, Springer 1990.

Edward E. Leamer, 1990, Let's Take the Con Out of Econometrics, American Economics Review, 73 (March 1983), 31-43.

<<I have proposed a form of organised sensitivity analysis that I call “global sensitivity analysis” in which a neighborhood of alternative assumptions is selected and the corresponding interval of inferences is identified.
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.>>

RULE FOUR: find sensitivities before sensitivities find you;

RULE FOUR : find sensitivities before sensitivities find you;

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 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.

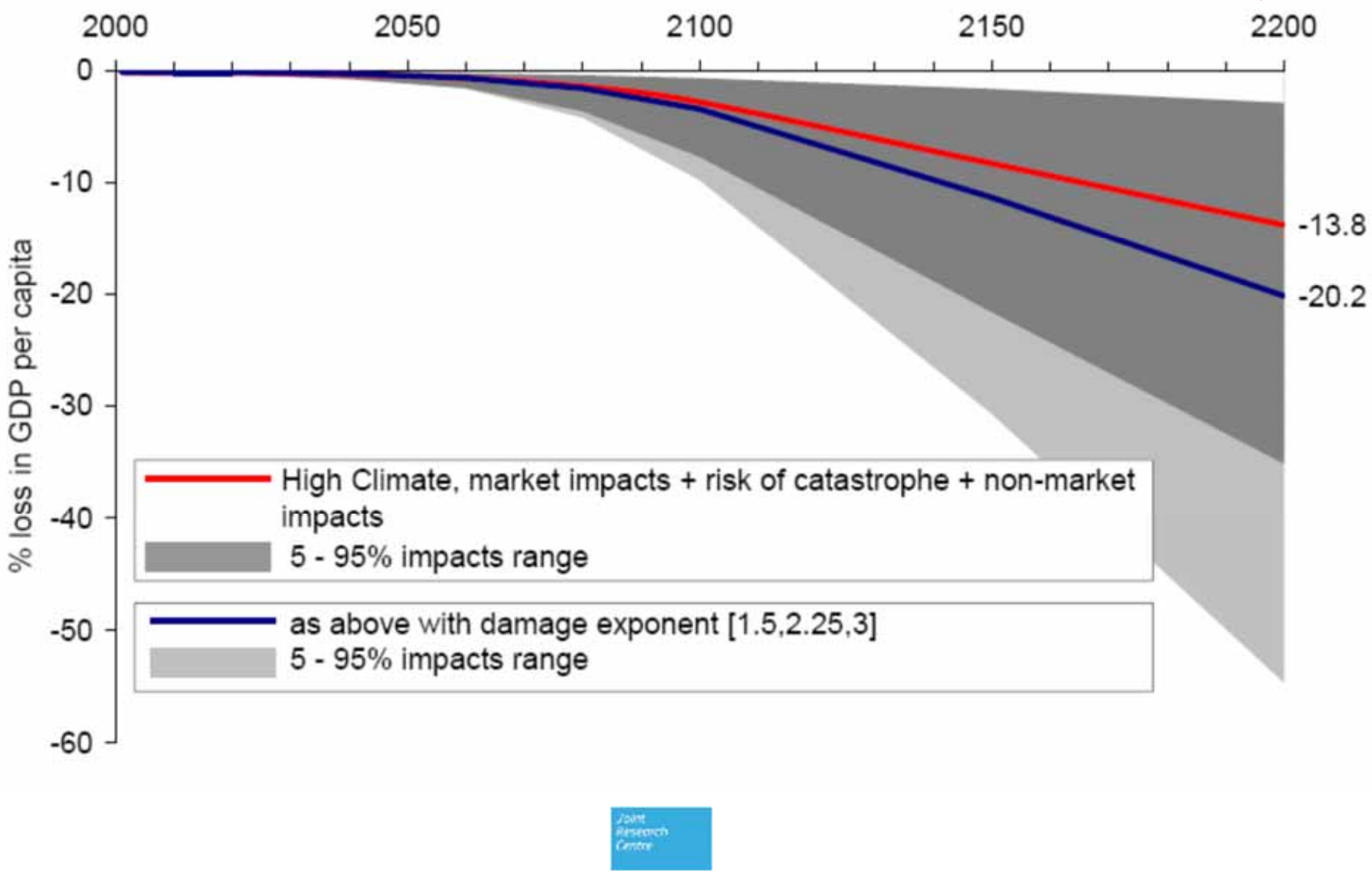
Nordhaus W., Critical Assumptions in the Stern Review on Climate Change, SCIENCE, 317, 201-202, (2007).

RULE FOUR : find sensitivities before sensitivities find you;

Stern's Review – Technical Annex to postscript (a sensitivity analysis of a cost benefit analysis)

The Stern - Nordhaus exchange on *SCIENCE*
Nordhaus → falsifies Stern based on 'wrong' range of discount rate (~ you GIGOing)
Stern → 'My analysis shows robustness'

My problems with it:

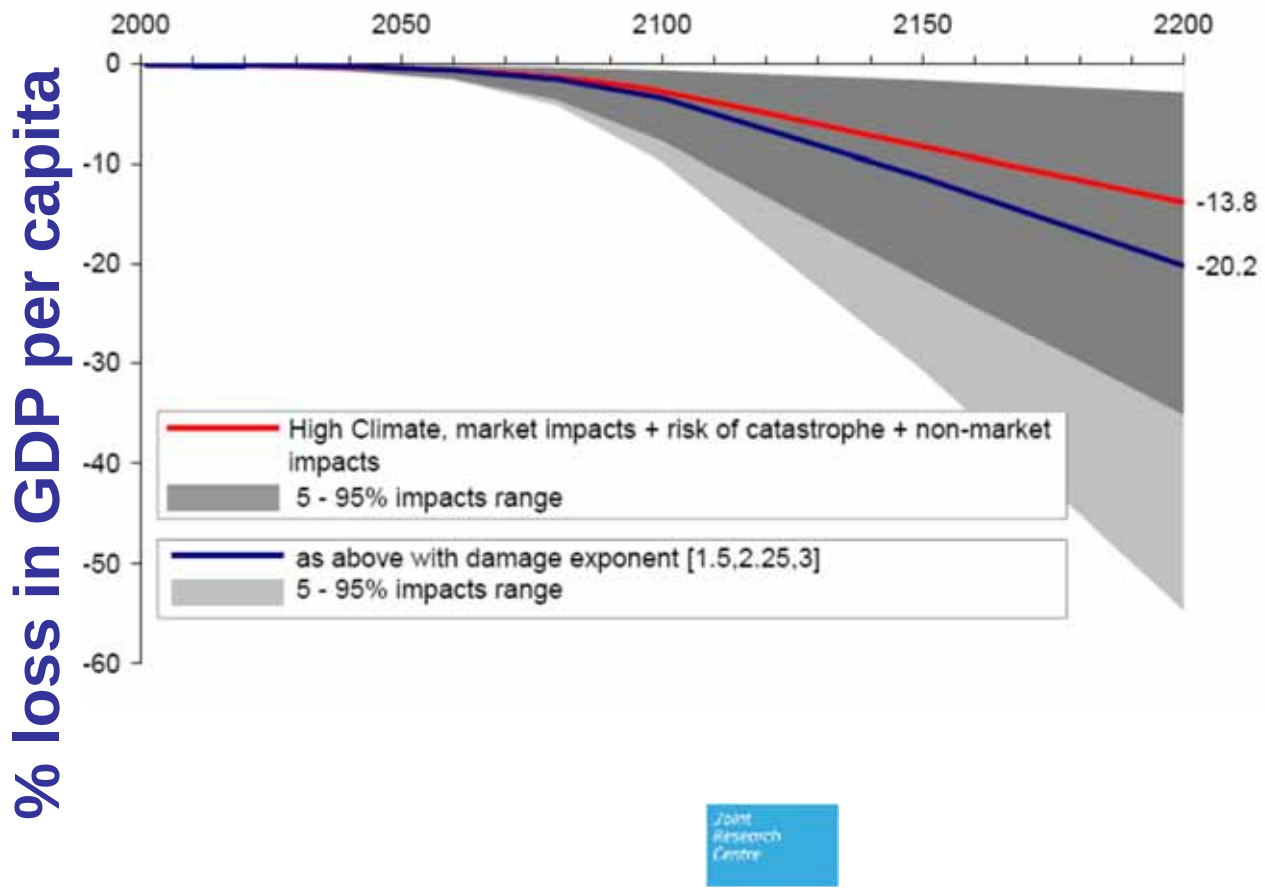


… but foremost Stern says:

changing assumptions → important effect

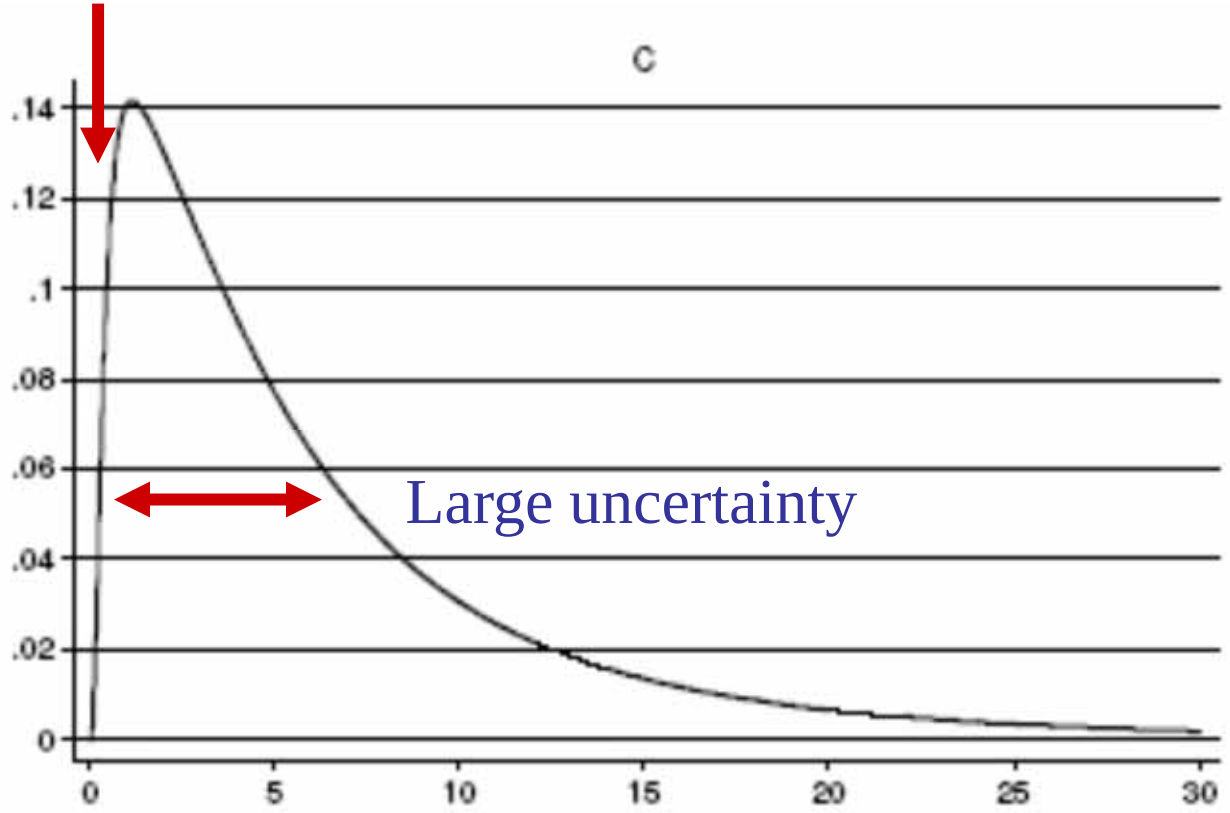
when instead he should admit that:

changing assumptions → all changes a lot



How was it done? A reverse engineering of the analysis

Missing points

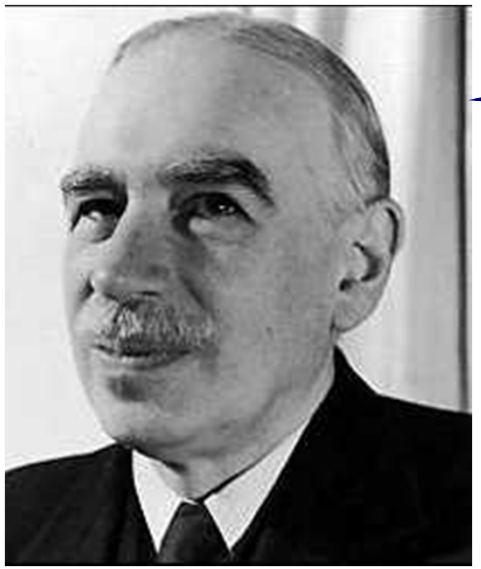


% loss in GDP per capita

RULE FOUR : find sensitivities before sensitivities find you;

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

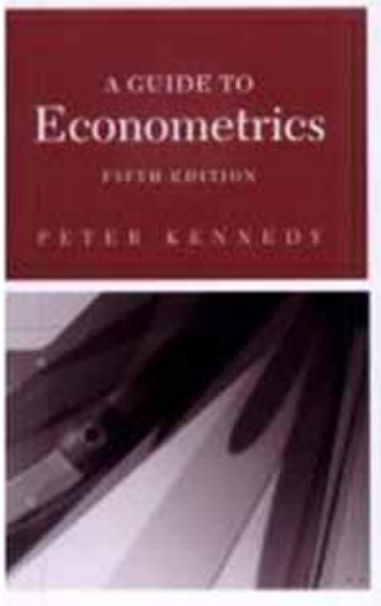
… precisely wrong



RULE FOUR : find sensitivities before sensitivities find you;

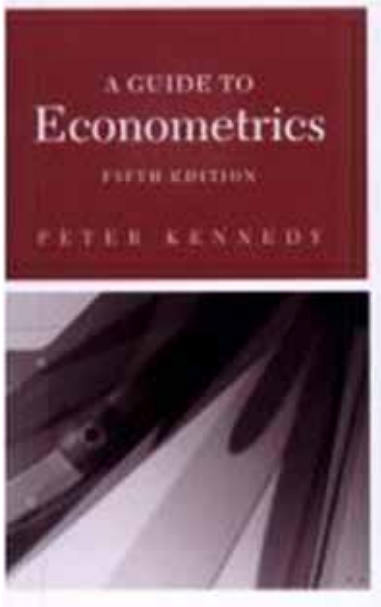
Peter Kennedy, A Guide to Econometrics. Anticipating criticism by applying sensitivity analysis. This is one of the ten commandments of applied econometrics according to Peter Kennedy:

<<Thou shall confess in the presence of sensitivity.
Corollary: Thou shall anticipate criticism >>



RULE FOUR : find sensitivities before sensitivities find you;

<<When reporting a sensitivity analysis, researchers should explain fully their specification search so that the readers can judge for themselves how the results may have been affected. This is basically an `honesty is the best policy' approach, […]'.>>



RULE FIVE: aim for transparency

European Commission

Doubts raised over Europe's green energy plan

'Host of questions' from advisers

Economic model lacks transparency

By Pilita Clark in London

The credibility of a European energy review has been cast into doubt by specialist advisers who say plans to cut carbon emissions up to 2050 are based on an economic model owned by a single Greek university that cannot be independently scrutinised.

The energy experts have "raised a host of questions" on how the European Commission's use of a non-transparent model could affect the review, according to a leaked report by advisers chosen by Brussels to comment on the "Energy Roadmap 2050" proposals.

The economic model, known as Primes, is owned by the National Technical University of Athens and is designed to show how the use of different mixes of energy sources affect the wider economy.

The European Commission has used it for many years to help guide the bloc's energy policies. But industry critics complain that its assumptions are impossible to question because the model is privately owned. One think group, Business Europe, has called for the Commission to use other, more transparent models.

The forecasting reform of the energy road map, which will review the effect of choosing gas, more wind and less nuclear power to meet Europe's green targets, has heightened concerns about the model's transparency, the expert advisory group report shows. One of the group's three meetings was "devoted largely" to how the Commission was using the Primes model to produce different energy mix scenarios for the road map.

"There was considerable debate about the role of the oil fuel price assumptions in the Primes model," said the report by the group, which is chaired by Dieter Helm, an Oxford university economics professor, and includes bodies such as the International Energy Agency.

There were also questions on "the costs of different technologies" and "the assumptions of public intervention by comparison but not by individuals".

The group's key concern was "about the transparency of the Primes work, and in particular the properties rights in the algorithms and detailed internal workings of the model", says the report, which is marked "final draft".

"The model remains the private property of the National Technical University of Athens," it says.

"The consequence is that independent parties cannot replicate the results. This is a substantial matter for the Commission, but members of the group pointed out that it does have obvious consequences for the credibility of the road map."

The advisory group recommends that the Primes model be made publicly available "so that its results can be replicated by interested parties".

Pantelis Capros, an economist from the National Technical University of Athens who built the Primes model, told the Financial Times he agreed that transparency was important and would try to find a way of the model's workings were made public. "It is not the code itself – not the software".

A spokeswoman for the energy commissioner, Günther Oettinger, said she could see concerns in an unpublished document. The final version of the advisory group's report would be released with the energy road map next month.

Prof Capros has been an energy consultant for many years and has held positions on bodies ranging from Greece's energy regulator to the country's Public Power Corporation.

> 'Independent parties cannot replicate the results because the model is private property'

A mining plant in Germany: the viability of plans to cut EU emissions has been called into question by experts

Joint Research Centre

RULE FIVE: aim for transparency

"Experts have "raised a host of questions" about how the European Commission's use of a non-transparent model could affect the energy review, according to a leaked report by energy specialists chosen by Brussels to advise on the forthcoming "Energy Roadmap to 2050"
FT November 6, 2011

RULE FIVE: aim for transparency

"The credibility of a European energy review has been cast into doubt by experts who point out that long-term plans to cut carbon emissions are based on an economic model owned by a single Greek university that cannot be independently scrutinised."

Part IX

Office of Management and Budget

Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies; Notice; Republication

The OMB about transparency

http://www.whitehouse.gov/omb/inforeg/

RULE FIVE: aim for transparency

[models should be made available to a third party so that it can] use the same data, computer model or statistical methods to replicate the analytic results reported in the original study.

[…] The more important benefit of transparency is that the public will be able to assess how much an agency's analytic result hinges on the specific analytic choices made by the agency.

Friday, February 22, 2002
Graphic - Federal Register, Part IX
Office of Management and Budget
Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies; Notice; Republication
http://www.whitehouse.gov/omb/inforeg/

This was 2002

Reproducibility

a necessary condition for

Transparency

a necessary condition for

Legitimacy

House Republicans Aim To Limit Power Of Environmental Protection Agency

This is 2014

The Huffington Post | by Robin Wilkey (/robin-wilkey)

Posted: 02/07/2014 6:18 pm EST | Updated: 02/08/2014 10:59 am EST

The bill, dubbed the Secret Science Reform Act would force the EPA to publicly release its research on a topic before issuing a policy recommendation, and require that the research be "reproducible." Supporters claim the bill will increase transparency in public policy, while opponents have accused the bill's authors of trying to “keep the EPA from doing its job.”

AUTHENTICATED
U.S. GOVERNMENT
INFORMATION
GPO

I

113TH CONGRESS
2D SESSION

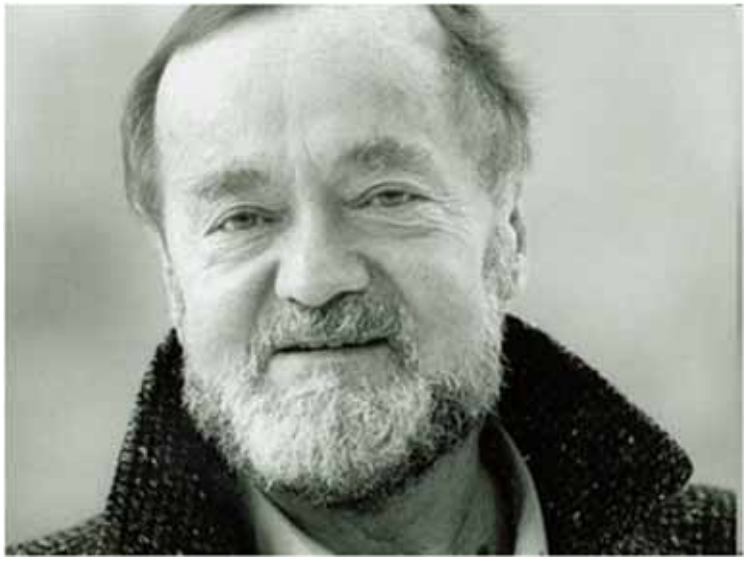
H. R. 4012

To prohibit the Environmental Protection Agency from proposing, finalizing, or disseminating regulations or assessments based upon science that is not transparent or reproducible.

http://beta.congress.gov/bill/113th-congress/house-bill/4012
Accessed May 2014

Joint Research Centre

RULE SIX: Do the right sums



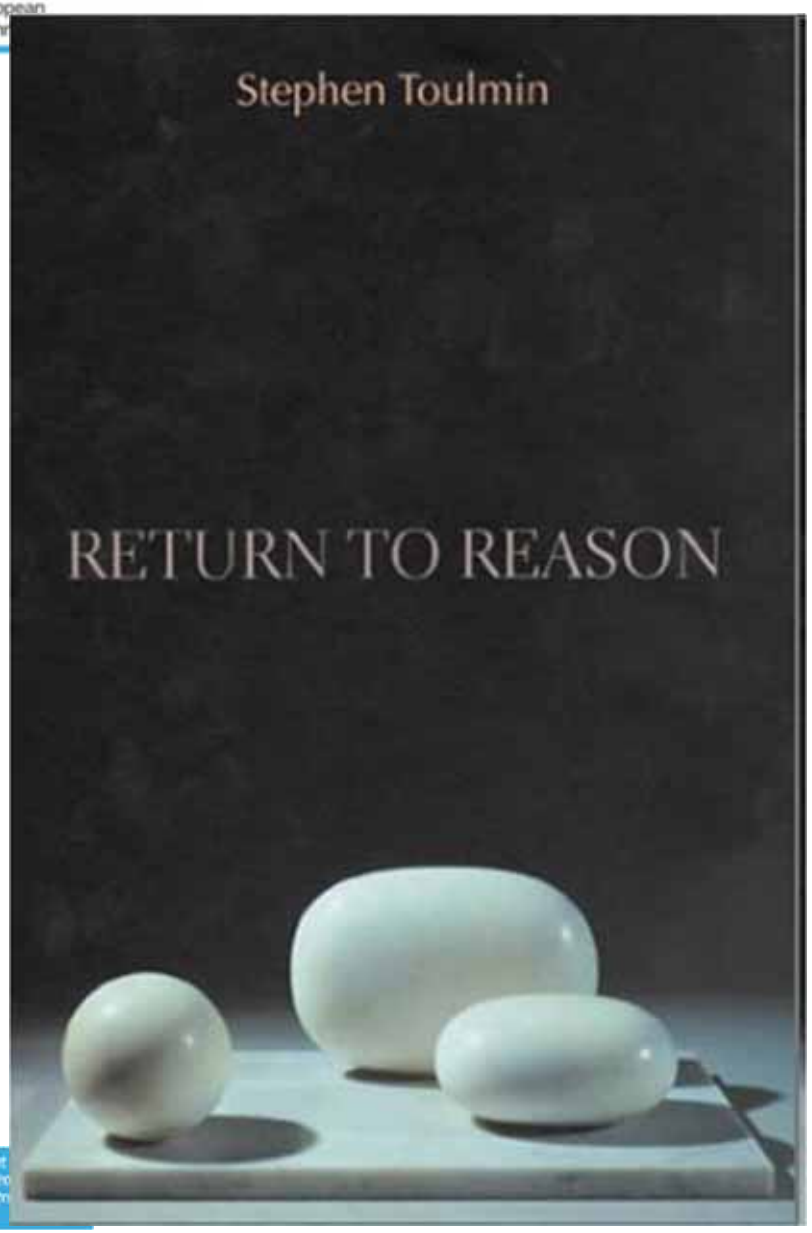
Do the sum right

Versus

Do the right sums

(*Stephen Toulmin*)

A plea for reasonableness versus rationality



RULE SIX: Do the right sums

Peter Kennedy's commandment of applied econometrics: 'Thou shall answer the right question', Kennedy 2007

Ex: GMO treated as a 'risk to health issue'

RULE SIX: Do the right sums

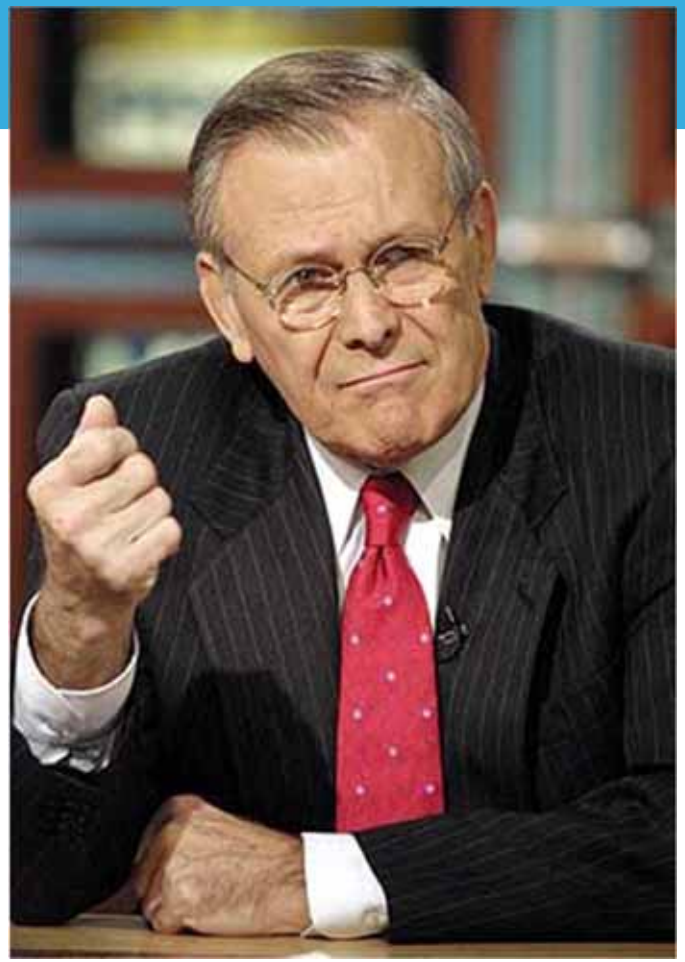
Final Report of the PABE research project funded by the Commission of European Communities, Contract number: FAIR CT98-3844 (DG12 - SSMI), December 2001

- Why do we need GMOs? What are the benefits?
- Who will benefit from their use?
- Who decided that they should be developed and how?
- Why were we not better informed about their use in our food, *before* their arrival on the market?
- Why are we not given an effective choice about whether or not to buy and consume these products?
- Do regulatory authorities have sufficient powers and resources to effectively counter-balance large companies who wish to develop these products?

…

RULE SIX: Do the right sums

The spectre of type III errors: Donald Rumsfeld version: "Reports that say that something hasn't happened are always interesting to me, because as we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns -- the ones we don't know we don't know."



RULE SEVEN: Explore diligently the space of the assumptions

Environmental Modelling & Software 25 (2010) 1508–1517

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

RULE SEVEN: Explore diligently the space of the assumptions

"The uncertainties which are more carefully scrutinised are usually those which are the least relevant" (*lampposting*, Jeroen van der Sluijs).

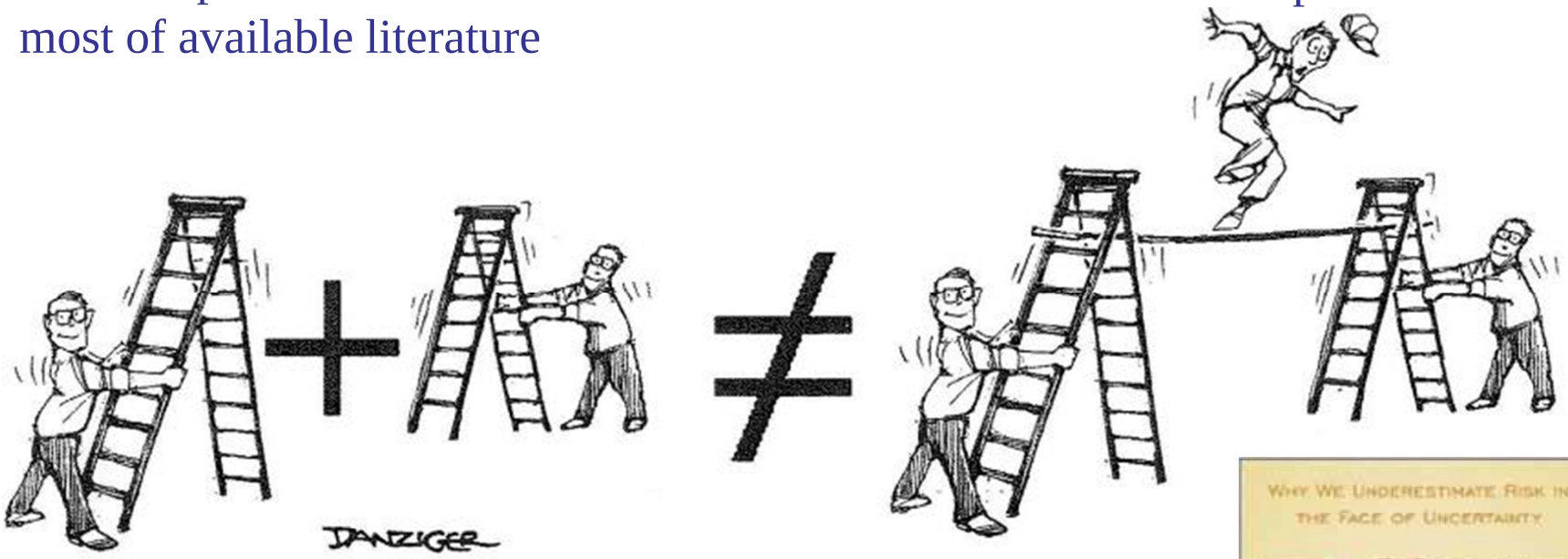
Nassim Nicholas Taleb calls this 'The delusion of uncertainty' .

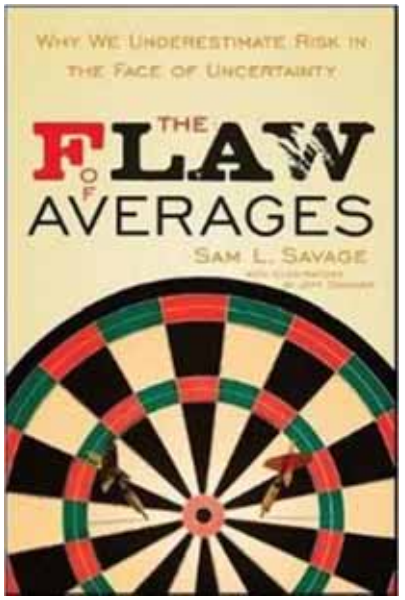
Uncertainty can be instrumentally amplified or downplayed

RULE SEVEN: Explore diligently the space of the assumptions

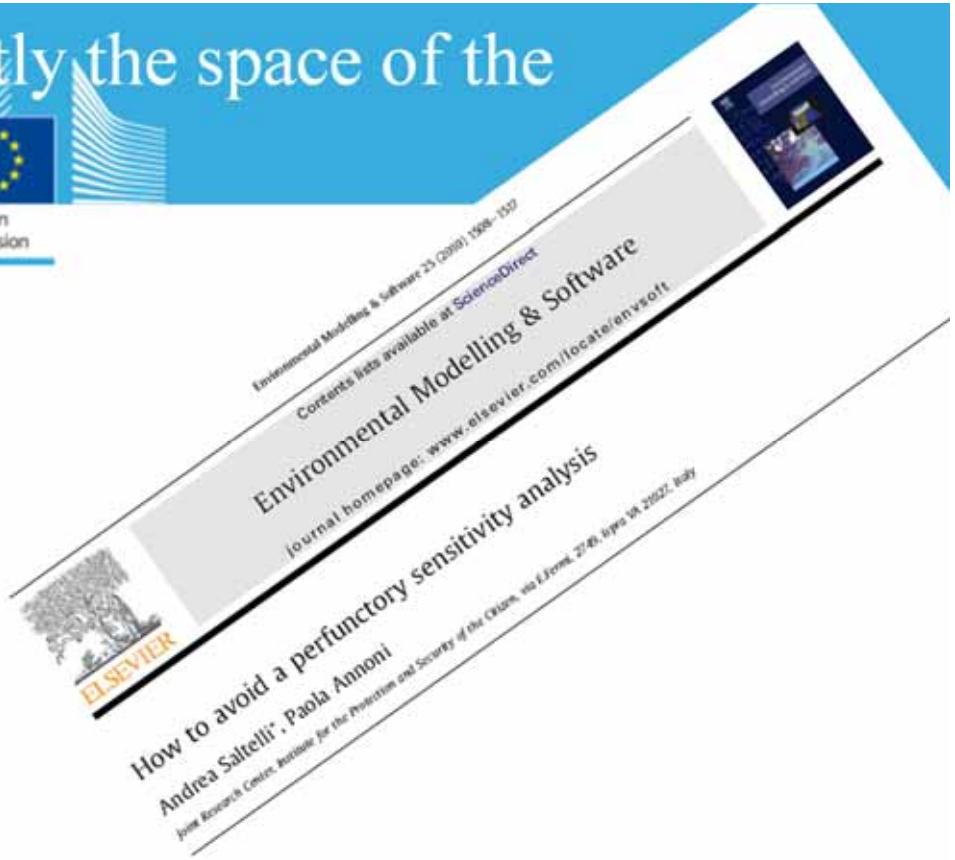
How coupled ladders are shaken in most of available literature

How to shake coupled ladders





RULE SEVEN: Explore diligently the space of the assumptions



The most popular SA practice seen in the literature is that of 'one-factor-at-a-time' (OAT). This consists of analyzing the effect of varying one model input factor at a time while keeping all other fixed.

While the shortcomings of OAT are known from the statistical literature, its widespread use among modelers raises concern on the quality of the associated sensitivity analyses

European Commission

EPILOGUE

Joint Research Centre

Examples of troubles to be anticipated :

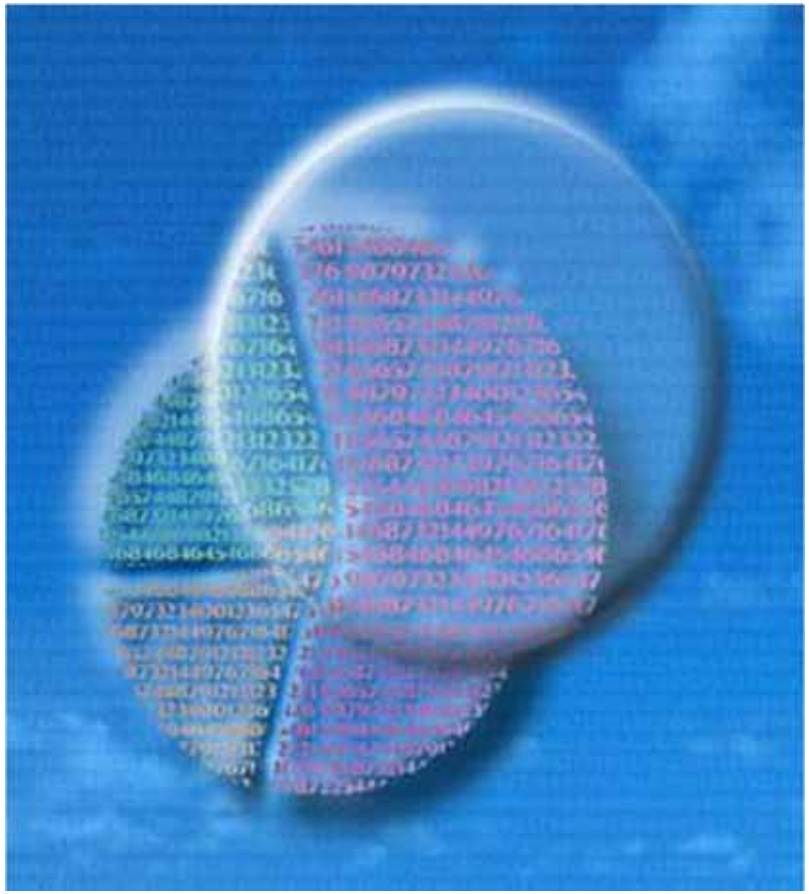
➔<<You treated X as a constant when we know it is uncertain by at least 30%>>

➔<<Beware: It would be sufficient for a 5% error in X to make your statement about Z fragile>>

➔<<Your model is but one of the plausible models – you neglected model uncertainty>>

➔<< You have maximized instrumentally your level of confidence in the results>>
➔<<Your model is a black box – why should I trust your results?>>
➔<<You have artificially inflated uncertainty>>
➔<<Your framing is not socially robust>>
➔<<You are answering the wrong question>>

All of the above can be used to defend an assessment as well as to invalidate one.



END