Methods: sensitivity auditing +

Andrea Saltelli
Centre for the Study of the Sciences and the Humanities, University of Bergen, and Open Evidence Research, Open University of Catalonia; August 27th–September 1st, 2018,
Course Numbers for policy, Castelldefels (Barcelona)
CAETERIS ARE NEVER PARIBUS

sensitivity analysis, sensitivity auditing, science for policy, impact assessment
= more material on my web site
Problematic quantifications
The myth of scientific quantification via risk or cost benefit analyses, including of the impact of new technologies, has been at the hearth of the critique of the ecological moment (e.g. Schumacher, 1973; Winner, 1986; Funtowicz and Ravetz, 1994)


[...] quality is much more difficult to 'handle' than quantity, just as the exercise of judgment is a higher function than the ability to count and calculate.

Quantitative differences can be more easily grasped and certainly more easily defined than qualitative differences: their concreteness is beguiling and gives them the appearance of scientific precision, even when this precision has been purchased by the suppression of vital differences of quality.

Frames

Most analyses offered as input to policy are framed as cost benefit analysis or risk analyses.

Funtowicz and Ravetz ➔ poor quality in science for policy ➔ post normal science

J. Ravetz and S. Funtowicz
Post–Normal Science as a reaction to cost benefit and risk analysis applied to ecological problems:

“How much is a songbird worth?”

Example: deconstruction of the economics of climate change

About a paper (Nordhaus 1991) on the economics of the greenhouse effect “since the paper displays considerable sophistication in the handling of uncertainties in data.”

“the paper by Nordhaus is liberally sprinkled with caveats...”

One such caveat is – in the words of William Nordhaus – the difficulty to move from the “terra infirma of climate change to the terra incognita of the social and economic impacts of climate change”, but:

Sensible policies on global warming should weight the costs of slowing climate change against the benefits of slower climate change. Ironically, recent policy initiatives, such as the Kyoto Protocol of 1997, have been introduced without any attempt to link the emissions controls with the benefits of the lower emissions.

— William Nordhaus
“[Although ] in his rhetoric at least, the author shows a clear awareness of the presence of the various sorts of uncertainty, […] he does not successfully manage the problems of uncertainty.”

<table>
<thead>
<tr>
<th>Sectors</th>
<th>Billions (1981 $)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Severely impacted sectors</strong></td>
<td></td>
</tr>
<tr>
<td>Farms</td>
<td></td>
</tr>
<tr>
<td>Impact of greenhouse warming and CO₂</td>
<td>−10.6 to +9.7</td>
</tr>
<tr>
<td>fertilisation</td>
<td></td>
</tr>
<tr>
<td>Forestry, fisheries, other</td>
<td>Small + or −</td>
</tr>
<tr>
<td><strong>Moderately impacted sectors</strong></td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>+</td>
</tr>
<tr>
<td>Water transportation</td>
<td>?</td>
</tr>
<tr>
<td>Energy and utilities</td>
<td></td>
</tr>
<tr>
<td>Energy (electric, gas, oil)</td>
<td></td>
</tr>
<tr>
<td>Energy demand</td>
<td>−1.65</td>
</tr>
<tr>
<td>Non-electric space heating</td>
<td>1.16</td>
</tr>
<tr>
<td>Water and sanitary</td>
<td>−?</td>
</tr>
<tr>
<td>Real estate</td>
<td></td>
</tr>
<tr>
<td>Land-rent component</td>
<td></td>
</tr>
<tr>
<td>Estimate of damage from sea-level rise</td>
<td>−1.55</td>
</tr>
<tr>
<td>Loss of land</td>
<td>−1.55</td>
</tr>
<tr>
<td>Protection of sheltered areas</td>
<td>−0.90</td>
</tr>
<tr>
<td>Protection of open coasts</td>
<td>−2.84</td>
</tr>
<tr>
<td>Hotels, lodging, recreation</td>
<td>?</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
</tr>
<tr>
<td>Central estimate</td>
<td></td>
</tr>
<tr>
<td>Billions, 1981 level of national income</td>
<td>−6.23</td>
</tr>
<tr>
<td>Percentage of national income</td>
<td>−0.26</td>
</tr>
</tbody>
</table>

Sources for Table 6: Underlying data on impacts are summarised in EPA (1988). Translation into national-income accounts by author. Details are available on request.
"The hyper-precision in the expression of the key number -0.26% [...] shows that this is one of those 'magic numbers' designed to produce confidence in the existence of a hard core of objective fact deep inside the mass of intuitive fuzz."

For Nordhaus - based on a 'hunch' this -0.26% could become -2% …

### Table 1

<table>
<thead>
<tr>
<th>Sectors</th>
<th>Billions (1981 $)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Severely impacted sectors</strong></td>
<td></td>
</tr>
<tr>
<td>Farms</td>
<td></td>
</tr>
<tr>
<td>Impact of greenhouse warming and CO₂</td>
<td>-10.6 to +9.7</td>
</tr>
<tr>
<td>fertilisation</td>
<td></td>
</tr>
<tr>
<td>Forestry, fisheries, other</td>
<td>Small + or -</td>
</tr>
<tr>
<td><strong>Moderately impacted sectors</strong></td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td>+</td>
</tr>
<tr>
<td>Water transportation</td>
<td>?</td>
</tr>
<tr>
<td>Energy and utilities</td>
<td></td>
</tr>
<tr>
<td>Energy (electric, gas, oil)</td>
<td></td>
</tr>
<tr>
<td>Energy demand</td>
<td>-1.65</td>
</tr>
<tr>
<td>Non-electric space heating</td>
<td>1.16</td>
</tr>
<tr>
<td>Water and sanitary</td>
<td>?</td>
</tr>
<tr>
<td><strong>Real estate</strong></td>
<td></td>
</tr>
<tr>
<td>Land-rent component</td>
<td></td>
</tr>
<tr>
<td>Estimate of damage from sea-level rise</td>
<td></td>
</tr>
<tr>
<td>Loss of land</td>
<td>-1.55</td>
</tr>
<tr>
<td>Protection of sheltered areas</td>
<td>-0.90</td>
</tr>
<tr>
<td>Protection of open coasts</td>
<td>-2.84</td>
</tr>
<tr>
<td><strong>Hotels, lodging, recreation</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
</tr>
<tr>
<td>Central estimate</td>
<td></td>
</tr>
<tr>
<td>Billions, 1981 level of national income</td>
<td>-6.23</td>
</tr>
<tr>
<td>Percentage of national income</td>
<td>-0.26</td>
</tr>
</tbody>
</table>

Sources for Table 6: Underlying data on impacts are summarised in EPA (1988). Translation into national-income accounts by author. Details are available on request.
A more recent paper:

Climate Models as Economic Guides
Scientific Challenge or Quixotic Quest?

The uncertainties associated with mathematical models that assess the costs and benefits of climate change policy options are unknowable. Such models can be valuable guides to scientific inquiry, but they should not be used to guide climate policy decisions.
An audacious study:
“[…] the report forecasts—at the level of individual counties in the U.S.—energy costs and demand, labor supply, mortality, violent crime rates, and real estate property prices up to the year 2100 […]”
“The report presents the amount of computer power and data generated as evidence of the scientific legitimacy of the enterprise.

The authors note, however, that out of an abundance of caution they did not model deterioration in cognitive performance as temperatures rise.”
Advocates for better integrated assessment models (IAM)
After a list of criticism moved to the realism of Integrated Assessment Models:

“[…] the point is that estimates based on these models are very sensitive to assumptions and are likely to lead to gross underestimation” p.139
Things to be incorporated in ‘formal modelling’

“Damage to social, organizational or environmental capital […]
Damage to stock of capitals and land […]
Damage to overall factor productivity […]
Damage to learning and endogenous growth”, p. 145

‘formal modelling’ as to produce ‘numbers’?
N. Stern suggests using different mathematical models, including dynamic stochastic general equilibrium models (DSGE).

Philip Mirowski: a critique of DSGE as used in economics; inquiries by the US senate and the Queen of the England about their failure to predict the crisis.
Everybody in the profession knows that DSGE work under the economists' standard 'caeteris paribus' hypothesis (=all the rest being equal)
But

Caeteris are never paribus
Sensitivity auditing
EC impact assessment guidelines: what do they say about sensitivity auditing?

… where there is a major disagreement among stakeholders about the nature of the problem, … then sensitivity auditing is more suitable but sensitivity analysis is still advisable as one of the steps of sensitivity auditing.
Sensitivity auditing, […] is a wider consideration of the effect of all types of uncertainty, including structural assumptions embedded in the model, and subjective decisions taken in the framing of the problem.

[…]
The ultimate aim is to communicate openly and honestly the extent to which particular models can be used to support policy decisions and what their limitations are.
“In general sensitivity auditing stresses the idea of honestly communicating the extent to which model results can be trusted, taking into account as much as possible all forms of potential uncertainty, and to anticipate criticism by third parties.”
The rules of sensitivity auditing

Rule 1: Check against rhetorical use of mathematical modelling;

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

Rule 3: Check if uncertainty been instrumentally inflated or deflated.
The rules of sensitivity auditing

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

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

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

Rule 7: Perform a proper global sensitivity analysis.
The rules of sensitivity auditing can be used as columns for NUSAP pedigree matrix

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

Jeroen van der Sluijs

http://www.nusap.net/
Some examples:
Sensitivity auditing: the OECD PISA study
Do PISA data justify PISA-based education policy?
International PISA tests show how evidence-based policy can go wrong

A condensed version of the article
With PISA the OECD gained the centre-stage in the international arena on education policies, which led to important controversies.

http://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics
Critical remarks by the 80 signatories of the letter:

- Flattening of curricula (exclusion of subjects)
- Short-termism (teaching to the test)
- Promoting “life skills to function in knowledge societies”
- Stressing the student
- … ➔ Stop the test!
- A more participatory run of the study would be advisable
Figure 1

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

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

http://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/thehighcostofloweducation/
PISA’s daring quantifications:

“If every EU Member State achieved an improvement of 25 points in its PISA score (which is what for example Germany and Poland achieved over the last decade), the GDP of the whole EU would increase by between 4% and 6% by 2090; such an 6% increase would correspond to 35 trillion Euro”

Our study identifies both technical and normative issues:

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

2) Non open data, which makes SA impossible
Our study identifies both technical and normative issues:

3) Flattening curricula (do all countries wish to prosper by becoming knowledge societies?)

4) Power implications: power in the use of evidence. OECD (unelected officers and scholars) becoming a global super-ministry of education
Some examples:
Sensitivity auditing/Quantitative storytelling: Golden Rice’s story
While Greenpeace and other organizations oppose genetically engineered food, more than 100 Nobel laureates are taking a stand on the side of GMOs. Here's a look at each side's arguments. (Jenny Starrs/The Washington Post)
From the Noble laureates’ letter:

“Greenpeace has spearheaded opposition to Golden Rice, which has the potential to reduce or eliminate much of the death and disease caused by a vitamin A deficiency (VAD), which has the greatest impact on the poorest people in Africa and Southeast Asia.

[...] a total of one to two million preventable deaths occur annually as a result of VAD, [...] VAD itself is the leading cause of childhood blindness globally affecting 250,000 – 500,000 children each year. Half die within 12 months of losing their eyesight.”
From the Noble laureates’ letter:

“[…] Opposition based on emotion and dogma contradicted by data must be stopped.

How many poor people in the world must die before we consider this a "crime against humanity"?”
Opposing evidence on Golden Rice

Nutritionally: not enough beta carotene

Golden rice not authorized yet

More politically viable alternative successful

Dangerous colour

Low yield of the modified variety …

“What climate, vaccines and GMOs have in common”

https://theconversation.com/forcing-consensus-is-bad-for-science-and-society-77079
Some examples:
Sensitivity analysis: the case of the Stern review
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

The Stern – Nordhaus exchange on *SCIENCE*

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

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

3) Stern infers: My analysis shows robustness’
My problems with it:
but foremost Stern says:
changing assumptions $\rightarrow$ important effect
when instead he should admit that:
changing assumptions $\rightarrow$ all changes a lot

% loss in GDP per capita
How was it done? A reverse engineering of the analysis

Missing points

Large uncertainty

% loss in GDP per capita
Sensitivity analysis, also by reverse engineering

(A) delta

(B) eta

(C) scenario

(D) market

(E) gamma

Sensitivity analysis, also by reverse engineering
Same criticism applies to Nordhaus – both authors frame the debate around numbers which are …

… precisely wrong

Quantitative story-telling
“There is only a perspective seeing, only a perspective “knowing”; and the more affects we allow to speak about one thing, the more eyes, different eyes, we can use to observe one thing, the more complete will our “concept” of this thing, our “objectivity”, be.”
Frames: The expression ‘tax relief’ is apparently innocuous but it suggests that tax is a burden, as opposed to what pays for road, hospitals, education and other infrastructures of modern life (Lakoff, 2004).


Lakoff, G., 2004–2014, Don’t think of an elephant: know your values and frame the debate, Chelsea Green Publishing.
Frames

For Akerlof and Shiller – against what the ‘invisible hand’ would contend – economic actors have no choice but to exploit frames to ‘phish’ people into practices which benefit the actors not the subject phished.
QST tests frames/narratives for:

- Misconstruction, internal contradictions, technical errors
- Feasibility (compatibility with processes outside human control);
- Viability (compatibility with processes under human control, in relation to both the economic and technical dimensions); and
- Desirability (compatibility with a multitude of normative considerations relevant to a plurality of actors).
Frames as hypocognition &
Socially constructed ignorance
For Rayner (2012) “Sense-making is possible only through processes of exclusion. Storytelling is possible only because of the mass of detail that we leave out. Knowledge is possible only through the systematic ‘social construction of ignorance’ (Ravetz, 1986)”

Rayner’s (2012) strategies to deal with “uncomfortable knowledge”.

1. Denial: “There isn’t a problem”

2. Dismissal: “It’s a minor problem”

Rayner’s (2012) strategies to deal with “uncomfortable knowledge”.

3. Diversion: “Yes I am working on it” (In fact I am working on something that is only apparently related to the problem)

Rayner’s (2012) strategies to deal with “uncomfortable knowledge”.

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

“Uncomfortable knowledge” can be used as a gauge of an institution’s health.

The larger the “uncomfortable knowledge” an institution needs to maintain, the closer it is to its ancient régime stage (Funtowicz and Ravetz, 1994).

Why frames ‘stick’

“If is difficult to get a man to understand something when his salary depends upon his not understanding it.”

Upton Sinclair
Some examples:
Sensitivity auditing/Quantitative storytelling: The Ecological Footprint
Footprints to nowhere

Mario Giampietro a,c, Andrea Saltelli b,*

a Institute of Environmental Science and Technology (ICTA), Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
b Institute for the Protection and Security of the Citizen (IPSC), The European Commission, Joint Research Centre, TP 361, 21027 Ispra, VA, Italy
c Catalan Institution for Research and Advanced Studies (ICREA), Passeig Lluís Companys, 23, 08010 Barcelona, Spain


Alessandro Galli, Mario Giampietro, Steve Goldfinger, Elias Lazarus, David Lin, Andrea Saltelli, Matthias Wackernagel, Felix Müller, 2016, Questioning the ecological footprint, Ecological Indicators, 69, 224–232.
How many Chinas does it take to support China?

- **CHINA**: 2.5

What about some other countries?

- **FRANCE**: 1.6
- **INDIA**: 1.8
- **U.S.A.**: 1.9
- **EGYPT**: 2.4
- **GREECE**: 3.1
- **U.K.**: 3.5
- **ITALY**: 4.0
- **SWITZERLAND**: 4.2
- **QATAR**: 5.7
- **JAPAN**: 7.1

- **WORLD**: 1.5
Based on two “accounts (biocapacity and footprint) representing the supply and demand of renewable biological resources, and the area of forest required to offset human carbon emissions (the carbon footprint)” the EF tells mankind how many planets are being used.
The change of world footprint in time (1961-2006)

In this period:
- world population doubled
- economic activity increased 6-fold
- food production increased 2 ½ times

... but on the supply side: nothing changed!

The only measured change in EF is generated by a very creative protocol converting energy use into hectares of planet . . .
The footprint is almost entirely driven by energy consumption, which corresponds to carbon emission which are in turn sequestrated by forests; [...] Carbon sequestration rate is hence what drives the results.

But this number could be made negative as well as infinity depending on what number one picks ... it is totally volatile.
Is the EF a rhetorical device?

- The implausible accuracy (Earth overshoot day = August 2!)
- Offsetting a flow with a stock (Kg of CO2 per year versus square meters of land)
- The anti-trade bias (CMEPSP, 2009, p. 71)
- The total dependence upon energy related pressures
- Paradoxical policy implications (e.g. in Agriculture)

Is the EF a rhetorical device?

- The EF is inconsistent with its stated purpose of measuring demand on ecosystems
- The EF depends mostly from a dimensionally flawed energy emissions assessment
- One cannot accept EF’s flaws on the ground that the EF has normative virtues; EF’s rhetoric muddles the sustainability debate
“EF measurements, as currently constructed and presented, are so misleading as to preclude their use in any serious science or policy context. [...] less than half the area of the United States planted with eucalypts could essentially give us an EF equal to one Earth—an approach that no ecologist would recommend.”


See also follow up:


Some examples:
Quantitative storytelling: Cost Benefit Analyses
The myth of scientific quantification via risk or cost benefit analyses, including of the impact of new technologies, has been at the hearth of the critique of the ecological moment (e.g. Schumacher, 1973; Winner, 1986; Funtowicz and Ravetz, 1994)

Consume GMO because they are safe
GMO as a food scare

The Economist, Vermont v science, The little state that could kneecap the biotech industry, May 10th 2014
Citizens’ worries (Marris, 2001, excerpts)

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

US National Academy of Sciences report on genetically engineered crops:

“Products of new technologies should be regulated not only on the basis of their benefit-risk profiles, but also on their societal context and need”

As noted in the field of economics, mathematization is cyclical.

Reinert, Erik S., 2000, Full circle: economics from scholasticism through innovation and back into mathematical scholasticism Reflections on a 1769 Price essay: “Why is it that economics so far has gained so few advantages from physics and mathematics?”, Journal of Economic Studies 27,4/5, 364 –376.
‘Decisionism’ was high after WW2 (RAND corporation, linear programming, decision analysis), then the ecological critique of the 70’s; then Milton Friedman and the neoliberals brought back faith e.g. in econometrics/counterfactual analysis, today in a new crisis⋯
The End

@andreasaltelli